Ireland's pre-1940 daily precipitation data: data rescue, quality assurance and analysis of extremes

Ciara Ryan BA, MSc

Thesis Submitted for the Degree of Doctor of Philosophy

Irish Climate Analysis and Research Units (ICARUS), Department of Geography, Maynooth University, National University of Ireland

February 2021

Head of Department:

Professor Gerry Kearns

Research Supervisor:

Professor Conor Murphy

Table of Contents

Research Funding

The research presented in this thesis has been made possible through the financial support of the Irish Research Council Employment Based Scholarship and the Maynooth University John and Pat Hume Scholarship programme. Additional funding was provided by the Maynooth University Research Incentivisation Scheme (RIS) and the Irish Climate Analysis and Research UnitS (ICARUS) for the purchase of the Metis EDS professional digital scanner.

The aim of the IRC Employment Based Scholarship is to embed the scholar in a professional research and innovation environment for training and development whilst simultaneously providing postgraduate education leading to a Doctoral Degree; and to facilitate research collaboration, knowledge transfer and networking between scholars and the host HEIs and Irish-based employers.

As part of the scheme, I was employed by Met Éireann, the Irish National Meteorological Service for the duration of the award (3-year fixed term) with 50% of my time based with the employment partner. Upon completion of my scholarship, I applied for and was offered a permanent position with Met Éireann. In my role as a Research Meteorologist in the Climate, Research and Applications Division (CSRAD) I am continuing my research on data rescue activities and the development of long-term climate series, including the development of long-term gridded precipitation datasets for Ireland.

Abstract

Over much of the globe, the temporal extent of meteorological records is limited, yet a wealth of data remains in paper or image form in numerous archives. This research presents daily rainfall data and metadata for Ireland transcribed from historical manuscripts and printed copies of rainfall registers located in Met Éireann's archives. To facilitate the transcription of rainfall observations from paper records, the historical manuscripts were scanned and integrated into Met Éireann's digital archives. The transcription from digital image to data format was undertaken in collaboration with students at Maynooth University as part of a novel crowdsourcing initiative to integrate data rescue activities into the classroom. In total, 3,616 station years of rainfall data (~1.32 million daily values) and associated metadata for the period 1864-1940 were transcribed and made available.

Utilising the rescued data to extend current station records, this research provides the first long-term assessment of changes in extreme precipitation observed for 30 daily station series across Ireland. Quality control of rescued data was carried out before selected long-term stations were tested for homogeneity using RHtests software. Eleven extreme precipitation indicators were calculated on an annual basis for two fixed periods, 1910-2019 and 1940-2019, and analysed to determine spatial and temporal trends in the frequency, intensity and magnitude of observed precipitation. The persistence of trends for the full record length and different periods of analysis was assessed for all stations and indicators. Results show an overall tendency towards increasing trends, with increases in precipitation extremes and intensity especially notable in the east and southeast of the island. The findings suggest that the contribution of heavy and extreme precipitation events to annual totals are increasing, despite the lack of a robust trend in annual totals.

The dataset produced is fully traceable and will have a long legacy of impact for understanding historical extremes, climate variability and change in Ireland, while the procedures for data rescue are already inspiring similar approaches in other University settings. Moreover, the work provides a relatively rare attempt to homogenise daily precipitation data, with a novel study design implemented to ensure the robustness of results. While data rescue activities can be slow and painstaking, the dataset produced resurrects the lifetimes work of many hundreds of weather observers across Ireland and sheds light on the unfolding changes in Irish precipitation.

Acknowledgements

Many people have supported, challenged and inspired me throughout the course of my degree. First is my primary supervisor Professor Conor Murphy who introduced me to the subject of historical climatology for which I am very grateful. I would like to thank Conor for his meticulous readings of many drafts submitted, for his dedication and time, and for his constant encouragement, motivation and guidance. I would not be where I am now without his support. Thank you also to my Employment Mentor, Mary Curley, whose interest and support made it possible for me to undertake this study. Mary's knowledge and expertise in the field of climatology helped me to achieve the objectives of this research.

I acknowledge with gratitude the Irish Research Council Employment Based Scholarship and the Maynooth University John and Pat Hume Scholarship programme for the funding that made this study possible. Thanks also to Maynooth University and the Irish Climate Analysis and Research UnitS (ICARUS) for funding provided through the Research Incentivisation Scheme (RIS) to purchase the Metis EDS professional digital scanner which was fundamental to the development of this research and ongoing data rescue efforts.

Thanks to Met Éireann for their role as Employment Partner and for supporting my research through the IRC Employment Based Programme. I am thankful to my colleagues in the Climate Services, Research and Applications Division for welcoming me to the organisation and supporting me in the completion of this thesis. In particular, I would like to thank the following people: Eoin Moran, Séamus Walsh, Dr. Saji Varghese, Keith Lambkin, Dr. Patrick Fournet, Tony O'Neill and Conor Lally.

Thanks to Maynooth University's Department of Geography and my colleagues in the Irish Climate Analysis and Research Units (ICARUS), in particular, Professor Peter Thorne for making the department and ICARUS such a supportive environment for postgraduates. I would like to express my gratitude to and commend Maynooth University students who participated in this study as part of the GY313 Climate Change course. To my friends in ICARUS, past and present, thank you for making the PhD experience enjoyable.

To my close friends and family, I would like to say a big thank you for the numerous ways in which you helped, supported and tolerated me throughout this experience. A special thank you to my son Alex for his patience, love and much needed hugs. Finally, I would like to dedicate this work to my Mam, Dr. Denise Ryan who inspired me to follow in her footsteps and return to education as a "mature" student.

List of Figures

List of Tables

1. Introduction

1.1 Background

With increasing concerns about the impacts of climate change, understanding changes in the characteristics of precipitation is increasingly important given that climate extremes often exhibit different behaviours than changes in average conditions (Seneviratne et al., 2012). Changes in the frequency, magnitude and intensity of precipitation extremes have far-reaching impacts on society and infrastructure, water resources, industry, and the management of flood hazards. The UN Intergovernmental Panel on Climate Change (IPCC) Fifth Annual Assessment Report (AR5) states that over most of the mid-latitude land masses, extreme precipitation events are very likely to be more intense and more frequent in a warmer world (IPCC, 2014). The relationship between temperature and precipitation extremes, governed by the Clausius-Clapeyron (C-C) relation, is used to examine projected changes in precipitation extremes (e.g. Blenkinsop et al., 2015). Observational evidence and climate models suggest that intensification of precipitation extremes is expected to occur at a rate of ∼ 6–7 % per 1°C warming given that other atmospheric conditions remain constant (Trenberth et al., 2003; Pall et al., 2007; Wang et al., 2017; Ali et al., 2018). Subsequent research suggests an accelerated rate of change for short-duration extreme precipitation events (Lenderink & van Meijgaard, 2008; Kendon et al., 2014).

Research into precipitation extremes has progressed significantly since the establishment of the World Meteorological Organization (WMO) Expert Team on Climate Change Detection and Indices (ETCCDI) (Zwiers et al., 2012). The approach proposed by the ETCCDI in 2002 facilitates the objective measurement of extreme events through the provision of standardised software and motivates analyses of climate extremes at the global scale (Peterson and Manton, 2008). The characteristics of precipitation (e.g. frequency, magnitude, intensity) can be assessed using various metrics. The ETCCDI define 13 precipitation metrics including; percentile-based indices such as annual contribution from very wet days (R95pTOT), and extremely wet days (R99pTOT) which represent the magnitude of precipitation falling above the 95th/99th percentiles respectively; absolute indices which represent the magnitude of the maximum precipitation values within a given year (e.g. maximum 1-day precipitation amount

(RX1day) and maximum 5-day precipitation amount (RX5day); threshold indices which are defined as the number of days in which a precipitation value falls above a fixed threshold including the number of heavy precipitation days > 10 mm (R10mm) and number of very heavy precipitation days > 20 mm (R20mm); duration indices which define periods of excessive wetness or dryness, such as longest consecutive spell of wet days (CWD) and longest consecutive spell of dry days (CDD); and other indices which include annual precipitation total (PRCPTOT) and the simple daily intensity index (SDII). The development of these standard climate indices has allowed derived products to be more widely shared than the underlying measurements themselves (Alexander et al., 2019; Thorne et al., 2018).

Despite uncertainties in total precipitation changes due to the combined effect of anthropogenic GHG-related global warming and natural decadal variations (Trenberth et al., 2011; Gu and Adler, 2015), detectable increases in precipitation extremes have been reported (Alexander et al., 2006; Min et al., 2011; Westra et al., 2013). Observational data indicate that, globally, there are more regions showing significant increasing trends in extreme precipitation amounts, intensity, and frequency than regions with decreasing trends, although there is widespread variability in spatial patterns (Westra et al., 2013; Donat et al., 2016; Dunn et al., 2020) and over the period of record analysed (Papalexiou & Montanari, 2019; Sun et al., 2020).

Previous studies (e.g. Alexander et al., 2006; Donat et al., 2013; 2016) have utilised gridded data to examine changes in precipitation extremes with results indicating expected changes in Asia, Europe and regions of the United States. Donat et al. (2016) investigated global changes in PRCPTOT and Rx1day indices across wet and dry regions using the HadEX2 observational data set. The study detected statistically significant ($p \leq 0.05$) increases in both precipitation indices in the dry regions (i.e., central and northeast Asia, central Australia, northwestern North America, as well as north and southwestern Africa) and in the Rx1day index in the wet regions (i.e., Southeast Asia, India, eastern South America that includes southern Brazil, Uruguay and eastern Argentina, the southeastern United States, Europe, as well as small regions in northern tropical and eastern coastal Australia, eastern tropical Africa and southeastern Africa), with slope values of 1–2% per decade.

More recently, Dunn et al. (2020) used an updated version of the HadEX dataset (HadEX3) derived from daily, in situ observations at 17,000 locations over the period 1901–2018 to assess changes in 12 ETCCDI precipitation indices. Overall, the results show that globally there are more heavy precipitation events that are also more intense and contribute a greater fraction to the total precipitation measured. Very little change was observed in the occurrence of heavy precipitation days (R10mm, Figure 1.1) over the entire period of the data set, though there are indications of a slight increase in the most recent decades and notable regional differences. A relatively uniform increase was shown in the contribution from very wet days (days exceeding the 95th percentile of daily precipitation, R95pTOT, Figure 1.2), with an extra 1–2% of precipitation falling during very wet days. Moreover, the study observed a significant increase in R95pTOT from the 1970's. Consistent with previous findings (e.g. Westra et al., 2013), the study reports strong increases in maximum 1-day precipitation, (Rx1day, Figure 1.3), particularly in the eastern half of North America, as well as the eastern parts of southern South America, parts of India, and China, with smaller increases observed over Europe.

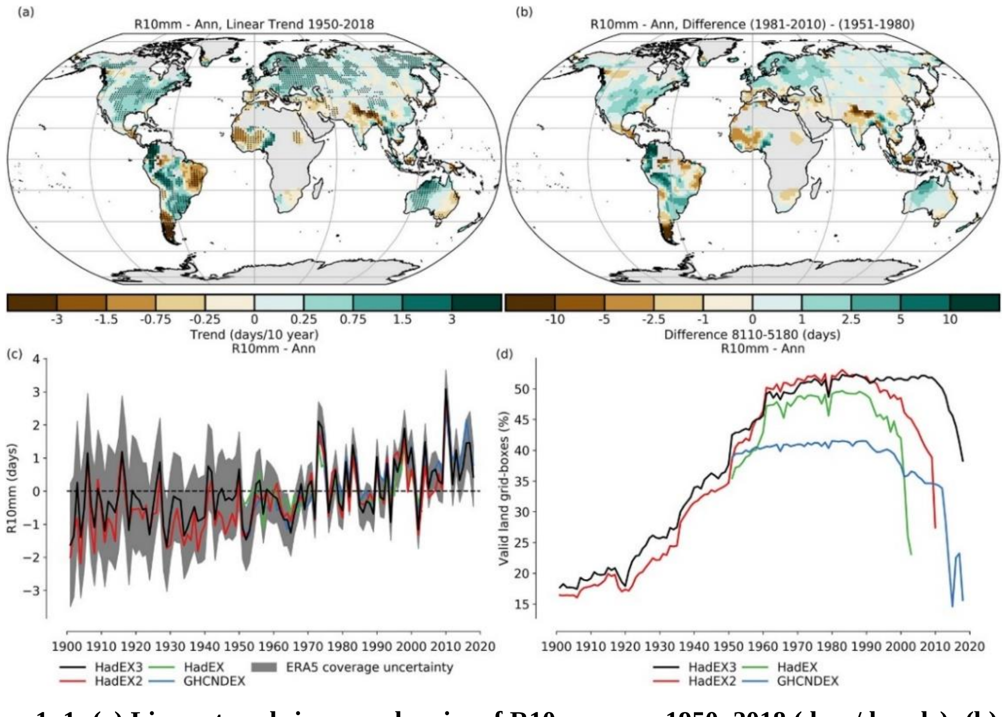


Figure 1. 1: (a) Linear trends in annual series of R10mm over 1950–2018 (days/decade); (b) difference between two 30-year periods (1981–2010 and 1951–1980); (c) time series of HadEX3 compared with HadEX, HadEX2, and GHCNDEX; (d) time series of land fraction containing grid boxes valid data (Dunn et al., 2020).

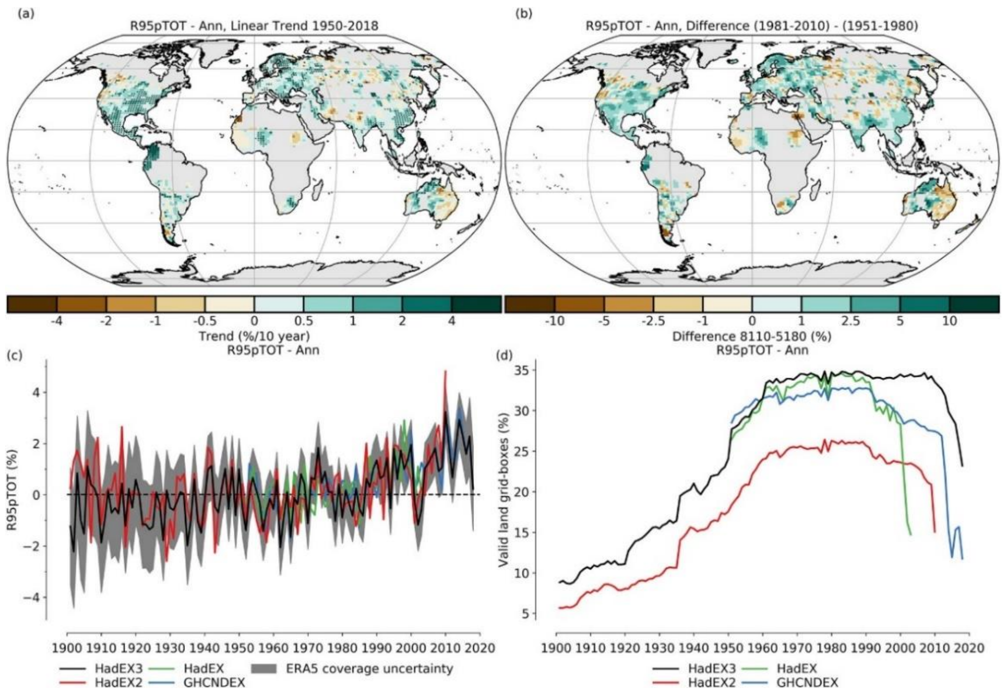


Figure 1. 2: (a) Linear trends in annual series of R95pTOT over 1950–2018 (%/decade); (b) difference between two 30-year periods (1981–2010 and 1951–1980); (c) time series of HadEX3 compared with HadEX, HadEX2, and GHCNDEX; (d) time series of land fraction containing grid boxes valid data (Dunn et al., 2020).

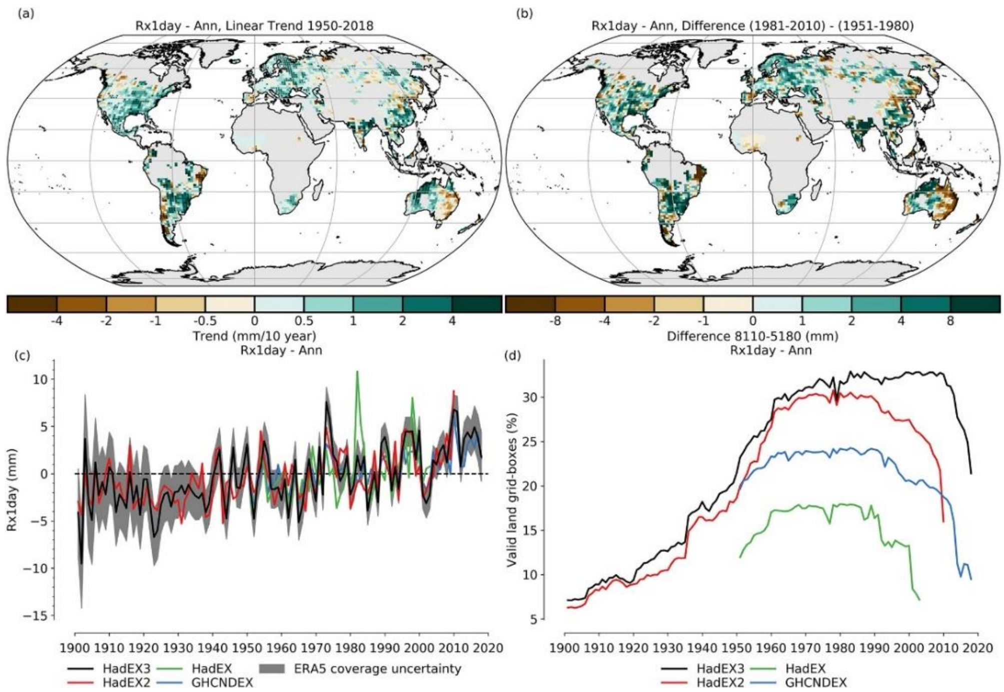


Figure 1. 3: (a) Linear trends in annual series of Rx1day over 1950–2018 (mm/decade); (b) difference between two 30-year periods (1981–2010 and 1951–1980); (c) time series of HadEX3 compared with HadEX, HadEX2, and GHCNDEX; (d) time series of land fraction containing grid boxes valid data (Dunn et al., 2020).

The analysis of historical precipitation events is crucial for understanding climate variability and change, identifying emerging trends and contextualising past events. A key challenge is understanding how these extremes are evolving at regional and local scales. However, detecting changes in observational data can be a challenging task. In Ireland, research into precipitation extremes has largely focused on relatively short records spanning the period of available digital observations, typically 1940 onwards (e.g. Kiely, 1999; Sheridan, 2001; Walsh and Dwyer, 2012). A limited number of studies have used key meteorological indicators to examine changes in precipitation extremes for Ireland (e.g. McElwain and Sweeney, 2007). These, however, have been restricted to short timescales due to a lack of available daily data. Analysis of extreme precipitation events based on observational data is strongly influenced by data quality, record length, and spatial and temporal distribution (O'Gorman, 2015; Slater et al., 2020). Due to natural climate variability, trends that are identified in short records often do not reflect the long-term climatic changes (Khan et al., 2020). The challenges posed by the lack of available long-term observations have prompted the establishment of international data rescue initiatives and the development of long-term climate series (e.g. Allan et al., 2016; Ashcroft et al., 2018; Hawkins et al., 2019). Although such datasets lengthen the period available for analysis they are nevertheless prone to discontinuities resulting from changes in site location, observing equipment and practices over time (Thorne et al., 2018). Confidence in trend detection analyses is dependent on the quality and homogeneity of the underlying data series. As such, a comprehensive quality control process and homogeneity assessment which utilises metadata regarding changes to observational practices is essential.

1.2 Trends in precipitation extremes in Europe and the British-Irish (BI) Isles

While large-scale analyses (e.g. Westra et al., 2013; Donat et al., 2016; Dunn et al., 2020) are necessary to provide a global picture of changes in extreme precipitation, understanding regional and local changes is vital for the assessment of impacts. The 5th Assessment Report of the IPCC (2014) states that the frequency and intensity of heavy precipitation events has likely increased in Europe, as a result of anthropogenic forcing, which leads to an intensification of the water cycle. Existing analyses of trends in precipitation extremes at the European scale (e.g. Zolina et al., 2012; van den Besselaar et al., 2013) show that there is some evidence of a general increase in extreme precipitation although seasonal and regional differences exist. In a recent study of 3510 European stations over the period 1950-2018, Sun et al. (2020) show that the percentage of stations showing increasing trends in Rx1day is the largest among all the continents included in the analysis, with 70% of stations showing positive trends of which 10.6% are deemed significant. Only 1.3% of stations showed significant negative trends, and these are randomly distributed across Europe. The study revealed similar findings for Rx5day, however, the fraction of stations showing increasing and significantly increasing trends were slightly higher (Table 1.1). Using a smaller number of stations (691), the study also assessed trends in Rx1day and Rx5day over the period 1900-2018 (Table 1.2). Results show a larger proportion of stations reporting increasing and significantly increasing trends in both extreme indices over the longer period of record (Sun et al., 2020).

Table 1. 1: Percentage of stations with increasing, decreasing, statistically significant increasing, and statistically significant decreasing trends in Rx1day and Rx5day based on the Mann–Kendall test over the IPCC AR6 reference regions during the 1950–2018 period. Table adapted from Sun et al. (2020) to show results for European regions.

Rx1day							Rx5day			
Reference Region	No. of stations	Increase (%)	Decrease (%)	Significant increase (%)	Significant decrease (%)		Increase (%)	Decrease (%)	Significant increase (%)	Significant decrease (%)
Northern Europe	1988	76.1	23.9	14.4	1.2		79.9	20.1	18.4	1.1
Western and central Europe	1236	65	35	6.6	1.1		63	37	10	2.1
Eastern Europe	163	73.6	26.4	7.4	1.8		79.1	20.9	14.7	0.06

Table 1. 2: Percentage of stations with increasing, decreasing, statistically significant increasing, and statistically significant decreasing trends in Rx1day and Rx5day based on the Mann–Kendall test over the IPCC AR6 reference regions during the 1900–2018 period. Table adapted from Sun et al. (2020) to show results for European regions.

Rx1day							Rx5day			
Reference Region	No. of stations	Increase (%)	Decrease (%)	Significant increase (%)	Significant decrease (%)		Increase (%)	Decrease (%)	Significant increase (%)	Significant decrease (%)
Northern Europe	314	79.9	20.1	22	0.6		90.1	9.9	38.5	0.3
Western and central Europe	323	72.8	27.2	18	0.3		76.8	23.2	25.4	1.5
Eastern Europe	54	88.9	11.1	38.9	0		88.9	11.1	46.3	0

Studies on extreme precipitation over the British-Irish Isles, have revealed an intensification of daily precipitation in the UK, with a trend towards fewer light and medium events and an increase in heavy precipitation events (Osborn et al., 2000; Fowler and Kilsby, 2003; Maraun et al., 2008; Simpson and Jones, 2014). Osborn et al. (2000) reported an increased contribution of heavy precipitation events to total precipitation, especially in winter. Maraun et al. (2008) used additional station data to provide an update to work by Osborn et al. (2000) and noted a seasonal component to trends in the contribution of heavy events to precipitation in the UK, with an increase in winter precipitation intensity and similar trends evident in spring and (to a lesser extent) autumn. Summer rainfall intensity revealed changes that are more consistent with inter-decadal variability than any overall trend. Fowler and Kilsby (2003) reported an increase in the intensity of heavy multi-day (5 and 10 day) precipitation events in northern and western parts of the United Kingdom, particularly in autumn and winter, while Tye (2015) confirmed significant increases in the probability of multi-day extreme precipitation events during the late summer and autumn. Jones et al. (2013) provide an updated study of extreme rainfall in the UK which focuses on changes to seasonal and annual maxima over the period 1961-2009. Their results show that changes in the magnitude of extreme events are spatially varied, with event magnitude in the south and east mainly influenced by seasonality whereas changes in the northern and western regions are influenced by changes in the North Atlantic Oscillation, in addition to seasonal fluctuations. Findings are consistent with previous studies (e.g. Osborn et al., 2000; Maraun et al., 2008) showing increases in winter, spring and autumn extreme precipitation events and decreases in short-duration summer rainfall.

In Ireland, work by Kiely (1999) noted an increase in the occurrence of extreme precipitation events during the second half of the 20th century, particularly in the west of the country, with a change point evident in the mid-1970s coinciding with a change point in the North Atlantic Oscillation. Leahy and Kiely (2011) subsequently updated this work with an additional ten years of observations and found evidence for changes in short-duration (1hr-24hr) precipitation extremes consistent with Kiely (1999). More recently, Walsh and Dwyer (2012) reported an increase in the frequency of heavy rain days (days with rainfall greater than 10 mm) over the period 1961-2010, however, regional variations are evident with occasionally conflicting trends from stations that are geographically relatively close.

Local and regional changes in the character of precipitation depend to a great extent on patterns of variability in atmospheric circulation. Previous studies for Europe and the British-Irish Isles highlight the importance of considering the influence of natural climate variability in the interpretation of trends in extreme precipitation analyses (e.g. Jones et al., 2013; Simpson and Jones, 2014). The North Atlantic Oscillation (NAO) is the dominant mode of natural climate variability in the British-Irish Isles (Hurrell, 1995) with positive phases of the NAO index associated with increased westerly airflow and positioning of the storm track over north-west Europe and a greater frequency of heavy rain events over the UK (Maraun et al., 2011) and Ireland (Leahy and Kiely, 2011). Observational records provide evidence that the NAO has shifted from a predominantly more negative phase before 1910 to a more positive phase in the 1920s, followed by a negative trend enduring until the late 1960s. A strong positive trend is observed until the mid-1990s followed by a downward trend that has persisted into the 21st century (Iles and Hegerl, 2017). For analysing trends in extreme precipitation, it is crucial to account for the influence of natural interannual and decadal variability (Dai, 2013). The influence of multi-decadal variability can often dominate the climate signal in shorter time series (Hegerl et al., 2015). As such, longer records are crucial to facilitate the analysis of trends over longer time scales. The greatest barrier to progress in this area is the limited availability of observed precipitation data at the appropriate spatio-temporal resolution for detection of climate-driven variability and change.

1.3 Data Rescue

Since the latter part of the 20th century, climate data have been stored in a digital format. However, for many decades prior to this, climate observations were recorded on handwritten paper documents. These observational records accumulate to hundreds of thousands of volumes in countless archives across the world (Brönnimann et al., 2018). Climate data rescue is the process of finding, imaging, digitising and processing non-digitised climate data (Wilkinson et al., 2019). While some efforts have been made in recent years to digitise these data (e.g. Brunet and Jones, 2011; Ashcroft et al., 2018; Hawkins et al., 2019; Pfister et al., 2019), the temporal extent of digitised records to this day remains grossly incomplete (Allan et al., 2011; Brunet et al., 2014). Long time series are of high relevance for climate analysis, especially for the analysis of extreme events (Kasper at al., 2015) and as input for model-based global and regional reanalysis of past weather patterns (Rennie et al., 2014; Stickler et al., 2014). However, the task of digitising these data is daunting, both logistically and in terms of the sheer volume of records that still require rescue (Brönnimann et al., 2006).

Climate data rescue involves the preservation of historical meteorological data at risk of being lost due to the vulnerability of the original paper record, poor storage, neglect, ill-informed decision making, or their being deliberately destroyed (WMO, 2016; Wilkinson et al., 2019). The main components of the data rescue process include (i) archiving of historical records and manuscripts; (ii) creating digital image files; (iii) transcription from paper to digital format; and (iv) digital storage and data accessibility.

Many national archives have invested time and resources into cataloguing historical meteorological records. However, many still hold important collections that are inadequately catalogued or not catalogued at all (Wilkinson et al., 2019). Ideally, once catalogued, a list of all materials should be posted online so that they are easily accessible. For instance, the Copernicus Climate Change Service (C3S) Data Rescue Service (https://datarescue.climate.copernicus.eu) facilitate the publication of data rescue activities through the provision of technical support services, best practice guidelines and project resources. Such registries provide information about ongoing data rescue activities and access to existing inventories which may help to avoid the duplication of time-consuming work.

The imaging of historical records preserves the record for future verification following transcription and the application of quality control processes. The WMO (2016) recommend compiling an image inventory which includes an archive reference, a brief description of the document, the type and frequency of data recorded and links to relevant metadata (e.g. Pfister et al., 2019). Imaging historical documents is a labour-intensive process, the output of which is highly dependent on the type of material being imaged and available imaging technologies as well as the size, format, and condition of the documents (Wilkinson et al., 2019). Moreover, approximately double the amount of time spent on imaging is needed for sorting and inventorying the images i.e., for every 1 day spent imaging, a further two days will be spent sorting and inventorying the images.

Transcription of historical observations from paper to digital format is a time-consuming process, with estimated keying speeds varying from 6 observations per minute to 1 observation every 3 minutes (Wilkinson et al., 2019). Various methodologies have been tried and tested, including software for speech recognition, optical character recognition (OCR) and manual keying (Brönnimann et al., 2006). Results from a recent trial to test the power of OCR software to transcribe weather records show ~50% accuracy in reading barometric pressure observations from the documents (Wilkinson et al., 2019). OCR inaccuracies result from legibility issues such as faded ink, poor handwriting, omission of decimal separators or changes in page layout. Manual keying, with a success rate of up to 99%, is recommended for the transcription of hand-written climate data. To reduce the propensity of transcription errors, double or triple keying of data, the use of form specific templates and 1-to-1 transcription (i.e., "key as you see") is advised (WMO, 2016; Ashcroft et al., 2018).

Once the data has been preserved in both digital image and digital numerical format it is important to make the data available to research communities and the public. In 2015, the WMO adopted its policy for the international exchange of climate data and products to support the implementation of the Global Framework for Climate Services (GFCS), stating that relevant data and products are to be made accessible among members on a free and unrestricted basis (https://gfcs.wmo.int/sites/default/files/IBCS%20MC-3-INF-4_en.pdf). As such, many data rescue initiatives submit their rescued images and data to open-access international repositories such as the Copernicus Climate Change Service (C3S) (https://datarescue.climate.copernicus.eu/projects) and the I-DARE Portal (https://www.idare-portal.org/). In addition, several data journals (e.g. Geoscience Data Journal) support data rescue activities through the publication of rescued data in an approved public repository alongside a description of the data and the data rescue process.

There are many advantages to undertaking data rescue activities such as:

- Increasing the temporal resolution of climate series to place current climate within a historical context.
- Validation of numerical model representations.
- Supporting climate change detection and attribution studies.
- Assessing the historical sensitivity of natural and human systems to climate variability.
- Extending the spatial distribution of observations to increase the accuracy of regional and global analysis of climate variability and change.
- Bridging the gap between palaeoclimate data and current observations.

- Assisting climate action, including mitigation and adaptation, particularly through the use of historical extremes to stress-test existing systems.
- Improving quality control and homogenisation processes by facilitating the comparison of station data with those from neighbouring sites.

Global data rescue efforts such as the International Atmospheric Circulation Reconstructions over the Earth Initiative (ACRE, Allan et al., 2011) and International Environmental Data Rescue Organization (IEDRO) have undertaken to preserve and digitise historical instrumental records. In Ireland, recent work to rescue historical meteorological records has largely focused on temperature data and monthly precipitation data. Mateus et al. (2019) reconstructed 12 long-term daily maximum and minimum air temperature series for Ireland over the period 1831-1968 building on the work of Butler et al. (2005) and McKeown et al. (2012) to extend the temperature records at Armagh Observatory and Markree, respectively. Murphy et al. (2018) compiled a 305-year continuous monthly precipitation series for the island of Ireland (1711–2016) utilising data from the UK and Ireland, including previously unpublished work from the 1970s by the British Meteorological Office. The comprehensive monthly series builds on the work of Noone et al. (2015) which developed a monthly precipitation series for 25 stations throughout Ireland for the period 1850–2015 using previously undigitised records. Subsequent research (Murphy et al., 2017; Noone et al., 2017; Wilby et al., 2016) has utilised this monthly dataset to examine historical drought events in the instrumental record, highlighting a hydro-climatic extreme that has received limited attention in Irish research. This extensive research makes a significant contribution to understanding climate variability in Ireland. The lack of long-term, quality controlled daily precipitation data, however, has hindered analysis of other historical extremes. Ireland possesses a rich heritage of historical meteorological observations that to date, have been largely unavailable for research purposes. Recent work by Mateus (2021) to develop an inventory of unexplored meteorological observations registered across the Island of Ireland in the period prior to the twentieth century highlights the extent of the data that requires digitisation.

Figure 1. 4: Hard copies of rainfall records and meteorological registers held in Met Éireann archives (Photos from Met Éireann stock images, 2021).

1.4 Metadata

Generally speaking, metadata are data about data. Metadata reflect how, where, when and by whom observations were collected and help us understand the context of the data. Station metadata should include details of any changes a station has undergone since observations commenced (Aguiler et al., 2003). The WMO (2020) provides a list of important elements that should be included in station metadata. These include, but are not limited to:

- The location and elevation of a station and dates of changes therein;
- The types, conditions and maintenance of instruments;
- Changes in the surrounding area including land use changes and vegetation;
- The frequency and time at which observations are made including time zone;
- Observer(s) information.

Station metadata are extremely valuable for data quality and homogeneity assessments. Changes in observational practices, instrumentation and station relocation can introduce data quality issues and breaks in the time series (Aguiler at al., 2003; Plummer et al., 2003;

Auer et al., 2005; Venema et al., 2012). While statistical methods can be used to detect inhomogeneities in a data series, reliable metadata can provide valuable information that can be useful in interpreting statistical homogeneity tests (Coll et al., 2020). For instance, metadata can indicate the nature and timing of the inhomogeneity and may be used for informing the magnitude of adjustments required (Ribeiro et al., 2016; Coll et al., 2018). Metadata are also useful to assess the quality and homogeneity of observations where changes affect a wide network of stations or where there are a limited number of neighbouring stations available to use as a reference (Lindau and Venema, 2016). Homogenisation methods are most effective when a combination of metadata and statistical tests are utilised (WMO, 2020).

According to the Global Climate Observing System monitoring principles (WMO, 2002), metadata should be documented and treated with the same care as the measurements. Nevertheless, the completeness of metadata for many networks is inadequate (e.g. Westerberg et al., 2009; Trewin, 2010), particularly in the case of historical observations. In addition to the vast quantities of observations that exist only in paper format, substantial quantities of metadata have not been digitised. As such, metadata must be considered an essential part of data rescue projects in order to make the data optimally useful for future users (Aguiler et al., 2003; WMO, 2020).

1.5 Citizen Science

The Expert Team on Data Rescue (ET-DARE), part of WMO's Commission for Climatology (CCL), recognises the potential of crowdsourcing as a data rescue strategy (WMO, 2016). With crowdsourcing, information is processed by online volunteers, often referred to as "citizen scientists". Traditionally, scientific inquiry has been performed by experienced professionals within academic or government organisations. In recent years, however, citizen science projects have increasingly become an intrinsic part of advancing scientific understanding (Bonney et al., 2014), with many recent climate data rescue initiatives incorporating crowdsourcing for data rescue activities as a fast and efficient methodology to digitise large amounts of observations (Brönnimann et al., 2018; Hawkins et al., 2019). Such projects have significant value in terms of the volume of information that can be derived over relatively short time frames and may be crucial to data rescue activities where a lack of funding can impede progress. For instance, the cost of rescuing weather observations from the Royal Navy World War 2 logs by professional contractors is estimated to be ~1 million GB£, with 8 million observations rescued. By contrast, the estimated cost of the original OldWeather citizen science project (www.oldweather.org/) to transcribe observations made by Royal Navy ships in the period 1913-1923 was ~100 k GB£ and rescued 7.1 million observations (see http://brohan.org/transcription_methods_review/).

Most citizen scientists work with professionals on projects that have been specifically designed either for the educational benefit of the volunteers, for the benefit of the project, or in many cases providing mutualistic benefits. The success of ongoing citizen science applications such as OldWeather.org (www.oldweather.org/), Rainfall Rescue (**https://www.zooniverse.org/projects/edh/rainfall-rescue**), and Data Rescue: Archives & Weather (DRAW, **https://citsci.geog.mcgill.ca/en/**) underscores the potential of crowdsourcing as a data rescue strategy. Moreover, studies show that the quality of crowdsourced data is extremely high. For example, during the course of the "Rainfall Rescue" project the correct value was obtained more than 95% of the time based on the transcriptions of three independent volunteers (Hawkins et al., 2019).

More recently, climate data rescue has been considered as a teaching tool in climate science education, particularly in the context of historical climatology (e.g. Creilson et al., 2008; Coleman et al., 2017). The integration of citizen science into the classroom is a relatively new, yet continually evolving concept that democratizes science education to promote critical thinking, problem solving and practical experience in students (Shah and Martinez, 2016). Given the benefits associated with citizen science both in terms of progressing data rescue activities and in increasing student engagement through active and inquiry-based learning it is likely that such initiatives will play an increasingly important role in future data rescue activities (Mitchell et al., 2017; Brönniman et al., 2018).

1.6 Quality Control

Assessing trends in extreme precipitation requires high quality data. Data quality issues may result from observer error, station discontinuity, defective instruments, imprecise transcription or post processing (Brunet and Jones, 2011; Brönnimann et al., 2018). Detecting and removing errant values in observational data is crucially important, since inaccurate data values may result in misleading and erroneous results (Westerberg et al., 2009; Durre et al., 2010; WMO, 2020). Quality control (QC) procedures vary from simple techniques, such as plotting the data against time (alone or together with neighbouring stations) or identifying data exceeding pre-fixed thresholds, to sophisticated methods that cross validate different meteorological elements at the same station and/or data from different stations. The World Meteorological Organization (WMO) provides guidelines on the QC of climate data in various documents (e.g. WMO, 1986; Plummer et al., 2003; Klein Tank et al., 2009; WMO, 2011). These include a range of tests that can be grouped into five main categories:

1. Basic integrity tests: These are QC tests that detect and flag gross erroneous values e.g. duplicate date control, rounding evaluation, detection of physically impossible, non-numeric and unrealistic values, and detection of improbable zero values.
2. Internal consistency tests: Plausibility checks which analyse each datum independently of other data. e.g. cloud cover during an extreme rainfall event.
3. Tolerance tests: These are QC tests that detect and flag those values considered outliers with respect to their own defined upper and lower limits i.e., climatological outliers.
4. Temporal consistency test: The purpose of these tests is to detect whether the observed values are consistent with the amount of change that might be expected in an element in any time interval e.g. to determine whether or not the month in question is consistent with the sample population of other such months for that station. The temporal check for outliers for a particular station is based on the premise that an individual monthly value should be similar (in statistical sampling sense) to values for the same month for other years.
5. Spatial consistency tests: These QC tests are designed to check data for consistency with neighbouring stations i.e., compare observations against values observed at the same time at adjacent stations that are affected by similar climatic influences.

Quality control (QC) of daily precipitation data is a challenging task mainly due to the large spatial and temporal variability of precipitation (Chen and Xie, 2008). Assuring the quality of a dataset is more difficult for sparse station networks (Hunziker et al., 2017) as many QC tests rely on the availability of suitable neighbouring stations to check coherency across the network (Kunkel et al., 2005; Durre et al., 2010; Vertacnik et al., 2015). Furthermore, the application of internal consistency tests that draw on the physical relationships between different climatological variables depend on the availability of observed variables (Hunziker et al., 2017). While extensive QC methods have been developed and applied to some global climate datasets (e.g., Kunkel et al., 2005; Durre et al., 2010), there is no widely accepted QC methodology for daily precipitation data (Hamada et al., 2011; Štepánek et al., 2013).

Hunziker et al. (2017) identify many common data quality issues that can be traced back to incorrect measurement practices by the observers. Sources of errors include discontinuities in the frequency of observed heavy precipitation events above a certain threshold value that may be attributed to the instrument's storage capacity which would require the observer to take more frequent measurements during extreme precipitation events. Similarly, inconsistencies in the precision with which small values are measured may introduce discontinuities in the time series which, to date, have not been adequately addressed (Rhines et al., 2015). Such inconsistencies may result from observer practices, unit conversions, or transcription errors. Quality issues also occur when smaller precipitation values are not recorded to the day on which the rain event occurred but rather get accumulated and accredited to the next measurement. Such inaccuracies will not significantly impact the analysis of monthly or annual trends. They could, however, have a significant impact on the assessment of wet days (≥1 mm) as days with rainfall <1mm are included in the calculation. Furthermore, other indices that use wet days, such as the CDD and CWD or the 'simple precipitation intensity index' (SDII), are strongly affected by the accumulation of small precipitation events (Hunziker et al., 2017). Similar issues in data quality may occur if inconsistencies arise in the frequency of observations on certain days. For instance, discontinuities may be observed in weekend records, often followed by an accumulated value on Monday (Wilby et al., 2017).

Various QC software applications such as RClimDex (Zhang and Yang, 2004), RClimDex-Extraqc (Aguilar and Prohom, 2011) and the ERA4CS INQC (Aguilar, 2020) offer a range of statistical tests that produce a combination of numerical and graphical outputs for testing the quality of daily climate data. Such programmes provide accessible and user-friendly QC methods for detecting errors in daily precipitation data. However, the tests provided in these programmes are not exhaustive and a thorough QC process will often require the application of additional QC tests, particularly in the case of historical observations.

Quality control techniques aim to verify that a reported data value is representative of what was intended to be measured and that it has not been contaminated by unrelated factors (WMO, 2016). However, QC methods can also be a source of inaccuracies, especially in daily data when these are analysed for changes in variability and extremes (WMO, 2020). The methods used for quality control of data have changed over time. In many cases, older historical data was subjected to very limited quality control checks; in particular, methods that involve spatial consistency checks with other stations have only become practical since the introduction of modern computer systems. Conversely, the use of automatic QC methods has led to a reduction in the level of manual intervention in quality control procedures, thereby increasing the potential for the detection of "false positives". As QC tests are designed to reduce the potential for erroneous data values, existing methods should be enhanced for the detection of systematically occurring data quality issues (Hunziker et al., 2017) and include strategies for evaluating QC procedures (e.g. Durre et al., 2008).

1.7 Homogenisation

Long-term series are often affected by non-climatic influences such as station relocations, instrumental and observational changes, and changes in the local environment. Such changes can adversely impact the quality of the time series through the introduction of inhomogeneities (Auer et al., 2005; Begert et al., 2005; Trewin, 2010). Some of these factors can cause abrupt shifts, while others cause gradual changes over time. Since these shifts are often of the same magnitude as the climate signal (Auer et al., 2007), the quality of long-term climate analysis depends on the homogeneity of the underlying time series (Vertačnik et al., 2015).

The aim of homogenisation techniques is the removal or reduction of any spurious non-climatic signal introducing inhomogeneities to the time series so that variability is caused only by changes in weather or climate (Freitas et al., 2013). Homogeneity tests can be broadly divided into "absolute" and "relative" methods. Absolute homogeneity methods are applied to individual candidate stations to identify statistically significant shifts in the section means (referred to as breaks or changepoints), while relative methods involve a comparison of the candidate station with a reference time series based on correlated neighbouring stations to identify breaks in the series. In general, reference series should experience the same climate influences of the candidate station so that only artificial breaks are identified and removed.

Relative homogenisation is more robust than absolute methods, provided station records are sufficiently correlated (Gubler et al., 2017) and, as with homogeneity techniques more generally, benefits from reliable metadata to account for shifts in the time series (Venema et al., 2012). However, relative approaches are often confounded by a lack of long records at neighbouring stations for comparison and by simultaneous changes in measuring techniques across a network (Peterson et al., 1998; Wijngaard et al., 2003). Spatial correlation between climatic time series depends on various factors that govern regional climate patterns, on the chosen climate variable, and on the time resolution of the series. Previous studies indicate that a minimum value of 50% common variance (r^2) is required. (Auer et al., 2005). A value less than 0.5 allows the potential discontinuities in series to disappear into statistical noise.

When neighbouring data series are unavailable or if the correlation between test series and reference series suggest a weak association, absolute methods are preferred (Wijngaard et al., 2003; Tsidu, 2012). The applicability of statistical methods highly depends on the properties of the target climatic variable and the temporal structure of the observational record. As such, the application of multiple methods is advised (Domonkos, 2013). The most widely used absolute homogeneity assessments include the Pettitt's test (Pettitt, 1979), the standard normal homogeneity test (SNHT) (Alexandersson, 1986), the cumulative deviations test (Craddock, 1979), the von Neumann's ratio test (von Neumann, 1941) and Bayesian approaches (Beaulieu et al., 2010). In recent years, different combinations of these tests have been used to assess the homogeneity of rainfall data in different climatic regions (e.g. Toreti et al., 2011; Santos and Fragoso, 2013; Yusof et al., 2019).

A comprehensive assessment of homogenisation techniques for climate series was included in the scientific programme of the COST Action HOME ES 0601 (Advances in Homogenisation Methods of Climate Series: An Integrated Approach). The HOME objective was to test the existing statistical homogenisation techniques and develop a standardised method for homogenising climate datasets. The study reports five important conclusions of test experiments: i) The reliability and accuracy of time series homogenisation highly depends on the selected statistical methods; ii) Methods that do not presume the homogeneity of reference series perform better than other methods; iii) Absolute homogenisation in automatic mode can worsen data quality rather than improving it; iv) Tested methods including tools for the joint detection and joint correction of multiple inhomogeneities outperform the other methods; v) Reliability and accuracy of the homogenisation products generally decrease with the increasing time resolution of homogenisation (Venema et al., 2012). A full list of the available homogenisation methods and software packages are provided in Table 1.3.

It is widely accepted that the reliability and accuracy of homogenisation products generally decrease with the increasing time resolution of homogenisation (Venema et al., 2012), with applications experiencing severe problems, mainly due to the decreasing correlation of daily climate series with increasing distance between stations. As such, there have been very limited applications of homogenisation methods at the daily time scale. Where methods have been applied they are generally carried out using daily temperature series due to its greater spatial coherence. Some of the most commonly used approaches for the detection and correction of daily precipitation data are described below:

(i) ACMANT (Domonkos, 2011, 2015) is fully automatic and performs break detection of temperature and precipitation data on either a monthly or a daily scale. ACMANT is better suited to the automation and rapid processing of larger networks but offers no scope for metadata integration or for graphical interpretation by the user. Furthermore, this method uses a technique to pattern scale the monthly adjustment to daily.

(ii) CLIMATOL applies the Standard Normal Homogeneity Test (SNHT) (Alexandersson and Moberg, 1997) as the basis for data homogenisation. This method is based on the application of an iterative process using a candidate time series together with a group of reference series. Climatol is more flexible in terms of handling fragmented series as reference series and does not require a common period between candidate and reference series. Climatol is one of the methods being considered in the ERA4CS INDECIS project. The objective of work package 3 of the INDECIS project is to assess the reliability of existing homogenisation packages for application to nine different variables at the daily scale. The results of the benchmarking showed that the benefits of homogenisation of daily rainfall seem quite limited in their reduction of errors. As no worsening is generally introduced in these series, the study concluded that homogenisation could contribute to eliminate outliers and infill any missing data, facilitating the calculation of climate indices from them (Guijarro, 2018).

(iii) The MASH method (Multiple Analysis of Series for Homogenisation) (Szentimrey, 1999) is a fully automated and objective relative homogeneity test procedure that does not assume homogeneity of the reference series. The candidate series is chosen from the available time series, and the remaining series are considered reference series. However, the selection of reference series is restricted by criteria for factors such as length of record and data completeness. MASHv3.03 is able to homogenise daily temperature and precipitation data where daily inhomogeneities can be derived from the monthly ones. Adjustments are applied to the daily series using techniques to pattern scale the monthly adjustment to daily data. The programme also allows for the automatic integration of metadata. The computational demand of MASH is much greater than other programmes (e.g. ACMANT) and therefore, is not suitable for the homogenisation of larger networks. Moreover, the benchmarking study of Killick (2016) to homogenise daily temperature suggests that current daily correction methods are not able to improve the series and furthermore, they failed to get the metadata to match the results produced by MASH.

(iv) RHtests (Wang and Feng, 2013) is fully automated and can homogenise on both monthly and daily scale. It is the most widely used homogenisation test globally. This is likely because it is published by and accessible via the homepage of the WMO expert team ETCCDI and is therefore considered the WMO recommended method. RHtestsV4 is an R software package (Wang et al., 2010) created to homogenise daily and monthly data. It is based on the penalized maximal *t* test and the penalized maximal *F* test, which are embedded in a recursive testing algorithm. The problem of uneven distribution of false alarm rate and detection power is also greatly alleviated by using empirical penalty functions. The time series being tested can have zero-trend or a linear trend throughout the whole period of record. The RHtestsV4 includes Quantile-Matching adjustments that are estimated with or without the use of a reference series. Additional functions such as mean-adjustments, choice of the segment to which the base series is to be adjusted (referred to as the base segment) and choices of the nominal level of confidence at which to conduct the test are facilitated. A software package called RHtests_dlyPrcp, which contains a set of functions for use in detection and adjustment of shifts in nonzero daily precipitation amounts has also been developed and made available online at: (http://cccma.seos.uvic.ca/ETCCDMI/software.shtml). This includes a quantile matching algorithm for adjusting shifts in nonzero daily precipitation series.

(v) The Wijngaard et al. (2003) approach involves the application of four homogeneity tests; SNHT, Buishand range (BR) test, Pettitt test and von Neumann ratio (VNR) to detect breaks in individual time series. The results are classified into three groups: useful, doubtful and suspect, depending on the number of tests rejecting the null hypothesis. Rather than attempting to correct for breaks detected in the series, the approach excludes series that are found to be inhomogenous.

Developing daily homogenised precipitation series is a challenging task due to the inherently high temporal and spatial variability of precipitation (Coll et al., 2018). Nevertheless, homogenised daily precipitation series will become increasingly important, both for the improved validation of climate model projections and for the assessment of variability and change in indices and extreme values. Ideally, homogenisation work should proceed in parallel with data rescue activities to allow for the development of dense observational networks and provide decision-makers with quality assured long-term series on which to base climate change adaptation decisions. Moreover, additional research is required to advance developments in the field of homogenising daily series both in the detection of breaks and in methods which would allow for the application of corrections to daily series (Szentimrey, 2011). If the goal is to establish data products that can be regularly updated and used operationally, further research is necessary to establish specific methods suitable for different time resolutions and within different climatic regions. This will be a considerable challenge for the research community and will likely demand new and highly collaborative initiatives (Coll et al., 2018).

Table 1. 3: Overview of homogenisation packages (WMO, 2020).

Package	*Resolution detection*[a]	*Detection method*	*Reference use*[b]	*Resolution correction*[c]	*Correction method*	*Primary operation*	*Metadata use*	*Variable*[d]	*Documentation*[e]	*Reference*
ACMANT[f]	**Year, month**	**Multiple breakpoints**	**Composite**	**Year, month, [day]**	**Joint (ANOVA)**	**Automatic**	**No**	**Any**	**User guide**	**Domonkos and Coll (2017)**
AnClim	**Year, month**	**Many**	**Composite, pairwise**	**Year, month, day**	**Several**	**Interactive, automatic**	**Yes**	**Any**	**Manuals**	**Štěpánek et al. (2009)**
Berkeley Earth	**Month**	**Splitting**	**Composite**	**n/a**	**n/a**	**Automatic**	**Yes**	**T**	**Article**	**Rohde et al. (2013)**
Climatol	**Month (serial), day**	**Splitting**	**Composite**	**Month (serial), day**	**Missing data filling**	**Automatic**	**Yes**	**Any**	**Manual and user guide**	**Guijarro (2018)**
GAHMDI HOMAD	**Month (serial), day**	**Multiple breakpoints**	**Selection**	**Day**	**Higher-order moment method**	**Automatic**	**Yes**	**T**	**None**	**Toreti et al. (2010, 2012)**
GSIMCLI	**Year, month**	**Multiple breakpoints**	**Composite**	**See footnote**[g]	**See footnote**[g]	**Automatic and interactive**	**No**	**T, p**	**Manuals**	**Ribeiro et al. (2017), Costa and Soares (2009)**
HOMER	**Year, season, month**	**Multiple breakpoints**	**Pairwise, joint**	**Year, month**	**Joint (ANOVA)**	**Interactive**	**Yes**	**Any**	**Basic user guide+courses**	**Mestre et al. (2013)**
iCraddock	**Year, season, month**	**Splitting**	**Pairwise**	**Year, season, month, day**	**Daily: smoothed monthly corrections**	**Interactive**	**Yes**	**Any**	**None**	**Craddock (1979), Brunetti et al. (2006)**
MASH	**Year, season, month**	**Multiple breakpoints**	**Composite**	**Month, [day]**	**Multiple comparisons**	**Automatic and interactive**	**Yes**	**Any**	**User guide**	**Szentimrey (2008, 2014)**
ReDistribu-tion Test	**Readings**	**Single breakpoints**	**No reference**	**n/a**	**n/a**	**Interactive**	**No (but it is interactive)**	**Wind**	**None**	**Petrovic (2004)**

Package	*Resolution detection*[a]	*Detection method*	*Reference use*[b]	*Resolution correction*[c]	*Correction method*	*Primary operation*	*Metadata use*	*Variable*[d]	*Documentation*[e]	*Reference*
RHtests	**Year, month, day**	**Splitting**	**Selection or no reference**	**Year, month, day**	**Multi-phase regression**	**Interactive**	**Yes**	**Any**	**User guide & courses**	**Wang (2008a &b). Wang and Feng (2013)**
R package SNHT	**Year, month**	**Splitting**	**Composite and pairwise**	**Month**	**Composite and multiple comparisons**	**Automatic**	**No**	**T**	**Help files**	**Haimberger (2007), Menne and Williams, (2009)**
PHA	**Year**	**Splitting**	**Pairwise**	**Year, [month]**	**Multiple comparisons**	**Automatic**	**Yes**	**T**	**Plain text notes**	**Menne et al. (2009)**

[a] If not noted otherwise in the column Resolution, “month” means detection is on multiple monthly series in parallel;

[b] Options are: Operator reference selection, averaging (composite reference), removing references with breaks from composite (no reference), pairwise, and joint detection; [c] Square brackets mean that the resolution is supported by the software, but the corrections are not computed at that resolution; [d] Options are: T = temperature; p = precipitation; any = Gaussian and log-normal distribution or additive and multiplicative models; [e] A user guide is limited to a few pages and shorter than a manual;

[f] ACMANT can detect breaks in both annual averages and seasonal cycle in parallel;

[g] The corrections are computed at the resolution of the data (annual or monthly series). Corrections are applied to a metric computed by GSIMCLI (user-defined: percentile, mean or median) of the probability density function (pdf) of the candidate station, which is estimated using composite references.

1.8 Research gaps, aims and objectives

In light of the above review the following research gaps were identified:

Research Gap 1: A fundamental challenge to understanding natural and anthropogenic signals of change in Irish climate is the rescue of historical observational records currently held in paper records and beyond scientific scrutiny. The opportunity exists to rescue and digitise data from the archived records in Met Éireann to provide long-term climate series for Ireland.

Research Gap 2: To date, little attention has been given to the role that students might play in efforts to rescue historical climate data. In particular, there is potential to explore procedures and methods for integrating data rescue activities into the classroom, to evaluate the learning outcomes for students and to assess the ability of students to produce reliable transcriptions.

Research Gap 3: Assessing trends in extreme precipitation requires high quality data. Homogenisation of daily data is a challenging task and few studies have attempted to detect and correct inhomogeneities in daily rainfall data. The application of homogenisation methods to daily rainfall data represents a significant research gap.

Research Gap 4: Knowledge of long-term changes in climate extremes is vital to better understand multidecadal climate variability and to place present day extreme events in a historical context. Analysis of trends in extreme precipitation in Ireland have largely been limited to the latter half of the 20th century due to data availability. This has resulted in a significant gap in knowledge of long-term changes in precipitation extremes for Ireland.

In light of these research gaps, the principal aim of this research is to rescue, quality assure and analyse daily precipitation data from archived paper records to develop a long-term daily precipitation network for Ireland.

Four main thesis objectives have been identified to address these research gaps:

- **Thesis Objective 1**: To integrate data rescue and quality assurance with teaching and learning at Maynooth University through development of approaches for citizen science in the classroom.

- **Thesis Objective 2**: To develop a long-term daily precipitation network for Ireland that takes advantage of the rich weather observing heritage on the island and to ensure the raw data and metadata is made available in a fully traceable and open way to underpin future research endeavours.
- **Thesis Objective 3**: To quality control and homogenise data to facilitate the analysis of variability and change.
- **Thesis Objective 4**: To conduct a statistical assessment of the derived quality assured long-term series to assess changes in the characteristics of extreme events.

The derived data will be hosted by Met Éireann and routinely updated and widely disseminated. The resultant long-term series will underpin climate risk management across a range of public and private sectors (water planning, flood management, agriculture, insurance, etc.) and make significant contributions to understanding extremes in Ireland. Attempts to homogenise daily data will also provide a valuable insight to the applicability of such approaches internationally.

1.9 Thesis structure

Following my 2019 departmental review I applied to progress to the PhD by Publication route. The application was approved by the departmental research student progress committee (DRSPC) in July 2019. As per the requirements of the PhD by Publication, three manuscripts were prepared for submission to peer-reviewed journals. The topics for each manuscript were developed to meet the key objectives of the research, together with suitable journals for publication i.e., wide readership, good impact factors, established publishers and modest open access costs. The final research thesis is presented in the form of a set of manuscripts which comprise the main body of the thesis, with each manuscript forming one chapter of the thesis. The manuscripts are framed by an Introduction, Discussion and Conclusion that provide context and demonstrate how the work makes an original contribution to knowledge. Details of the publications are provided below. Co-authors on each paper include colleagues at ICARUS, Maynooth, who assisted with trouble shooting and implementation of data rescue activities in the classroom; 476 students in the Department of Geography who undertook data rescue over three successive years of the GY313 Climate Change module; staff at Met Éireann who acted as enterprise partner supervisor (Mary Curley); provided support in accessing and cataloguing meteorological records (Mairéad Treanor; Conor Daly); and provided support and training with in-house quality assurance and data interpolation processes at Met Éireann (Séamus Walsh).

- **Chapter 2** presents a novel crowdsourcing initiative to integrate data rescue activities into the university classroom through the development of a research-led project to engage students in data rescue tasks for credit. Following the success of the initial project, a further two iterations have been executed across three cohorts of final year Geography students at Maynooth University, producing in excess of 3500 station years of historical daily rainfall data. The paper provides a detailed description of the methodology and access to project resources:
  Ryan, C., Duffy, C., Broderick, C., Thorne, P. W., Curley, M., Walsh, S., Daly, C., Treanor, M., and Murphy, C. (2018) Integrating data rescue into the classroom. Bulletin of the American Meteorological Society, 99 (9), 1757– 1764. https://doi.org/10.1175/BAMS-D17-0147.1

- **Chapter 3** presents the raw daily rainfall data and associated metadata transcribed during the course of the data rescue project presented in chapter 2. The paper provides details of the rescued station data and metadata, as well as a description of the process of checking for errors in the transcribed data and details of dataset use: Ryan, C., Murphy, C., McGovern, R., Curley, M., Walsh, S. and 476 students (2020) Ireland's pre-1940 daily rainfall records. Geoscience Data Journal, 1–13. https://doi.org/10.1002/gdj3.1

- **Chapter 4** utilises the rescued data presented in chapter 3 to assess changes in the characteristics of extreme precipitation at 30 long-term stations across Ireland spanning the period 1874-2019**.** The paper describes the comprehensive quality control of the rescued data, the development of long-term series and the application of homogenisation methods before presenting the results of the trend analysis: Ryan, C., Murphy, C., Curley, M. and Walsh, S. (2021) Long-term trends in extreme precipitation indices in Ireland. International Journal of Climatology. submitted for review (26th February 2021).

- **Chapter 5** draws together the key findings from the three publications. This final chapter summarises the main research findings, discusses limitations, and identifies areas for further research.

2. Integrating data rescue into the classroom

2.1 Introduction

Historical observations are fundamental for advancing understanding of past changes in climate. Fundamentally, what you do not observe you cannot understand. Globally, however, the temporal extent of digitized records to this day remains grossly incomplete, with many records available in image or hardcopy format only (Allan et al., 2011; Brunet and Jones, 2011). The task of digitising these data is daunting both logistically, and in terms of the sheer volume of records that still require rescue.

The capacity to extend current observational data holdings is largely dependent on the resources available to carry out the digitisation and transcription process. Some records have been, and continue to be, rescued via professional keying but this practice is far from sufficient to key all known records. Engagement of non-experts or 'citizen scientists' on a voluntary basis has become increasingly integral to the rescue and refinement of observational data across multiple scientific disciplines (Bonney et al., 2014). The success of ongoing citizen science applications, e.g. OldWeather.org (www.oldweather.org/) and DataRescue@Home (www.data-rescue-at-home.org), underscores the potential of crowdsourcing as a data rescue strategy.

To date, however, the potential of university students to engage in data rescue activities, for credit, has not been systematically explored. Students represent a pool of interested, qualified and instructor-guided talent that could, if harnessed, significantly accelerate progress in historical data rescue. Student-scientist partnerships (e.g. The Global Learning and Observations to Benefit the Environment (GLOBE) Program) are collaborations that engage students, teachers and scientists in authentic research projects delivering context rich, hands-on approaches to science (Mitchell et al., 2017; Allan et al., 2011). Students gain important insights into processes underpinning research, while scientists gain access to otherwise unobtainable information (Vitone et al., 2016; Harnik and Ross, 2003).

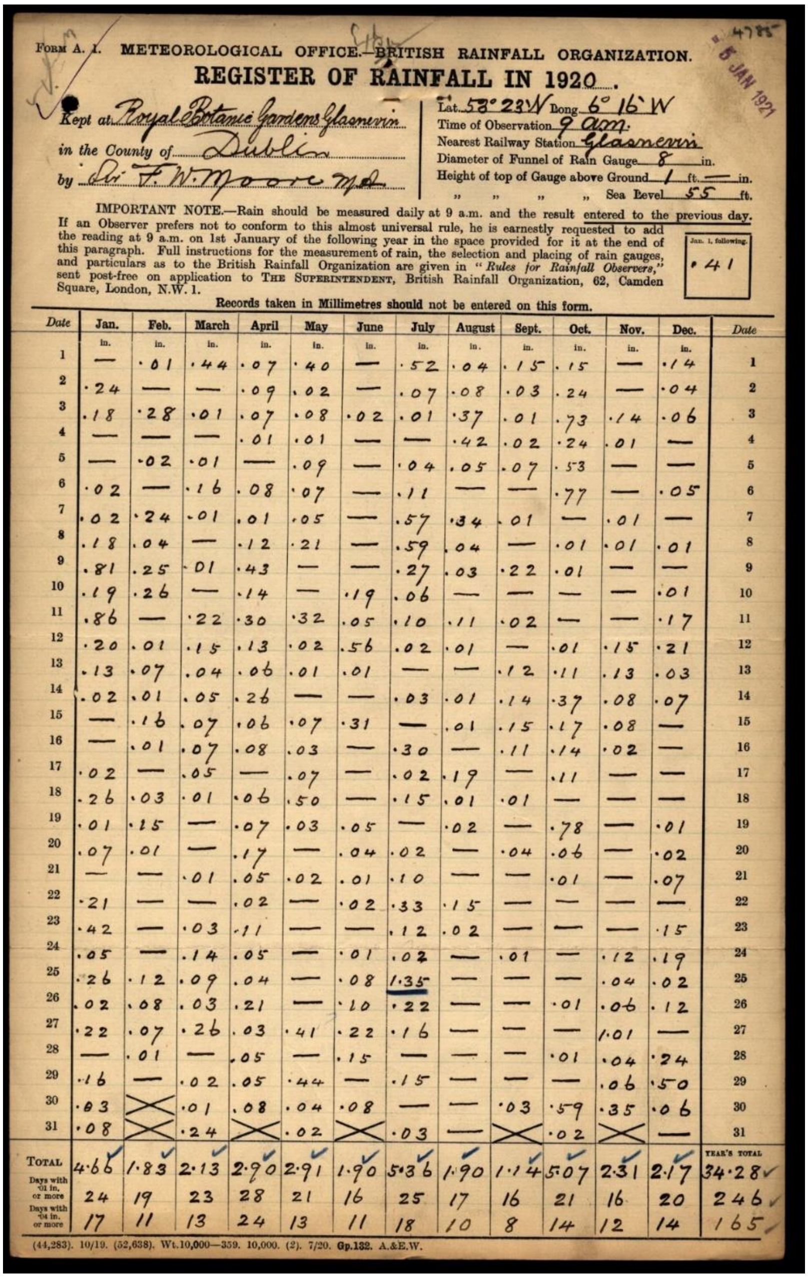
The present study describes a novel research-led accredited assignment as part of the Geography programme at Maynooth University, Ireland, in which final year undergraduate students successfully transcribed more than 1,300 station years of daily precipitation data and associated metadata for the period 1860-1939. The student rescued data relates to pre-1940 (post-1940 already digitised) daily rainfall observations from 43 stations across Ireland. Most records commence in the early 1900s, however, a few extend into the 1800s. The earliest record transcribed commences in 1866, the shortest in 1927. In this article we explore i) the potential for integrating data rescue activities into the classroom; ii) the ability of students to produce reliable transcriptions and; iii) the learning outcomes for students. We provide an overview of the assignment, its management and the student-aids implemented. Details of the workflow employed to evaluate student performance and facilitate the creation of a 'corrected' data series are provided. In the final section we discuss the students' perceptions of the project and reflect upon learning outcomes. Our experience demonstrates the substantial potential to extend such approaches to other universities, providing an enriched learning experience for the students, and a lasting legacy to the scientific community. Project resources, including transcription templates, MATLAB code and student-aids, are provided as SI.

2.2 Overview of assignment and student-aids

The assignment was set in the historical climatology component of the final year Climate Change course which focuses on the importance of long-term records for understanding past and contemporary climate change. The course took place in semester one (September – December) and was a large class with 142 undergraduate students enrolled during the 2016-17 academic year. The assignment, which accounted for 33% of the course marks, was designed as a research-led teaching experiment to reflect on the importance of historical climatology, to provide insights into the process of data rescue and to explore the power of the crowd in rescuing and transcribing meteorological data. As a lead-in to the assignment, students attended a talk given by a guest lecture from Met Éireann who emphasised the scientific and cultural importance of the data they would be working with. To coincide with the transcription component of the coursework students were required to submit a written reflection on the importance of historical climatology, utilising material and publications presented in the module. Learning outcomes were designed to provide students with:

- A first-hand experience of working with historical climate observations.
- A critical appreciation of the processes involved in data rescue, digitisation and quality assurance procedures that are essential to understanding past climate variability and change.
- First-hand experience of the powerful contribution that citizen science can make to the study of climatology, geography and other disciplines.

Scans of more than 1300 pre 1940 annual rainfall sheets (e.g. Figure 2.1) containing daily precipitation values and associated station metadata, together with templates used in transcribing the data (Figure 2.2) were provided by Met Éireann. Of the sheets provided, 274 had been previously transcribed (keyed once) by Met Éireann. These data were used to benchmark student performance.


FORM A. 1. METEOROLOGICAL OFFICE.—BRITISH RAINFALL ORGANIZATION.

4785

5 JAN 1921

REGISTER OF RAINFALL IN 1920.

Kept at Royal Botanic Gardens Glasnevin
in the County of Dublin
by Sir F. W. Moore M.D.

Lat. 53° 23' N Long. 6° 16' W
Time of Observation 9 a.m.
Nearest Railway Station Glasnevin
Diameter of Funnel of Rain Gauge 8 in.
Height of top of Gauge above Ground 1 ft. — in.
" " " " Sea Level 55 ft.

IMPORTANT NOTE.—Rain should be measured daily at 9 a.m. and the result entered to the previous day. If an Observer prefers not to conform to this almost universal rule, he is earnestly requested to add the reading at 9 a.m. on 1st January of the following year in the space provided for it at the end of this paragraph. Full instructions for the measurement of rain, the selection and placing of rain gauges, and particulars as to the British Rainfall Organization are given in "*Rules for Rainfall Observers*," sent post-free on application to THE SUPERINTENDENT, British Rainfall Organization, 62, Camden Square, London, N.W. 1.

Jan. 1, following: ·41

Records taken in Millimetres should not be entered on this form.

Date	Jan.	Feb.	March	April	May	June	July	August	Sept.	Oct.	Nov.	Dec.	Date
	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	
1	—	·01	·44	·07	·40	—	·52	·04	·15	·15	—	·14	1
2	·24	—	—	·09	·02	—	·07	·08	·03	·24	—	·04	2
3	·18	·28	·01	·07	·08	·02	·01	·37	·01	·73	·14	·06	3
4	—	—	—	·01	·01	—	—	·42	·02	·24	·01	—	4
5	—	·02	·01	—	·09	—	·04	·05	·07	·53	—	—	5
6	·02	—	·16	·08	·07	—	·11	—	—	·77	—	·05	6
7	·02	·24	·01	·01	·05	—	·57	·34	·01	—	·01	—	7
8	·18	·04	—	·12	·21	—	·59	·04	—	·01	·01	·01	8
9	·81	·25	·01	·43	—	—	·27	·03	·22	·01	—	—	9
10	·19	·26	—	·14	—	·19	·06	—	—	—	—	·01	10
11	·86	—	·22	·30	·32	·05	·10	·11	·02	—	—	·17	11
12	·20	·01	·15	·13	·02	·56	·02	·01	—	·01	·15	·21	12
13	·13	·07	·04	·06	·01	·01	—	—	·12	·11	·13	·03	13
14	·02	·01	·05	·26	—	—	·03	·01	·14	·37	·08	·07	14
15	—	·16	·07	·06	·07	·31	—	·01	·15	·17	·08	—	15
16	—	·01	·07	·08	·03	—	·30	—	·11	·14	·02	—	16
17	·02	—	·05	—	·07	—	·02	·19	—	·11	—	—	17
18	·26	·03	·01	·06	·50	—	·15	·01	·01	—	—	—	18
19	·01	·15	—	·07	·03	·05	—	·02	—	·78	—	·01	19
20	·07	·01	—	·17	—	·04	·02	—	·04	·06	—	·02	20
21	—	—	·01	·05	·02	·01	·10	—	—	·01	—	·07	21
22	·21	—	—	·02	—	·02	·33	·15	—	—	—	—	22
23	·42	—	·03	·11	—	—	·12	·02	—	—	—	·15	23
24	·05	—	·14	·05	—	·01	·02	—	·01	—	·12	·19	24
25	·26	·12	·09	·04	—	·08	1·35	—	—	—	·04	·02	25
26	·02	·08	·03	·21	—	·10	·22	—	—	·01	·06	·12	26
27	·22	·07	·26	·03	·41	·22	·16	—	—	—	1·01	—	27
28	—	·01	—	·05	—	·15	—	—	—	·01	·04	·24	28
29	·16	—	·02	·05	·44	—	·15	—	—	—	·06	·50	29
30	·03	X	·01	·08	·04	·08	—	—	·03	·59	·35	·06	30
31	·08	X	·24	X	·02	X	·03	—	X	·02	X	—	31
TOTAL	4·66	1·83	2·13	2·90	2·91	1·90	5·36	1·90	1·14	5·07	2·31	2·17	YEAR'S TOTAL 34·28
Days with ·01 in. or more	24	19	23	28	21	16	25	17	16	21	16	20	246
Days with ·04 in. or more	17	11	13	24	13	11	18	10	8	14	12	14	165

(44,283). 10/19. (52,638). Wt.10,000—359. 10,000. (2). 7/20. Gp.132. A.&E.W.

Figure 2. 1: Example of an annual rainfall sheet distributed to students.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	**Metadata**		**Day**	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Month**	**M_total**	**M_rainday**	**M_wetday**	**QCmonth**
2	image_no	1163	**1**		0.01	0.44	0.07	0.4		0.52	0.04	0.15	0.15		0.14	**1**	4.66	24	17	4.66
3	station	Royal Botanic Ga	**2**	0.24			0.09	0.02		0.07	0.08	0.03	0.24		0.04	**2**	1.83	19	11	1.83
4	county	Dublin	**3**	0.18	0.28	0.01	0.07	0.08	0.02	0.01	0.37	0.01	0.73	0.14	0.06	**3**	2.13	23	13	2.13
5	observer	Sir F.W.Moore	**4**				0.01	0.01			0.42	0.02	0.24	0.01		**4**	2.9	28	24	2.9
6	lat_degree	53	**5**		0.02	0.01		0.09		0.04	0.05	0.07	0.53			**5**	2.91	21	13	2.91
7	lat_min	23	**6**	0.02		0.16	0.08	0.07		0.11			0.77		0.05	**6**	1.9	16	11	1.71
8	lon_degree	6	**7**	0.02	0.24	0.01	0.01	0.05		0.57	0.34	0.01		0.01		**7**	5.36	25	18	5.49
9	lon_min	16	**8**	0.18	0.04		0.12	0.21		0.59	0.04		0.01	0.01	0.01	**8**	1.9	17	10	1.96
10	year	1920	**9**	0.81	0.25	0.01	0.43			0.27	0.03	0.22	0.01			**9**	1.14	16	8	1.14
11	start_month	1	**10**	0.19	0.26		0.14			0.19	0.06				0.01	**10**	5.07	21	14	5.07
12	start_day	1	**11**	0.86		0.22	0.3	0.32	0.05	0.1	0.11	0.02			0.17	**11**	2.31	16	12	2.31
13	funnel_diameter_	8	**12**	0.2	0.01	0.15	0.13	0.02	0.56	0.02	0.01		0.01	0.15	0.21	**12**	2.17	20	14	2.17
14	gauge_height_ft	1	**13**	0.13	0.07	0.04	0.06	0.01	0.01			0.12	0.11	0.13	0.03					
15	gauge_height_inch	0	**14**	0.02	0.01	0.05	0.26			0.03	0.01	0.14	0.37	0.08	0.07					
16	gauge_pattern		**15**		0.16	0.07	0.06	0.07	0.31		0.01	0.15	0.17	0.08						
17	railway	Glasnevin	**16**	0.02	0.01	0.07	0.08	0.03		0.3		0.11	0.14	0.02						
18	railway_distance		**17**	0.26		0.05		0.07		0.02	0.19		0.11							
19	elevation_msl	55ft	**18**	0.01	0.03	0.01	0.06	0.5		0.15	0.01	0.01			0.01					
20	jan_1_following	0.41	**19**	0.07	0.15		0.07	0.03	0.05		0.02		0.78		0.02					
21	year_total	34.28	**20**		0.01		0.17		0.04	0.02		0.04	0.06		0.07					
22	year_rainday	246	**21**			0.01	0.05	0.02	0.01	0.1			0.01							
23	year_wetday	165	**22**	0.21			0.02		0.02	0.33	0.15				0.15					
24	time_of_observati	9:00 AM	**23**	0.42		0.03	0.11			0.12	0.02				0.19					
25	comments_sheet		**24**	0.05		0.14	0.05		0.01	0.02		0.01		0.12	0.02					
26	observer_notes		**25**	0.26	0.12	0.09	0.04		0.08	1.35				0.04	0.12					
27			**26**	0.02	0.08	0.03	0.21		0.1	0.22			0.01	0.06						
28			**27**	0.22	0.07	0.26	0.03	0.41	0.22	0.16				1.01	0.24					
29			**28**		0.01		0.05		0.15				0.01	0.04	0.5					
30			**29**	0.16		0.02	0.05	0.44		0.15				0.06	0.06					
31			**30**	0.03		0.01	0.08	0.04	0.08			0.03	0.59	0.35						
32			**31**	0.08		0.24		0.02		0.03			0.02							

Sheet1

Figure 2. 2: Template distributed to students for transcribing the data. Columns D-O contain the 31x12 array of values extracted for comparison against the corresponding file. Differences detected by the MATLAB code between the two corresponding files are highlighted in yellow. Columns Q-S contain respective values for monthly totals, total number of rain days per month and total number of wet days per month as recorded by the observer. Column T (red) shows the quality check used by the students to compare automatically generated monthly totals to the monthly totals recorded by the observer and inputted to columns D-O.

Upon receipt of the digital images, the file format was converted from PNG file type to JPEG, reducing file size from 17MB to ~1.4MB, without any apparent loss of resolution. This facilitated the distribution of images to students and avoided potential logistical issues concerning students accessing large files. Each student was assigned eighteen annual rainfall sheets to transcribe, and each sheet was assigned twice (i.e. to two different students). Such double key data entry is a widely used method of quality control to detect miskeyed information. For the purpose of this project, double key entry was also necessary for the allocation of student grades as a component of the grade was based on transcription accuracy which could only be ascertained through comparison. The assignment was explained in an hour-long lecture that outlined both the assignment and the rationale. Individual directories containing eighteen, randomly selected, annual rainfall sheets were created and distributed to each student via Dropbox along with an Excel template for keying the data. Students were provided with a link from which they downloaded their personalised directory and performed the transcription component of the coursework. Students were given five weeks in which to complete and submit their transcribed data.

Several additional student-aids were implemented:

- A simple video tutorial was produced to describe the different sections of the annual rainfall sheets and to demonstrate the transcription process and posted to the university's online learning platform (Moodle).
- An automated quality assurance check was integrated into the Excel template to generate a monthly total based on the daily values transcribed. Comparison against the monthly totals recorded on the original sheet provided students with a basic quality check to ensure transcribed monthly totals were consistent with the originals (Figure 2.2).
- An online discussion forum was set up on the Moodle course page through which student queries could be addressed. Students were invited to post questions relating to, for example, differences identified in monthly totals, difficulties interpreting metadata or to clarify illegible values.
- An in-class check-in clinic was organised at the mid-way point of the assignment to highlight frequently asked questions from the online forum and to allow students the opportunity to raise questions in class.

Once completed, students were requested to compile the transcribed files into a compressed zipped directory, maintaining a consistent file naming convention, and upload the files to the online course portal. Developing a file naming convention that uniquely identifies the image and student (e.g. image number_student number.xlsx) was essential to data post-processing. This allowed for a comparison between equivalent sheets based on image number and assignment of grades to students based on an evaluation of their performance.

2.3 Evaluating student performance

Figure 2.3 provides an overview of the workflow implemented once student transcriptions were received. Having verified that consistent naming conventions existed for all student files, an initial comparison of the double keyed data was performed. Student files corresponding to the 274 previously transcribed files provided by Met Éireann were also included. MATLAB code was developed to compare double keyed data by extracting the 31 x 12 array representing daily values for one year and highlight cells where differences exist (Figure 2.2). From this, two directories were generated; (i) a consistent directory comprising student transcribed files indicating zero differences between double keyed sheets and (ii) an inconsistent directory of student transcribed files containing differences between double keyed sheets.

Utilising results from the first comparison, a ‘correct’ version of each file in the inconsistent directory was created by teaching staff. These corrected or ‘master’ files were manually created by examining highlighted differences against the original scanned images and adjusting transcribed values accordingly. Because this required consideration of solely a small subset of the total values transcribed, it was achieved in a matter of hours by two members of staff.

With master files now available for all original annual rainfall sheets (i.e., those showing no differences, together with the manually corrected files) a second comparison was performed to evaluate i) all student transcriptions against the corresponding master files to assess student errors and; ii) student transcriptions against Met Éireann transcriptions for common sheets to benchmark student performance relative to Met Éireann. This second comparison generated a new set of consistent and inconsistent directories containing files with zero differences and files with greater than zero differences, respectively. Results were derived from these two output directories.

Student grades were assigned based on the number of differences within any transcribed file i.e., a file containing zero differences obtained full marks; for files showing differences, marks were deducted for each incorrect entry. Figure 2.4a displays the frequency (%) of student submissions being less than or equal to x, where x is the number of errors per transcribed file (bounded by 0 and 365 for all years except leap years when 366 are possible), together with a bar graph (Fig. 2.4b) categorising the total number of incorrect files by actual number of errors per transcribed sheet. Sixty-two per cent of student transcribed files show no errors when compared against the corrected master files. In 96% of student transcribed files fewer than 5% of data entries were incorrect i.e., 96% of files had fewer than 20 errors. A review of all incorrect files reveals that 57% of the transcribed files containing errors had fewer than 5 errors, 90% had fewer than 20 errors and only 3% had greater than 40 errors. Cumulative error across all 2556 files transcribed by students reveals a percentage error of less than 1%. In the unlikely event that both students were wrong (i.e. that they produced identical errors) the errors would not be detected in the comparison. Such errors would, however, be identified during the subsequent application of more comprehensive quality assurance techniques.

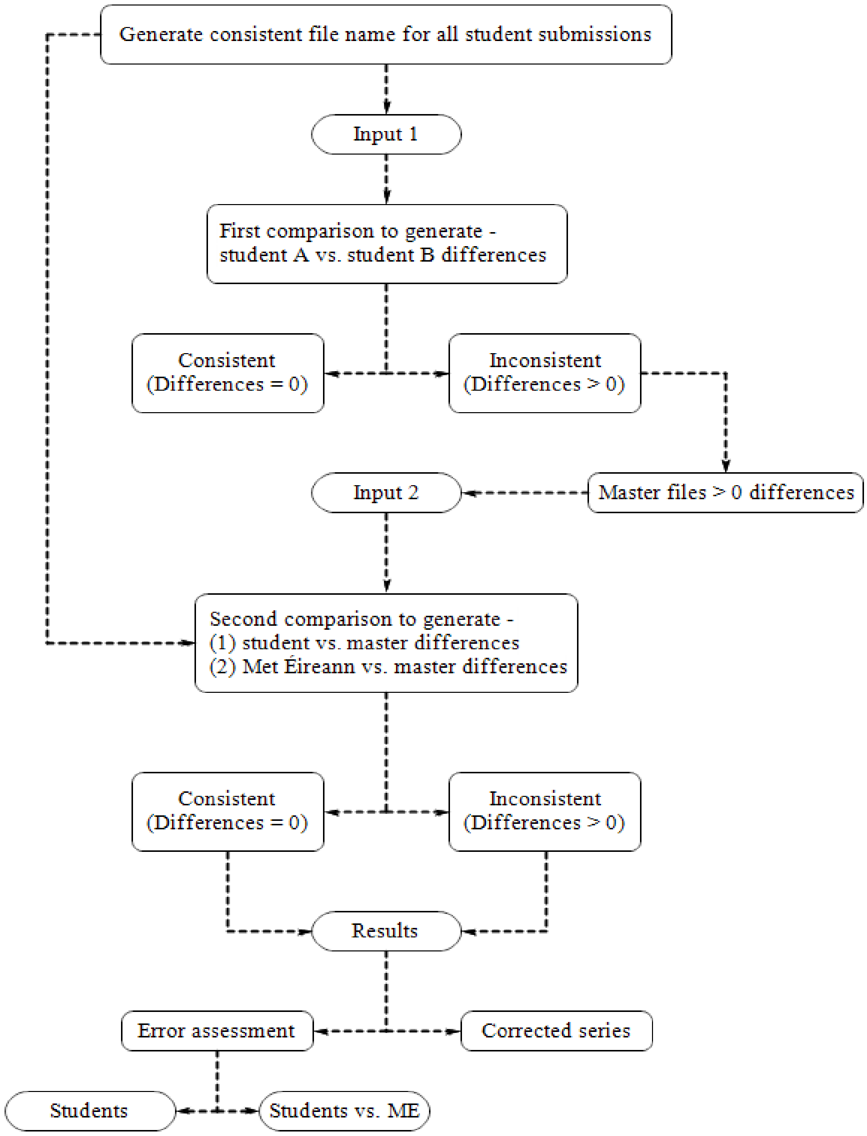


Figure 2. 3: Flowchart outlining the workflow from receiving student submissions to final output and assessment of errors.

2.4 Benchmarking students against professionals

A final assessment was carried out utilizing the 274 sheets transcribed by both Met Éireann and the students. This provided a benchmark against which student performance could be evaluated. Figure 2.4c displays the proportion of incorrect files by error category. While the students have a smaller number of incorrect files overall (39% for students compared to 49% for Met Éireann), the majority of Met Éireann's incorrect files lie in the lowest error propensity category. Fifty-two per cent of incorrect files from students contain fewer than 5 errors, 47% of incorrect files are accounted for by the 6-40 error categories with the remaining 1% of incorrect files showing greater than or equal to 40 errors. Eighty-five per cent of incorrect files from Met Éireann contain fewer than 5 errors. However, Met Éireann also has a greater number of incorrect files showing greater than 40 errors.

Upon further investigation it was noted that different approaches to the transcription process i.e., row-based or column based, had an impact on the number of errors produced within individual files. Specifically, the row-based approach employed by Met Éireann resulted in the majority of incorrect files falling into the lowest error category (i.e. ≤5 errors). Alternatively, the column-based approach adopted by the students produced a greater number of files containing errors in the intermediate categories (i.e. 6 – 10, 11 – 20, and 21 – 40). Utilising a column-based approach, the incorrect placement of a data value occasionally propagated down the entire monthly column. While the individual values were correct, differences were flagged due to the data values being input to the incorrect date/cell. An example of errors arising from the two different approaches is highlighted in Figure 2.2.

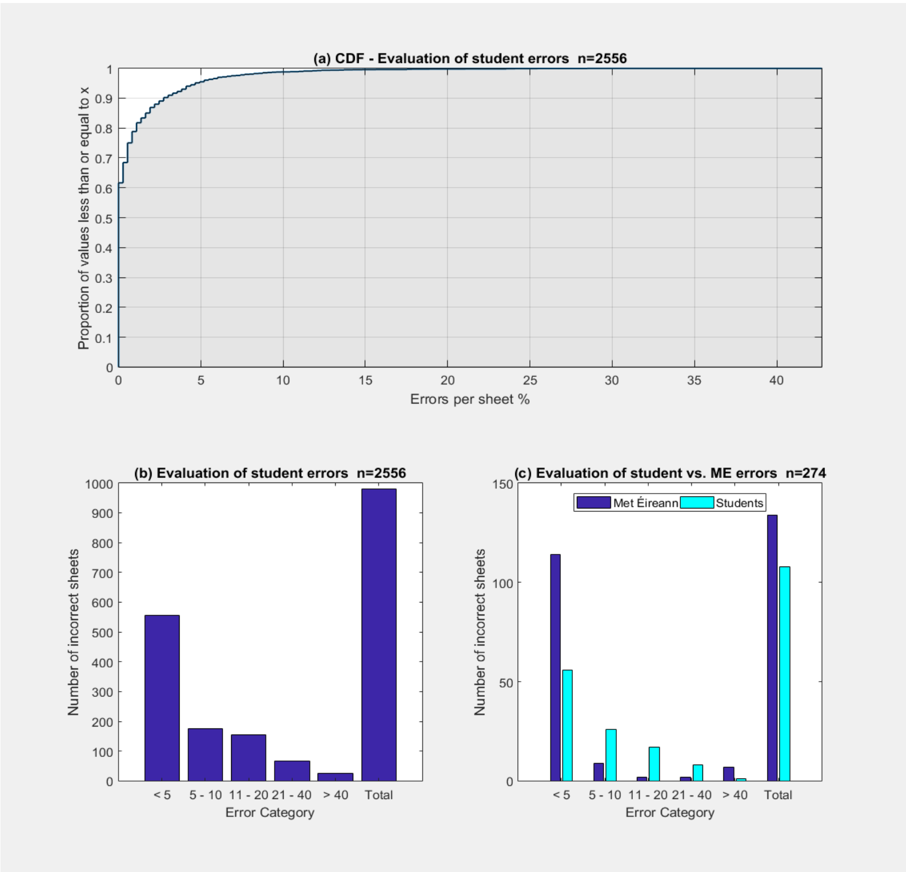


Figure 2. 4: (a) Empirical CDF showing the frequency (%) of student submissions (n = 2556) being less than or equal to x, where x is the percentage of errors per sheet; (b) Total number of incorrect student sheets by number of errors per sheet; (c) Comparison of errors by error category for sheets transcribed by both students and Met Éireann (n = 274).

2.5 Evaluating the student experience

On completion of the project feedback was provided to the students. Firstly, we explained how the evaluation was performed and what the results told us about the accuracy with which the students had transcribed the data. Secondly, data transcribed by the students for a local station were collated and presented to the class to exemplify their contribution to understanding changes in historical rainfall. We discussed trends and notable events present in the early record and how it commensurately nuanced our understanding of long-term local climate. To investigate students' perspectives of the project and reflect upon their perceptions of learning outcomes, a formal assessment was conducted via anonymous questionnaire completed in class in the final week. Table 2.1 presents the list of statements provided to students and the extent to which proportions of the class who responded agreed or disagreed with each of these statements.

The response was positive across all aspects of the assignment. The majority of students (>90%) gained insights into the process of data rescue and an appreciation of the role of historical data in climate research. The assignment afforded students a first-hand experience of working with raw historical climate data and dataset development; in particular seeing the processes behind the cataloguing and imaging of historical rainfall sheets, the importance of double entry to avoid gross errors, and the value of metadata recorded by the original observers. Over 90% of participants stated that they could see value in their contribution (extending available digital records) to understanding Irish rainfall trends. Presentation of the long-term time series at the end of the project supported this consensus. Notably, 80% of respondents stated that they would prefer to participate in assignments like this over other, more traditional, assignments. Students were given the opportunity to continue working with the data, in smaller groups, in the second semester as part of a research methods workshop whereby they could conduct statistical analysis on the transcribed data and present findings in individual research reports.

Students were asked to identify which student-aids they found most useful in completing the assignment. The most popular were the online discussion forum and video tutorial, at 44% and 36%, respectively. The management of these resources demanded a significant investment of time by teaching staff. In particular, the online discussion forum which received more than 500 queries from students. Nevertheless, effective management of the project and student-aids facilitated the development of the 'corrected' dataset by notably reducing the propensity for errors and the amount of time required to carry out post-processing of the data. While questions have been raised over the accuracy and reliability of citizen science-produced datasets, a number of projects have demonstrated the potential for enhancing data quality through practical management (van der Velde et al., 2017; Kosmala et al., 2016). Despite the success of this initial experiment a number of issues arose that require consideration in future iterations:

- Given that the misplacement of daily values, rather than incorrect entry of data was the main source of error, simple additions to the excel template used to transcribe the data, for instance, delineating the margins of the columns and rows could significantly reduce propensity for errors. Additionally, adopting the row-based approach used by Met Éireann to transcribe the data could further reduce errors associated with the incorrect placement of data values.
- Reducing the number of sheets allocated to the students could potentially increase student motivation and reduce the number of errors further.
- Having successfully integrated the transcription component into the coursework the next iteration of the assignment will build in basic quality assurance and statistical analysis on the resulting dataset. Development of such skills will further deepen the research experience offered to students.

Table 2. 1: List of statements provided to students to assess learning outcomes and the extent to which proportions of the class (%) agreed or disagreed with these statements (n = 62).

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Total
The assignment provided insights into the process of data rescue	58	40	0	2	0	100
I now have an appreciation of the role of historical data in climate research	55	39	5	2	0	100
Supports provided were sufficient to aid completion of the assignment	66	29	2	3	0	100
The process of working with the data/images was clear and easy to follow	41	41	10	7	2	100
The assignment provided insights into the power of citizen science	44	35	16	5	0	100
I would prefer to participate in CA like this over more traditional assignments	60	20	10	7	3	100
The workload was appropriate to the level of credit given	45	32	11	8	3	100
Overall, I found the assignment to be a valuable learning experience	44	44	6	6	0	100

I was more motivated than is usual for me in doing this assignment	35	23	39	3	0	100
I feel that I have contributed to understanding Irish rainfall trends	44	47	10	0	0	100
I think that future students would be pleased to participate in a similar assignment	52	37	5	5	2	100
I think that working with data is a useful skill to have gained	52	44	3	2	0	100
I did not gain any valuable skills from this assignment	0	2	11	34	53	100

2.6 Conclusion

We have outlined an initial successful attempt to integrate data rescue into the classroom using a research-led approach to teaching and learning that complies with the prerequisites of pedagogy within the university curriculum. Additionally, the project motivated students by engaging them in a practical exercise whereby their contribution adds considerable value to research. Such initiatives promote the development of mutualistic collaborations between national meteorological services and higher-level institutions with the paired objectives of accelerating science. At the same time, students gain a first-hand understanding of the processes that underpin data rescue and research.

Having established confidence in the transcription, the next steps in developing a long-term, daily rainfall network for Ireland will involve i) repeated iterations of this assignment with subsequent classes, ii) the application of comprehensive quality assurance and homogeneity techniques, and iii) analysis of the derived long-term record to assess changes in the characteristics of extreme rainfall events. Metadata transcribed by the students will be systematically extracted and catalogued to facilitate the process. A final objective is to make the data widely available to national and international researchers. To this end, the data shall be shared with the recently awarded Copernicus Climate Change Service Global Land and Marine Observations Database service led by co-author Thorne. It is hoped that the framework outlined in this paper may be integrated into teaching programmes across other universities and used to highlight the importance of historical meteorological sources to students and to encourage their involvement in data rescue efforts. We provide the developed student-aids to help realise this objective. These materials are available via both SI and https://www.maynoothuniversity.ie/icarus/data-rescue-classroom The latter shall be updated as we make modifications in subsequent years.

3. Ireland’s pre-1940 daily rainfall records

3.1 Introduction

There is an increasing demand for climatological information to support scientific understanding of the impacts of climate change. Accurate and reliable long-term meteorological records are essential for understanding climate trends and variability, evaluating reanalysis products and climate models and climate risk management (e.g. Murphy et al., 2019; Noone et al., 2017; Wilby et al., 2016; Matthews et al., 2014). Consequently, much work is being carried out internationally to rescue historical climate data and develop long-term climate series (e.g. Hawkins et al., 2019; Coll et al., 2019; Ashcroft et al., 2018; Allan et al., 2016; Cornes et al., 2012).

There is a long history of meteorological observation in Ireland. The earliest known weather observations were recorded at Kilkenny for the Duke of Ormond by John Kevan in 1682. From 1684 to 1686, William Molyneux and George Ashe, under the auspices of the Dublin Philosophical Society, documented weather at Trinity College, Dublin. When records resumed here in 1708 to 1709, Samuel Molyneux documented observations of temperature, rainfall and pressure. In 1787, Richard Kirwan established the first Irish meteorological station at Cavendish Row, Dublin. These records are the earliest surviving observations compiled with the aid of accurate instruments. Kirwan maintained the observations until 1808, by which time regular meteorological observations had commenced at the National Botanic Gardens (Dixon, 1987). Observations from Armagh Observatory date from 1796 to present, providing the longest continuous series from any single site on the island (Butler et al., 2007). In 1829, the Ordnance Survey Office located in Phoenix Park, Dublin began recording systematic meteorological measurements providing continuous readings up to the present day.

In 1859, George James Symons, working with a network of voluntary observers, set up a system for gathering and publishing rainfall records from across Britain and Ireland. The data from Symons’ observational network were published in British Rainfall (formerly English Rainfall), which for the year 1860 reports that there were then 168 stations in the network. No observations were included from sites in Scotland or Ireland, and only three from Wales (Glasspoole, 1952). The earliest available observations taken at Irish sites are from 1864. In 1900 the network founded by Symons became known as the British Rainfall Organisation and was later (in 1919) formally transferred to the aegis of the British

Meteorological Office. At this time there were approximately 5,000 observers contributing to the rainfall network (Glasspoole, 1952).

At the end of the 19th century, under the direction of the British Meteorological Department, led by Admiral Robert Fitzroy of the Royal Navy, an operational network of observing stations was established. In 1868, the Met Office equipped seven observatories, including two in Ireland, with autographic instruments, providing the first continuous record of a variety of meteorological parameters at a selection of locations (Eden, 2009). The records, along with the Phoenix Park records, which were maintained by the Royal Engineers, represent the early efforts of the British Meteorological Office to establish a network of meteorological stations across Britain and Ireland.

The Irish Meteorological Service, later Met Éireann, was established in December 1936 and subsequently took over responsibility for the network of Irish stations from the British Meteorological Office. The registers that pre-date the establishment of the Irish Meteorological Service were later transferred from the British Meteorological Office to the current service in Ireland. The substantial paper records are carefully preserved in Met Éireann's archives, but until now have largely remained in paper format for the years prior to 1941.

Met Éireann maintains the National Climate Database. The database comprises observations received by Met Éireann from the current network of staffed, synoptic, climatological and rainfall stations. Observations are quality controlled and archived in Met Éireann's database. Daily observations of rainfall have been digitised back to 1941 and temperature to 1961 (Walsh, 2013). Work to digitise and transcribe long-term daily minimum and maximum air temperatures is currently underway at the National University of Ireland (NUI), Galway (Mateus et al., 2019). Other long-term series include work by Murphy et al. (2018) to compile a 305-year continuous monthly rainfall series for the island of Ireland (1711–2016) utilising data from the UK and Ireland, including previously unpublished work from the 1970s by the British Meteorological Office. The comprehensive monthly series builds on the work of Noone et al. (2015) which developed a monthly rainfall series for 25 stations throughout Ireland for the period 1850–2015. However, until now there has not been a concerted effort to construct a long-term daily rainfall series for Ireland using historical records.

This paper describes recent work, undertaken as part of the PhD research of the lead author, to digitise and transcribe historical daily rainfall records from Met Éireann's extensive archive collection. The main objective of this work was to create a digital archive of the paper records of Ireland's longest meteorological stations, and from these imaged records and additional rainfall registers to extend the availability of long-term daily rainfall data prior to 1941. The data presented here are the raw data and associated metadata. It is envisaged that by presenting the data in its original state it can be easily integrated into current international data rescue initiatives, e.g. Copernicus Climate Change Service Global Land and Marine Observations Database, and that future research will have recourse to the raw data.

The remainder of the paper is organised as follows: Section 2 provides a description of the data. First we discuss the digitisation of historical records from Met Éireann's archives, then we outline the data transcription process and provide details of the stations transcribed, followed by a description of the metadata that was collected during the transcription process. In Section 3, we describe the process of checking for errors in the transcribed data. Section 4 provides details of dataset use, while Section 5 concludes with a proposed pathway to complete the imaging and transcription of stations and meteorological data that remain on paper.

3.2 Description of the data

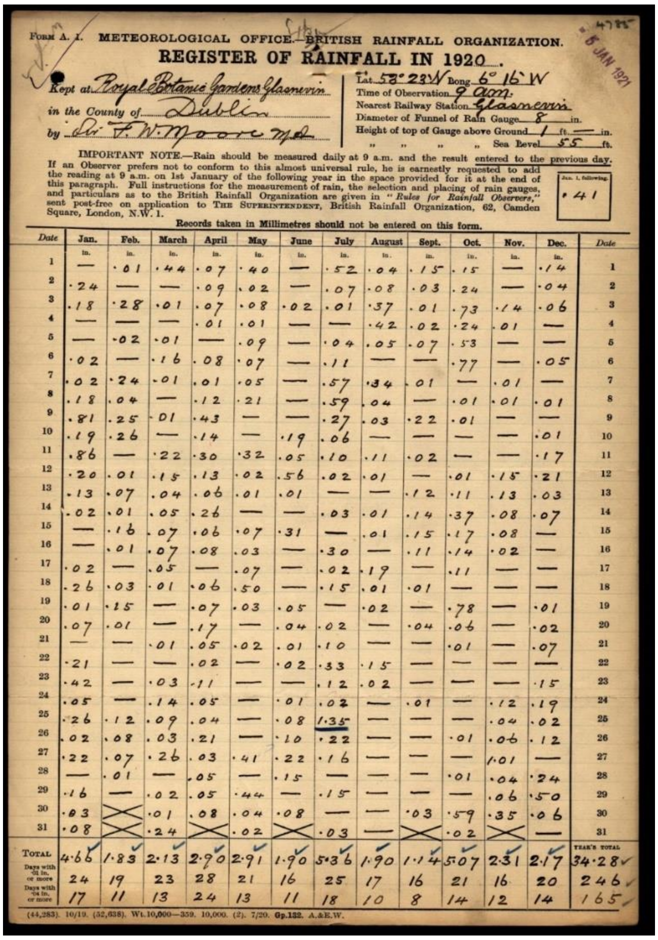
3.2.1 Data sources

The daily rainfall data was largely taken from annual Rainfall Registers and Meteorological Registers held in Met Éireann's archives. These registers are described in detail below. In addition, digital images of station records for Kells (Headfort) and the National University of Ireland (NUI), Galway, were obtained from the National Library of Ireland and NUIG, respectively.

3.2.2 Data imaging

Rainfall Registers

Previous work by Met Éireann focused on imaging the single-sheet annual rainfall registers collected from the early rainfall network, specifically the original handwritten rainfall observations that were taken by the volunteers who worked as part of Symons' rainfall network in Ireland (Treanor et al., 2011). In 2014, the Met Éireann Library received funding to collate, catalogue and preserve the registers in this collection. The records are known as the 'Rainfall Registers' and contain daily rainfall data and station metadata from various locations throughout Ireland. Readings were taken once a day and recorded on a standard form issued by Symons and later the British Rainfall Organisation (Figure 3.1). The collection includes readings from every county in the Republic of Ireland and intermittent observations taken in Antrim, Armagh, Derry, Down, Fermanagh and Tyrone in Northern Ireland. There are over 700 stations in total, which return data for varying time periods between 1864 and 1940.



FORM A. 1. METEOROLOGICAL OFFICE.—BRITISH RAINFALL ORGANIZATION.

REGISTER OF RAINFALL IN 1920.

Kept at Royal Botanic Gardens Glasnevin
in the County of Dublin
by Sir F. W. Moore M.A.

Lat. 53° 23' N Long. 6° 16' W
Time of Observation 9 a.m.
Nearest Railway Station Glasnevin
Diameter of Funnel of Rain Gauge 8 in.
Height of top of Gauge above Ground 1 ft. — in.
" " " " Sea Level 55 ft.

IMPORTANT NOTE.—Rain should be measured daily at 9 a.m. and the result entered to the previous day. If an Observer prefers not to conform to this almost universal rule, he is earnestly requested to add the reading at 9 a.m. on 1st January of the following year in the space provided for it at the end of this paragraph. Full instructions for the measurement of rain, the selection and placing of rain gauges, and particulars as to the British Rainfall Organization are given in "*Rules for Rainfall Observers,*" sent post-free on application to THE SUPERINTENDENT, British Rainfall Organization, 62, Camden Square, London, N.W. 1.

Jan. 1, following: .41

Records taken in Millimetres should not be entered on this form.

Date	Jan.	Feb.	March	April	May	June	July	August	Sept.	Oct.	Nov.	Dec.	Date
	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	
1	—	.01	.44	.07	.40	—	.52	.04	.15	.15	—	.14	1
2	.24	—	—	.09	.02	—	.07	.08	.03	.24	—	.04	2
3	.18	.28	.01	.07	.08	.02	.01	.37	.01	.73	.14	.06	3
4	—	—	—	.01	.01	—	—	.42	.02	.24	.01	—	4
5	—	.02	.01	—	.09	—	.04	.05	.07	.53	—	—	5
6	.02	—	.16	.08	.07	—	.11	—	—	.77	—	.05	6
7	.02	.24	.01	.01	.05	—	.57	.34	.01	—	.01	—	7
8	.18	.04	—	.12	.21	—	.59	.04	—	.01	.01	.01	8
9	.81	.25	.01	.43	—	—	.27	.03	.22	.01	—	—	9
10	.19	.26	—	.14	—	.19	.06	—	—	—	—	.01	10
11	.86	—	.22	.30	.32	.05	.10	.11	.02	—	—	.17	11
12	.20	.01	.15	.13	.02	.56	.02	.01	—	.01	.15	.21	12
13	.13	.07	.04	.06	.01	.01	—	—	.12	.11	.13	.03	13
14	.02	.01	.05	.26	—	—	.03	.01	.14	.37	.08	.07	14
15	—	.16	.07	.06	.07	.31	—	.01	.15	.17	.08	—	15
16	—	.01	.07	.08	.03	—	.30	—	.11	.14	.02	—	16
17	.02	—	.05	—	.07	—	.02	.19	—	.11	—	—	17
18	.26	.03	.01	.06	.50	—	.15	.01	.01	—	—	—	18
19	.01	.15	—	.07	.03	.05	—	.02	—	.78	—	.01	19
20	.07	.01	—	.17	—	.04	.02	—	.04	.06	—	.02	20
21	—	—	.01	.05	.02	.01	.10	—	—	.01	—	.07	21
22	.21	—	—	.02	—	.02	.33	.15	—	—	—	—	22
23	.42	—	.03	.11	—	—	.12	.02	—	—	—	.15	23
24	.05	—	.14	.05	—	.01	.02	—	.01	—	.12	.19	24
25	.26	.12	.09	.04	—	.08	1.35	—	—	—	.04	.02	25
26	.02	.08	.03	.21	—	.10	.22	—	—	.01	.06	.12	26
27	.22	.07	.26	.03	.41	.22	.16	—	—	—	1.01	—	27
28	—	.01	—	.05	—	.15	—	—	—	.01	.04	.24	28
29	.16	—	.02	.05	.44	—	.15	—	—	—	.06	.50	29
30	.03		.01	.08	.04	.08	—	—	.03	.59	.35	.06	30
31	.08		.24		.02		.03	—		.02		—	31
TOTAL	4.66	1.83	2.13	2.90	2.91	1.90	5.36	1.90	1.14	5.07	2.31	2.17	YEAR'S TOTAL 34.28
Days with ·01 in. or more	24	19	23	28	21	16	25	17	16	21	16	20	246
Days with ·04 in. or more	17	11	13	24	13	11	18	10	8	14	12	14	165

(44,283). 10/19. (52,638). Wt.10,000—359. 10,000. (2). 7/20. Gp.132. A.&E.W.

Figure 3. 1: Sample rainfall register. The annual forms report daily rainfall observations and station metadata for various locations throughout Ireland.

Meteorological Registers

Non-digitised records include a collection of meteorological registers consisting of 224 bound volumes, 128 folders and 353 folios of station records covering the period 1855-1976. The collection includes observations from Ireland's longest operating stations, many of which are still operational. The registers have been arranged by date and region and catalogued by their location in the archives at Met Éireann (Treanor et al., 2011). They provide a daily record of weather parameters at various locations throughout Ireland covering the mid-19th century until the latter part of the 20th century. Observers recorded data from a selection of instruments that measured temperature, pressure, precipitation, sunshine, and other fields. Readings were taken at least once a day and recorded on a British Meteorological Office issued broad-sheet form that was designed to contain a month's data (Figure 3.2). At the end of the month, the forms were posted back to the British Meteorological Office for centralised calculations and corrections.

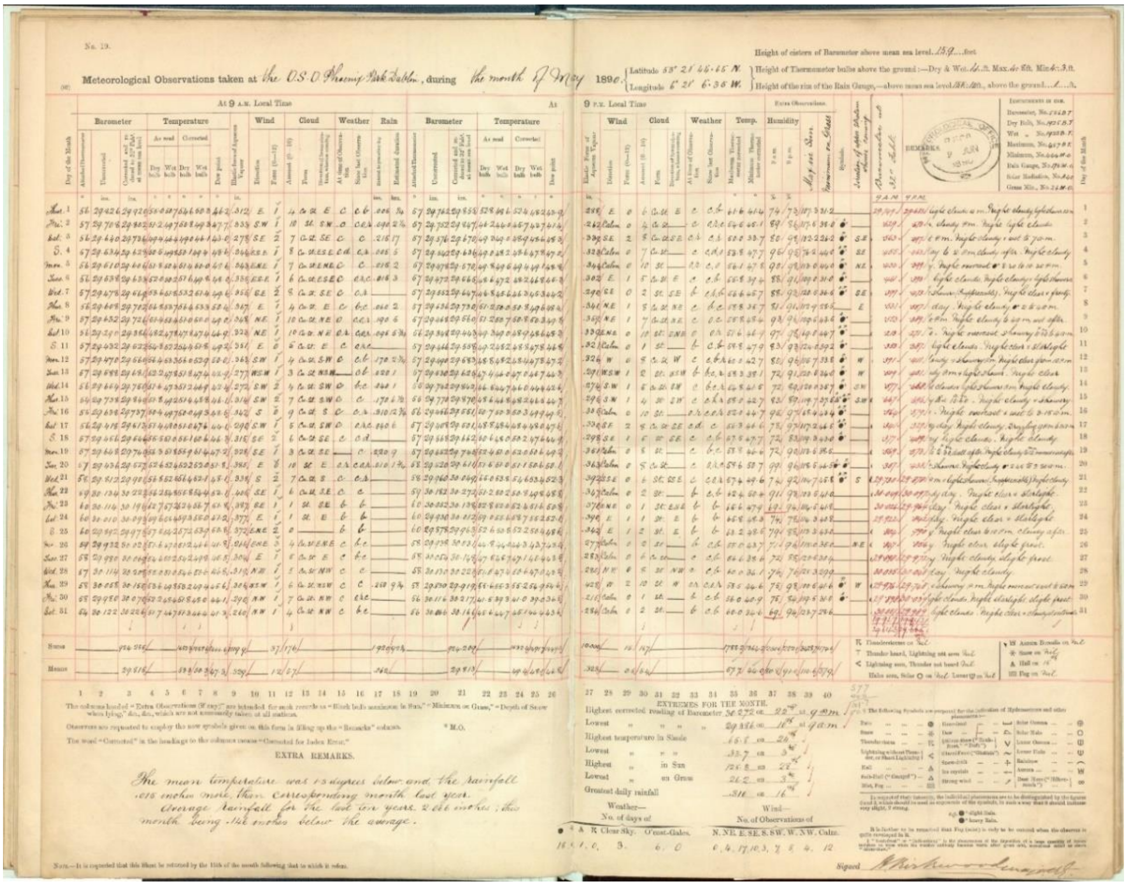


Figure 3. 2: Sample meteorological register. The monthly forms report daily observations for multiple parameters including temperature, pressure, precipitation, sunshine, wind and other parameters.

A Metis EDS professional digital scanner (Figure 3.3) was purchased to progress the digitisation of the historical meteorological registers. To date meteorological observations from eight long-term stations have been scanned and integrated into the digital database (Table 3.1). Work is ongoing to digitise the remaining manuscripts. Individual pages were scanned as high quality tif files and categorised using the standard naming convention adopted by Met Éireann. The scanned images were then uploaded to Met Éireann's database where they will be made available for research purposes.

Figure 3. 3: Metis EDS Gamma professional digital scanner used to image historical meteorological registers.

Station	Start year	End year
Birr	1873	1951
Blacksod	1884	1956
Fitzwilliam Square	1869	1935
Malin Head	1885	1955
Markree	1869	1968
Phoenix Park	1866	1959
Roches Point	1873	1956
Valentia	1873	1950

Table 3. 1: Meteorological registers imaged and integrated into Met Éireann's digital archives.

3.2.3 Data transcription

Daily rainfall observed at 114 sites throughout Ireland were transcribed as part of this work (Figure 3.4). Stations were selected based on record length, continuity and spatial distribution. Additional stations that could potentially be used to infill gaps in long-term data series were also included. The transcription from paper and digital image format to digital numerical format was largely undertaken by final year Geography students at Maynooth University as part of a novel crowdsourcing initiative to integrate data rescue activities into the classroom. Ryan et al. (2018) presented an innovative approach to data rescue by developing a research-led project to engage students in data rescue tasks for credit. The study explored i) the potential for integrating data rescue activities into the classroom, ii) the ability of students to produce reliable transcriptions, and iii) the achieved learning outcomes for students. The work was facilitated by the provision of student aids including written guidelines, transcription templates with an automated quality-assurance check, a video tutorial, in-class workshops and an online discussion forum. An evaluation of learning outcomes and student's perceptions of the project demonstrated a positive educational experience. Following the success of the initial project, a further two iterations have been executed across three cohorts of final year Geography students at Maynooth University, producing in excess of 3500 station years of historical daily rainfall data. A detailed description of the methodology and access to resources is provided by Ryan et al. (2018) (doi:10.1175/BAMS-D-17-0147.1).

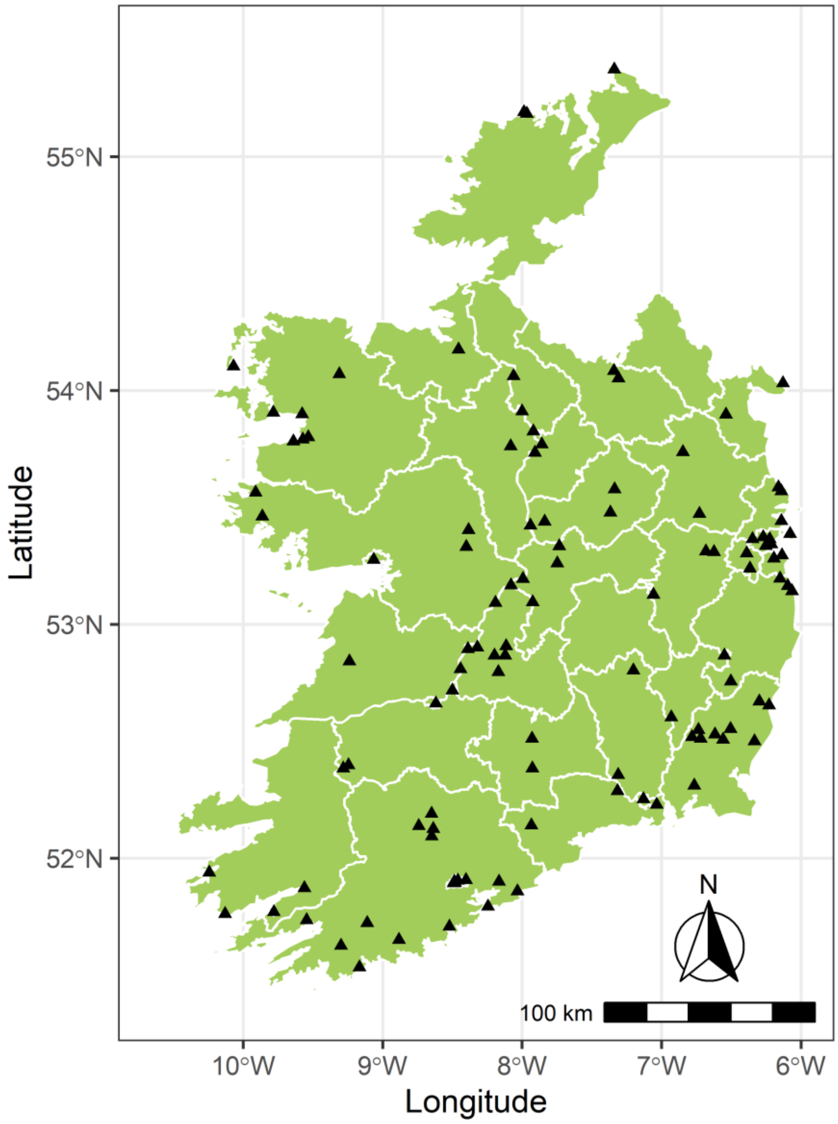


Figure 3. 4: Location of stations for which daily rainfall data was transcribed as part of this work. Individual folders containing a data file and metadata file are available for each station.

The majority of the data was transcribed from the annual rainfall registers. Additional rainfall data was extracted from the long-term meteorological registers listed in Table 3.1. Rainfall register observations were, for the most part, recorded in inches of rainfall. Data extracted from the meteorological registers were recorded in millimetres from the start of record up to 1914/15 when the unit of measurement changed to inches. Table 3.2 provides a quick reference to the different units of measurement used at each station. Detailed changes of units of measurement are provided in the individual station metadata files. The original units of measurement have been preserved here. The number of stations varies significantly throughout the years, with only a small number of stations available in the early record (Figure 3.5). The earliest observations recovered were taken at the National University of Galway (NUIG) in 1864. From 1900 the number of stations increases significantly. In total, 3616 station years (~1.32 million daily values) were transcribed. Details of all stations transcribed are provided in Table 3.2.

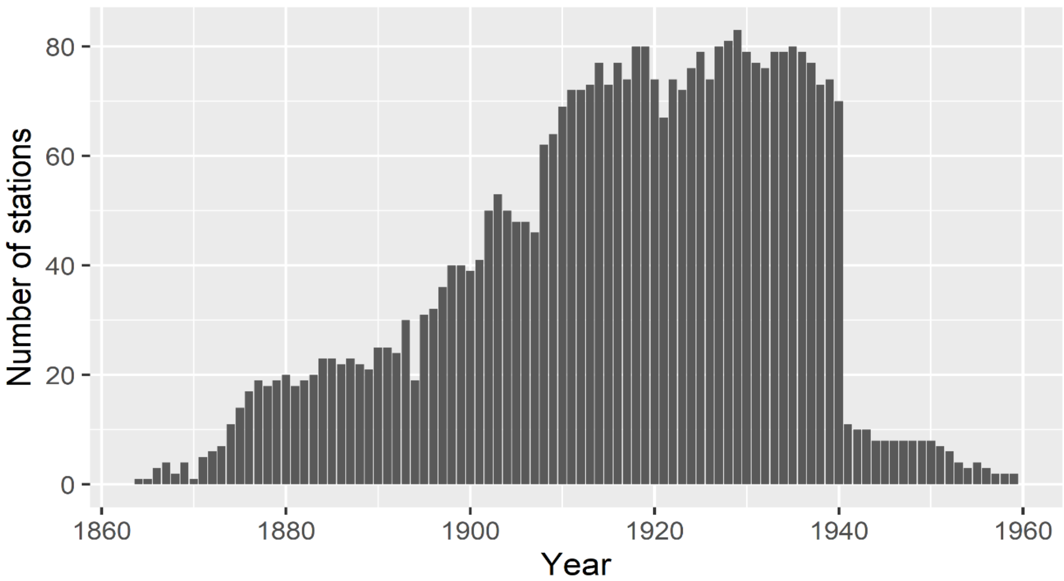


Figure 3. 5: Number of stations each year for which daily rainfall data was transcribed. Post 1940 data is available through Met Éireann's website.

Table 3. 2: Station details for data transcribed as part of this work. Individual folders containing data files and metadata files are provided for each station listed. Station coordinates are given in degrees, minutes, seconds (DMS).

STATION	LAT	LON	COUNTY	START YR	START MTH	START DAY	END YR	END MTH	END DAY	% NA	UNITS
ABBEYFEALE (PRESBYTERY)	522300	91700	LIMERICK	1936	1	1	1940	12	31	5.6	INCH
ABBEYFEALE (SPRINGMOUNT)	522350	91440	LIMERICK	1925	1	1	1940	12	31	3.2	INCH
AHASCRAGH (CLONBROCK)	532410	82300	GALWAY	1876	1	1	1940	12	31	14.1	INCH
ARDEE (LISRENNY)	535350	63220	LOUTH	1886	1	1	1913	7	31	0.0	INCH
ARDGILLAN	533509	60942	DUBLIN	1894	1	1	1938	2	28	2.6	INCH
ATHLONE (TWYFORD)	532620	75020	WESTMEATH	1872	1	1	1940	12	31	14.6	INCH
ATHLONE O.P.W.	532520	75630	WESTMEATH	1902	1	1	1940	12	31	2.7	INCH

BALLINACURRA	515400	81000	CORK	1904	4	1	1959	11	30	0.0	IN/MM
BALLYCUMBER (BELAIR)	532000	74400	OFFALY	1923	1	1	1940	12	31	8.5	INCH
BALLYHAISE (AGR.COLL.)	540305	71835	CAVAN	1923	1	1	1943	12	31	31.0	INCH
BALLYNAHINCH CASTLE	532740	95150	GALWAY	1913	1	1	1940	12	31	0.1	INCH
BANAGHER (CANAL HSE.)	531140	75940	OFFALY	1893	1	1	1940	12	31	46.1	INCH
BELGARD CASTLE	531810	62320	DUBLIN	1914	4	1	1940	12	31	8.2	INCH
BIRR CASTLE	530540	75530	OFFALY	1875	1	1	1951	12	31	2.0	IN/MM
BLACKSOD POINT S.W.S.	540610	100410	MAYO	1884	10	1	1956	10	31	1.6	IN/MM
BORRIS	523610	65550	CARLOW	1890	1	1	1895	12	31	19.6	INCH
BRAY (FASSAROE)	531146	60900	WICKLOW	1871	1	1	1916	12	31	0.0	INCH
CAHERAGH	513732	91755	CORK	1910	9	1	1935	12	31	0.3	INCH
CAHIR ABBEY	522300	75540	TIPPERARY	1924	1	1	1940	12	31	11.8	INCH
CASHEL (BALLINAMONA)	523038	75543	TIPPERARY	1911	1	1	1940	12	31	0.0	INCH
CASTLECOMER HSE	524810	71210	KILKENNY	1908	1	1	1940	12	31	6.4	INCH
CASTLECONNELL RECTORY	524300	83000	LIMERICK	1914	1	1	1939	12	31	0.9	INCH
CASTLEFORBES	534610	75130	LONGFORD	1911	1	1	1940	12	31	3.4	INCH
CASTLETOWNSHEND (GLENBARRAHANE)	513200	91000	CORK	1892	1	1	1940	12	31	5.1	INCH
CASTLETOWNSHEND (SEAFIELD)	513200	91000	CORK	1914	1	1	1940	12	31	0.0	INCH
CLONAKILTY(SHANNON VALE)	513900	85300	CORK	1910	1	1	1931	12	31	14.6	INCH
CLONGOWES WOOD COLL	531840	64100	KILDARE	1906	1	1	1940	12	31	32.6	IN/MM
CLOONDRA O.P.W.	534400	75430	LONGFORD	1902	1	1	1940	12	31	5.1	INCH
CLOVERHILL	540500	72030	CAVAN	1896	1	1	1933	7	31	0.2	INCH
CORK (PARK VIEW)	515412	82738	CORK	1910	1	1	1937	12	31	0.4	INCH
CORK (THE PALACE)	515340	82850	CORK	1895	1	1	1911	12	31	0.0	INCH
DERRYNANE ABBEY	514540	100750	KERRY	1871	1	1	1940	12	31	23.3	INCH
DROMOD (RUSKEY)	534930	75510	LEITRIM	1902	1	1	1940	12	31	2.6	INCH
DRUMSHANBO(LOUGH ALLEN SLUICES)	540340	80340	LEITRIM	1902	1	1	1940	12	31	2.6	INCH

DRUMSNA (ALBERT LOCK)	535440	80000	ROSCOMMON	1903	1	1	1940	12	31	2.6	INCH
DUBLIN (CLONTARF)	532150	61330	DUBLIN	1912	1	1	1940	12	31	17.2	INCH
DUBLIN(FITZWILLIAM SQUARE)	532008	61512	DUBLIN	1908	1	1	1937	9	30	0.0	INCH
DUBLIN (PHOENIX PARK)	532150	62050	DUBLIN	1881	1	1	1959	12	31	0.0	IN/MM
DUBLIN (RINGSEND)	532030	61250	DUBLIN	1912	1	1	1940	12	31	13.9	INCH
DUN LAOGHAIRE (HARBOUR YARD)	531740	60810	DUBLIN	1905	1	1	1940	12	31	0.2	INCH
DUNFANAGHY (ST.PATRICK'S)	551100	75800	DONEGAL	1880	1	1	1940	12	31	46.7	INCH
DUNMANWAY	514320	90637	CORK	1905	1	1	1937	6	30	3.1	INCH
ENFIELD (SUMMERHILL)	532820	64340	MEATH	1896	1	1	1940	12	31	2.2	INCH
ENNISCOE	540410	91842	MAYO	1874	1	1	1923	12	31	4.1	INCH
ENNISCORTHY (BALLYHIGHLAND)	523040	64312	WEXFORD	1871	1	1	1919	10	31	0.0	INCH
ENNISCORTHY (ENNISCORTHY RECTORY)	523024	63335	WEXFORD	1926	1	1	1929	12	31	0.0	INCH
ENNISCORTHY(MONART RECTORY)	523144	63706	WEXFORD	1903	1	1	1907	12	31	2.2	INCH
ENNISCORTHY (MONKSGRANGE)	523104	64658	WEXFORD	1925	1	1	1931	12	31	0.3	INCH
ENNISCORTHY (SUMMERVILLE)	523310	63018	WEXFORD	1908	1	1	1923	12	31	0.0	INCH
ENNISCORTHY (WOODBROOK)	523255	64410	WEXFORD	1902	1	1	1940	12	31	54.2	INCH
FOULKESMILL (LONGRAIGUE)	521838	64559	WEXFORD	1874	1	1	1940	12	31	11.2	INCH
GLENASMOLE D.C.W.W.	531420	62200	DUBLIN	1900	1	1	1940	12	31	13.2	INCH
GLENGARRIFF (ILNACULLIN)	514405	93245	CORK	1914	1	1	1940	12	31	7.4	INCH
GOREY (COURTOWN HOUSE)	523910	61345	WEXFORD	1908	1	1	1941	12	31	0.0	INCH
GOREY (RAM'S GATE)	524010	61800	WEXFORD	1928	1	1	1940	12	31	2.4	INCH
GOREY (WELLS)	523000	62000	WEXFORD	1893	1	1	1923	12	31	6.6	INCH
GREENORE	540150	60750	LOUTH	1876	1	1	1940	12	31	6.3	INCH
GREYSTONES(BURNABY LODGE)	530831	60356	WICKLOW	1928	1	1	1940	12	31	0.1	INCH
GREYSTONES (RATHDOWN HSE.)	530949	60540	WICKLOW	1918	1	1	1929	7	31	8.7	INCH
HACKETSTOWN RECTORY	525200	63300	CARLOW	1918	1	1	1940	12	31	0.0	INCH
HORN HEAD	551123	75916	DONEGAL	1892	1	1	1917	12	31	15.4	INCH

HOWTH CASTLE	532310	60440	DUBLIN	1913	1	1	1940	12	31	0.0	INCH
INAGH (MT. CALLAN)	525030	91418	CLARE	1908	1	1	1940	12	31	6.1	INCH
KELLS (HEADFORT)	534410	65050	MEATH	1893	10	1	1952	6	30	0.4	INCH
KENMARE (DERREEN)	514610	94650	KERRY	1912	1	1	1939	12	31	49.8	INCH
KENMARE (SHEEN FALLS)	515220	93340	KERRY	1921	1	1	1940	12	31	10.0	INCH
KILCONNEL (RECTORY)	531952	82404	GALWAY	1875	1	1	1886	12	31	0.1	INCH
KILLALOE (BALLINA)	524830	82630	CLARE	1866	1	1	1940	12	31	22.4	INCH
KINSALE (SCILLY HSE.)	514225	83120	CORK	1922	1	1	1940	12	31	0.0	INCH
KYLEMORE CASTLE	533350	95439	GALWAY	1889	1	1	1904	12	31	6.3	INCH
LIMERICK (MULGRAVE ST.)	523940	83710	LIMERICK	1924	1	1	1940	12	31	0.0	INCH
LISMORE CASTLE	520824	75600	WATERFORD	1909	1	1	1940	12	31	41.6	INCH
LOTA LODGE NO.1	515422	82407	CORK	1909	1	1	1937	3	31	3.5	INCH
LOTA LODGE NO.2	515422	82407	CORK	1911	1	1	1925	12	31	0.0	INCH
MALAHIDE (SEAMOUNT)	532630	60830	DUBLIN	1908	1	1	1939	3	31	3.2	INCH
MALIN HEAD	552220	72020	DONEGAL	1885	1	1	1955	12	31	0.1	IN/MM
MALLARANNY	535420	94700	MAYO	1919	1	1	1940	12	31	0.4	IN/MM
MALLOW (HAZELWOOD)	521125	83900	CORK	1928	1	1	1943	12	31	13.1	INCH
MALLOW (LONGUEVILLE)	520814	84426	CORK	1896	1	1	1937	12	31	0.0	INCH
MALLOW (OLD DRAMORE)	520537	83852	CORK	1890	1	1	1895	12	31	4.7	INCH
MALLOW (SUMMERHILL)	520732	83817	CORK	1896	1	1	1921	6	30	0.3	INCH
MARKREE CASTLE	541030	82720	SLIGO	1874	11	1	1940	12	31	0.0	INCH
MEELICK (VICTORIA LOCK)	531000	80450	OFFALY	1902	1	1	1940	12	31	2.6	INCH
MONASTEREVIN (MOORE ABBEY GARDENS)	530735	70333	KILDARE	1898	1	1	1918	12	31	0.8	INCH
MULLINGAR (BALLYNEGAL)	533440	72010	WESTMEATH	1908	1	1	1940	12	31	12.4	INCH
MULLINGAR(BELVEDERE HSE)	532840	72200	WESTMEATH	1898	1	1	1940	12	31	9.3	INCH
NENAGH (BALLYGIBBON)	525422	80653	TIPPERARY	1901	1	1	1912	12	31	0.0	INCH
NENAGH (CASTLE LOUGH)	525340	82320	TIPPERARY	1875	2	1	1940	12	31	19.3	INCH

NENAGH (CLASHNEVIN)	525200	80720	TIPPERARY	1884	1	1	1899	9	13	0.0	INCH
NENAGH (NENAGH LODGE)	525403	81918	TIPPERARY	1879	1	1	1889	12	31	4.9	INCH
NENAGH(ST.MARY'S RECTORY)	525200	81156	TIPPERARY	1919	1	1	1930	10	31	0.0	INCH
NENAGH (TRAVERSTON)	524746	81013	TIPPERARY	1907	1	1	1937	12	31	0.1	INCH
NEWPORT(BURRISHOOLE HSE.)	535352	93442	MAYO	1899	1	1	1919	12	31	0.1	INCH
NUIG	531635	90350	GALWAY	1864	1	1	1952	6	30	12.6	INCH
PILTOWN(KILDALTON ABBEY)	522120	71840	KILKENNY	1877	1	1	1940	12	31	45.4	INCH
PORTLAW (MAYFIELD)	521710	71900	WATERFORD	1881	1	1	1940	12	31	3.9	INCH
PORTUMNA O.P.W.	530530	81130	GALWAY	1929	1	1	1940	12	31	16.7	INCH
QUEENS COLLEGE CORK	515334	82925	CORK	1866	1	1	1907	12	31	40.7	INCH
ROCHES POINT	514735	81440	CORK	1873	7	1	1940	12	31	0.0	IN/MM
ROYAL BOTANIC GARDENS (GLASNEVIN)	532221	61618	DUBLIN	1880	1	1	1925	12	31	2.2	INCH
SHANAGARRY (KINOITH)	515130	80200	CORK	1910	1	1	1940	12	31	3.0	INCH
SHILLELAGH (COOLLATTIN)	524520	63010	WICKLOW	1908	1	1	1940	12	31	3.6	INCH
SKERRIES (MILVERTON HALL)	533410	60830	DUBLIN	1908	1	1	1940	12	31	9.1	INCH
STILLORGAN (FARMLEIGH)	531655	61138	DUBLIN	1915	1	1	1935	12	31	4.8	INCH
STRAFFAN HSE	531830	63730	KILDARE	1883	1	1	1940	12	31	0.1	INCH
STROKESTOWN (CASTLENODE)	534540	80500	ROSCOMMON	1908	1	1	1940	12	31	0.0	INCH
TURRAUN (BORD NA MONA)	531540	74450	OFFALY	1913	1	1	1940	12	31	10.8	INCH
UCC (UNIV.COLL.CORK)	515359	82910	CORK	1910	1	1	1940	12	31	3.3	INCH
VALENTIA (OBSERVATORY)	515617	101436	KERRY	1875	1	1	1950	12	31	0.0	IN/MM
WATERFORD(BROOK LODGE)	521345	70205	WATERFORD	1876	1	1	1931	12	31	20.6	INCH
WATERFORD (TYCOR)	521510	70750	WATERFORD	1888	1	1	1939	12	31	15.5	INCH
WESTPORT(MURRISK ABBEY)	534655	93823	MAYO	1900	1	1	1904	12	31	0.0	INCH
WESTPORT (ST. HELENS)	534726	93420	MAYO	1903	1	1	1920	12	31	11.5	INCH
WESTPORT HSE	534800	93200	MAYO	1911	1	1	1940	12	31	3.1	INCH

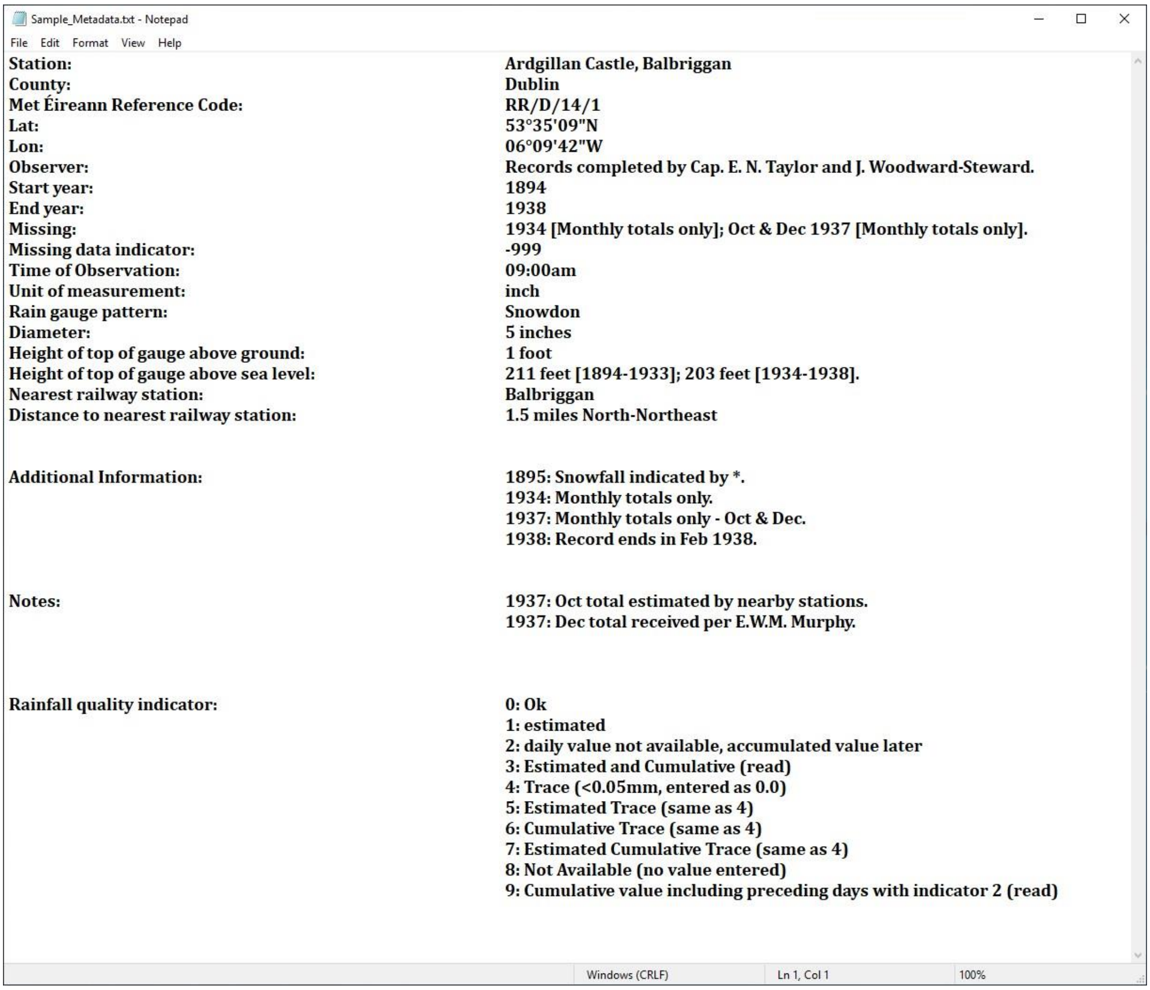
3.2.4 Metadata

In addition to the rainfall observations, metadata for each station was extracted from the individual station records and recorded in separate text files (Figure 3.6). These files provide information on station name and location, observer, record length, missing data, diameter of gauge, changes in gauge height, gauge pattern, time of observation, unit of measurement and distance to nearest railway station. A section for 'Additional Information' provides details of observations made while transcribing the data, for instance, the presence of a multiday accumulation. A separate section for 'Notes' provides a transcription of any handwritten notes recorded on the original record by the observer. For example, an observer may note an exceptional rainfall event or a leak in the gauge. These notes are transcribed verbatim.

The British Meteorological Office coordinated the meteorological network by supplying instruments and instructions for observations. Nevertheless, before standard equipment and procedures for meteorological observations were introduced from the late-nineteenth century to mid-twentieth century, the design and placement of rain-gauges varied considerably. Observer bias, instrument changes, sampling periods, as well as external factors relating to site exposure (e.g. proximity to buildings) affect the accuracy and consistency of observations (Green et al., 2008; Daly et al., 2007; Kunkel et al., 2005). In general, such changes were recorded on the station record and subsequently transcribed to the metadata files. Detailed metadata presented here provides a comprehensive account of station changes and can be used as an aid in determining the reliability of station records. This information will be particularly valuable in helping to explain the presence of any abrupt shifts identified when further quality assurance and homogenisation techniques are applied (Aguiler et al., 2003).

3.3 Error checking

At each stage of the transcription process, quality assurance measures were employed to preserve the integrity of the data being rescued. Keying guidelines were developed ensuring conformity to World Meteorological Organisation (WMO) standards (WMO, 2016). Monthly totals were examined against the derived sum of the daily entries to identify potentially incorrect data entries. The data was double keyed and the entries from different transcribers compared. Where the entries agreed, the value was provisionally accepted as the raw data value. If the values disagreed, the original record was manually examined to ascertain the true observed value. An examination of errors across all transcriptions revealed a percentage error of less than 1%. Multiday accumulations were identified and flagged using the original records as a reference. A description of numerical flag values is included in the metadata files. These indicator flags will facilitate the re-distribution of multiday accumulations to the respective days on which no observation was recorded. This will be undertaken prior to the application of further, more sophisticated, quality-assurance techniques. As a final check for transcription errors, the upper and lower 1% of observations (non-zero precipitation) were examined for each individual station record. Values identified as outliers were cross-checked against the original record.



Sample_Metadata.txt - Notepad

File Edit Format View Help

Station:	Ardgillan Castle, Balbriggan
County:	Dublin
Met Éireann Reference Code:	RR/D/14/1
Lat:	53°35'09"N
Lon:	06°09'42"W
Observer:	Records completed by Cap. E. N. Taylor and J. Woodward-Steward.
Start year:	1894
End year:	1938
Missing:	1934 [Monthly totals only]; Oct & Dec 1937 [Monthly totals only].
Missing data indicator:	-999
Time of Observation:	09:00am
Unit of measurement:	inch
Rain gauge pattern:	Snowdon
Diameter:	5 inches
Height of top of gauge above ground:	1 foot
Height of top of gauge above sea level:	211 feet [1894-1933]; 203 feet [1934-1938].
Nearest railway station:	Balbriggan
Distance to nearest railway station:	1.5 miles North-Northeast
Additional Information:	1895: Snowfall indicated by *. 1934: Monthly totals only. 1937: Monthly totals only - Oct & Dec. 1938: Record ends in Feb 1938.
Notes:	1937: Oct total estimated by nearby stations. 1937: Dec total received per E.W.M. Murphy.
Rainfall quality indicator:	0: Ok 1: estimated 2: daily value not available, accumulated value later 3: Estimated and Cumulative (read) 4: Trace (<0.05mm, entered as 0.0) 5: Estimated Trace (same as 4) 6: Cumulative Trace (same as 4) 7: Estimated Cumulative Trace (same as 4) 8: Not Available (no value entered) 9: Cumulative value including preceding days with indicator 2 (read)

Windows (CRLF) Ln 1, Col 1 100%

Figure 3. 6: Sample metadata file providing information regarding station name and location, observer, record length, missing data, diameter of gauge, changes in gauge height, gauge pattern, time of observation, unit of measurement and distance to nearest railway station. A section for 'Additional Information' provides details of observations made while transcribing the data, for instance, the presence of a multiday accumulation. A separate section for 'Notes' provides a transcription of any handwritten notes recorded on the original record by the observer.

3.4 Dataset use

The data are freely available from the edepositIreland data centre (http://hdl.handle.net/2262/91347). The dataset comprises daily rainfall data for 114 stations at various locations throughout Ireland for varying time periods. Individual station folders contain two files: a data file in ASCII format and a corresponding metadata text file as described in section 3.2.4. Data files consist of five columns providing the observation date (year, month, day), followed by rain value and indicator value. The indicator value provides information about the nature of the observation and identifies multiday accumulations. A key to the respective indicator values is provided in the metadata files. Rainfall values run continuously from start date to end date of the data recovery period, with missing values denoted using a -999 indicator. Work is currently underway to produce a network of long-term, quality assured daily rainfall stations using the datasets whose generation is described in this paper. Preliminary quality assurance checks have been applied to assess the accuracy of the transcription process. A second, comprehensive set of quality assurance techniques will be applied to detect both systematic and non-systematic errors, this will be described in a subsequent paper. Post 1940 daily rainfall records are readily available from Met Éireann's climate database. The newly transcribed data will be added to the database and used to extend these station records back to the late 19th Century. Once joined, the full series will be homogenized and analysed to assess variability and changes in the characteristics of rainfall events over the long-term record.

3.5 Future Work on Data Rescue

The importance of historical climate data is being increasingly acknowledged for its role in supporting effective climate risk management through reanalysis and validation of climate models. As a consequence, climate data rescue has experienced a substantial rejuvenation in recent years, with a number of national and international projects underway e.g. Copernicus Climate Change Data Rescue Service (C3S) and the International Data Rescue (I-DARE). The data presented here mark a significant and innovative effort to progress data rescue efforts for Ireland. Nevertheless, a considerable amount of data have not yet been digitised and exist only in hard copy format. Met Éireann, as part of their current operational business plans, has considered opportunities to advance data rescue initiatives within the organisation and through collaboration with other agencies to create a comprehensive climate data bank and to facilitate the creation of high-quality data products. Met Éireann are involved in a number of ongoing projects and initiatives aimed at enhancing climate data availability and accessibility. These include a collaboration with The Central Statistics Office (CSO) to transcribe all parameters and metadata from the eight meteorological registers imaged during the course of this work (see Table 3.2 for station details). As part of this, transcription of the entire Phoenix Park series has recently been completed. The data series which spans the period 1829 to present is the longest continuous series for the Republic of Ireland and the second longest in Ireland after Armagh Observatory. The extent of the data available for the Phoenix Park makes it one of the most comprehensive series available worldwide.

Engagement of non-experts or 'citizen scientists' on a voluntary basis has become increasingly significant to the rescue and refinement of observational data across multiple scientific disciplines (Bonney et al., 2014). The success of ongoing citizen science applications - for example, OldWeather.org (www.oldweather.org/), the Weather Rescue Project (https://weatherrescue.wordpress.com/) and Data.Rescue@Home (http://www.data-rescue-at-home.org) underscores the potential of crowdsourcing as a data rescue strategy. Further, the data rescue project developed by Ryan et al. (2018) as a collaboration between Met Éireann and Maynooth University will continue as an integral part of the Climate Change module delivered by Prof. Murphy. It is hoped that the project will be developed further through the use of an online platform designed to host Ireland's historical meteorological records. Such an application would facilitate the extension of the data rescue project to other teaching programmes and significantly advance contributions from citizen scientists.

Climate data rescue must be viewed as a continuous, long-term activity (Brönnimann et al., 2018). The priority for Met Éireann over the coming years is to catalogue and image all historical records currently held in the archives for integration into the climate repository, an effort enhanced by the development of resources (including a digital scanner) as part of this work. Continued effort is also being made to recover records held in other libraries and institutes in Ireland and abroad. Met Éireann have recently recruited an archivist to document the vast historical record holdings including meteorological and climatological collections, weather diaries, monthly bulletins, annual reports, weather maps and site inspection reports. Providing access to these records, along with an inventory of what is available is important to promote collaborations and knowledge-sharing between interested stakeholders.

4. Long-term trends in extreme precipitation indices in Ireland

4.1 Introduction

Understanding long-term variability and trends in precipitation is critical for water resource management and assessing potential flood risk. Long-term series of precipitation have been produced for many regions to detect changes in the spatial and temporal variation of this essential climate variable (e.g. Ashcroft et al., 2018; Brunet et al., 2014; Auer at al., 2007; De Jongh et al., 2006; Moberg et al., 2006). In Ireland, previous research has focused on the development and analysis of long-term monthly precipitation series (e.g. Murphy et al., 2018; Noone et al., 2015), while analysis of daily precipitation has largely been limited to the post 1940 period, for which digitised observational and gridded precipitation products are available. Moreover, no previous analysis has assessed long-term changes in precipitation extremes for Ireland using the suite of indices defined by the CCl/ CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI).

Previous analysis of Irish annual precipitation totals reveals an increase in mean annual precipitation of approximately 5% in 1981–2010, compared with the 30-year period 1961–1990 (Walsh, 2016). In general, the highest rainfall accumulations occur in the western half of the country and at higher elevations, with rainfall decreasing towards the northeast (Figure 4.1). Most of the eastern half of the country gets between 750 and 1000 (mm) of rainfall per year. Rainfall in the west generally averages between 1000 and 1400 mm. In many mountainous districts rainfall amounts exceed 2000mm per year. An increase in the frequency of heavy rain days (days with rainfall greater than 10 mm) has been reported at the majority of stations analysed over the period 1961-2010, however, regional variations are evident with occasionally conflicting trends from stations that are geographically relatively close (Walsh and Dwyer, 2012). Sheridan (2001) reported an increase in annual accumulations over the period 1941-1999, particularly in the west of the country, with more notable increases reported since the early 1970s. McElwain and Sweeney (2007) analysed monthly data at Birr and Malin Head from 1890 to 2003 and found significant increases in annual totals at Malin Head with no significant trends detected in the annual precipitation series at Birr. Leahy and Kiely (2011) examined changes in short-duration (1hr-24hr) precipitation extremes and found strong evidence for increasing trends since the late 1970s, particularly in the west of the country, coinciding with changes in the North Atlantic Oscillation.

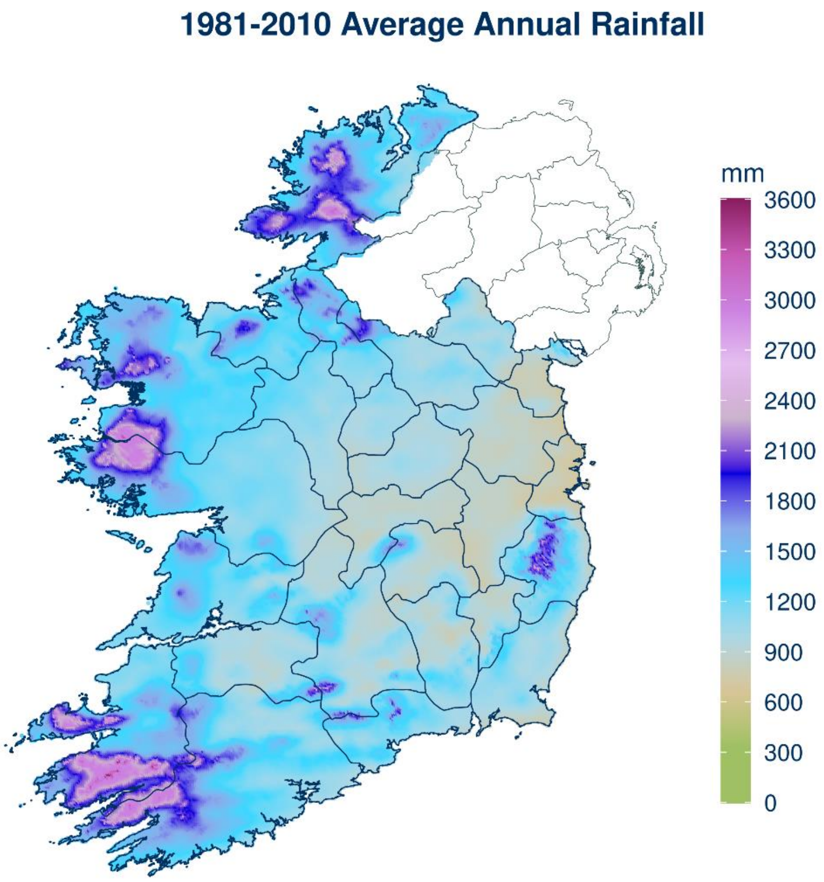


Figure 4. 1: Annual average rainfall (mm) accumulation over Ireland for the period 1981-2010 (derived from Met Éireann data).

Knowledge of changes in extreme precipitation events are a key aspect of monitoring climate change as increases or decreases in the frequency and/or magnitude of extreme precipitation events have high environmental and socio-economic impacts. However, the relative sparseness of long-term digitally available records at appropriate spatial and temporal scales hampers our ability to fully understand changes in precipitation extremes (Brunet and Jones, 2011). The spatial variability of precipitation, together with changes in observer practices, gauge location and design, mean that developing reliable, long-term precipitation series is often a challenging task (Wilby et al., 2017). Quality control procedures are crucial to identify both systematic and non-systematic errors that can confound the detection and interpretation of trends. Errors can occur as a result of changes in measuring techniques, observational and recording practices, and transfer of data from paper to digital format (Slater et al., 2020). A well-defined quality control process should be able to flag data errors that could compromise the analysis of natural climate variability, particularly in the study of extreme events (Llabrés-Brustenga et al., 2019). Homogenisation of data to remove spurious non-climatic features from long-term instrumental records is integral to the development of accurate climate data and the assessment of climate variability (Freitas et al., 2013; Vertačnik et al., 2015). While many methods have been developed to evaluate the homogeneity of monthly to annual-resolution climate data, homogenisation of daily data is a highly complex task that is still under development within the research community (Venema et al., 2018).

Recently, Ryan et al. (2020) completed the rescue of 3616 station years of daily precipitation data (~1.32 million daily values) for stations across Ireland. Details of the data rescue process and dataset development are presented in previous publications (Ryan et al., 2018; Ryan et al., 2020) to provide a fully traceable dataset. The availability of this data now makes it possible to assess changes in daily precipitation extremes back to the early 20th Century. While Ryan et al. (2020) implement basic quality assurance tests, in this paper we seek to conduct a more rigorous quality control of rescued data before extracting and homogenising long-term daily series for stations across Ireland. We then derive indices of annual extremes and assess these for evidence of trends.

The remainder of the paper is organized as follows: Section 2 provides a description of the data, the quality control (QC) and homogenisation techniques used to reduce errors in the dataset and methods used to test the direction and magnitude of trends. Section 3 presents the results, while Section 4 provides a discussion of key findings, limitations and directions for future work, before drawing conclusions in section 5.

4.2 Data and Methods

4.2.1 Quality control of data

The dataset used in this analysis was constructed from recently rescued and digitised historical (pre-1940) daily precipitation data (Ryan et al., 2020) and post 1940 daily precipitation data extracted from Met Éireann’s database. The latter contains daily and monthly precipitation records from 1941 onwards which have previously undergone quality control tests, full details of which are provided by Walsh (2016). These include a combination of visual checks, spatial techniques, cross-validation of daily interpolated values, comparison with daily radar accumulations, where available, and an automated system for the re-distribution of cumulative daily values. The pre-1940 dataset comprises daily precipitation data for 114 stations at various locations throughout Ireland for varying time periods (Figure 4.2).

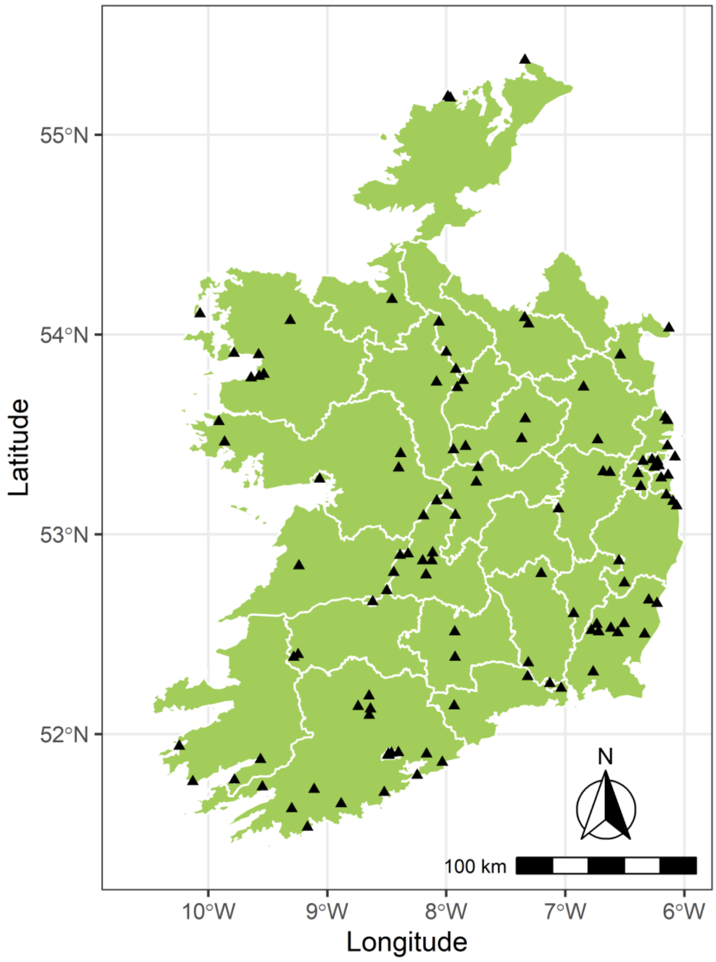


Figure 4. 2: Location of stations for which pre-1940 daily precipitation data were rescued and digitised. Individual folders containing a data file and metadata file are available for each station (Ryan et al., 2020).

Ryan et al. (2020) undertook an initial quality assessment of the pre-1940 digitised data. This involved manually cross-checking all flagged observations against the original record and metadata. At each stage of the transcription process, quality assurance measures were employed to preserve the integrity of the data being rescued. Keying guidelines were developed ensuring conformity to World Meteorological Organisation (WMO) standards (WMO, 2016). Monthly totals were compared to the derived sum of the daily entries to identify potentially incorrect data entries. The data were double keyed and the entries from different transcribers compared. Where the entries agreed, the value was provisionally accepted as the raw data value. If the values were different, the original record was manually examined to ascertain the true observed value. An examination of errors across all transcriptions revealed a percentage error of <1%. As a final check for transcription errors, the upper and lower 1% of observations (non-zero precipitation) were examined for each individual station record. Values identified as outliers were cross-checked against the original record. Following the application of these initial quality assurance checks the raw data were published to enable users to access the original data (Ryan et al., 2020). Additional quality control checks are undertaken here. These can be grouped into four categories: basic integrity tests, tolerance tests, temporal consistency tests, and spatial consistency tests. Details of the QC procedures deployed are provided in Table 4.1. Each test produced a list of values flagged at each station which were subject to manual inspection. The detection of false positives (e.g. extreme events flagged as outliers) is a common issue in QC assessments, however, manual inspection, combined with local/regional knowledge of climatological processes ensured that these important observations were retained.

Table 4. 1: Summary of Quality Control tests applied to the pre-1940 data. * Tests applied using RClimdex_extraQC software (Aguilar & Prohom, 2011).

QC test	Description
Monthly cross-check	Flag values when sum of daily rr data does not equal monthly rr total
Non-numeric values	Check file structure
Physically impossible values	Flag values where rr <0
Duplicate date control*	Flag any sequence of dates that appear more than once in time series. RClimdex_extraQC software (Aguilar & Prohom, 2011).
Unrealistic values*	Flag out of range values based on fixed threshold of 243.5mm i.e., highest daily total recorded in Ireland 18th Sep 1993 at Cloone, Co.Kerry. RClimdex_extraQC software (Aguilar & Prohom, 2011).
Anomalous sequence of zero rr	Flag if zeros persist for ≥ 1-month duration
Rounding problems evaluation*	Visual assessment to identify potential biases in the rounding of decimal values. RClimdex_extraQC software (Aguilar & Prohom, 2011).
Accumulated values	Flag months with only one value recorded which is in excess of 2 times the mean daily rr intensity for that month

Climatological outliers*	Flag outliers based on IQR exceedance: percentile-based approach and therefore suitable for asymmetric distribution. RClimdex_extraQC software (Aguilar & Prohom, 2011).
Temporal consistency	Flag outliers based on limits determined from a multiple of the interquartile range (IQR) calculated for each station/month. Outliers are identified using the sample distribution of each calendar month separately for each station. An outlier is flagged when $X_i - {}_q50 > f(\text{IR})$ where X_i is the monthly mean of the year i, ${}_q50$ is the median, and *f* is the multiplication factor. A value of $f = 4$ was used as the multiplicative factor as in Eischeid et al. (1995).
Spatial consistency	Check consistency with neighbouring stations: Number of dry days: flag months with 2 days more/less than max/min of nearest neighbour. Number of wet days (rr >= 0.2mm): flag months with 2 days more/less than max/min of nearest neighbour. Number of days with >= 5mm: flag months with 2 days more/less than max/min of nearest neighbour. Number of days with >=10mm: flag months with 2 days more or less than min/max nearest neighbour. Total Monthly Fall: (normalised as a fraction of the long-term average) flag months with ± 25% of the max/min normalised rr of nearest neighbours i.e., > 1.25 max of NN or < 0.75 min of NN.

For the majority of values flagged by the basic integrity tests, the value was adjusted following manual inspection, for example, in cases where zeros were used in place of the missing data indicator or undocumented cumulative totals. Values classified as suspect were accepted if they satisfied each of the following criteria: the value agreed with that recorded on the original record; similar observations were recorded at neighbouring stations; the value was deemed physically reasonable for that station/region/season. Values were rejected and set to missing if they were flagged as suspect and did not satisfy these criteria or if there was evidence reported in the metadata to suggest that the gauge was defective at the time of observation. In total, 59,274 values were flagged and investigated (~4.5% of the data), with slightly less than 1% of the values removed from the quality-controlled version of the dataset used to assess trends. Cumulative daily values which were identified and flagged during the transcription process were re-distributed to the respective days on which no observation was recorded. For cumulative values requiring redistribution, a first guess value for days on which no observation was recorded is estimated by means of inverse distance weighted interpolation; the estimated daily amounts are summed and the ratio of this total to the observed cumulative total is calculated; the final estimated daily values are obtained by adjusting the first guess values according to this ratio (Walsh, 2016).

4.2.2 Deriving long-term station series

The quality controlled pre-1940 rescued data was added to the database containing post 1940 data. For stations with matching station identifiers and locations, the pre-1940 data was joined to the post-1940 data to produce continuous station series. The method developed at Met Éireann by Walsh (2016) for developing gridded datasets was used to infill and extend station series using rescued data. First a regression-kriging interpolation model was used to generate monthly values for all stations in the database. For stations that had gaps in data, the monthly total was estimated using nearby station data. Next, for missing daily values, inverse distance weighted interpolation was used to disaggregate monthly interpolated values to provide a complete series of daily values. This method performed well in a previous assessment by Walsh (2016) using data from the Irish precipitation network. Finally, stations for analysis were selected. Given the lower station density in the pre-1940 period, only sites with neighbouring station data available were selected when developing long-term series. In total 36 stations were identified based on record length, continuity, completeness and proximity of pre-1940 station locations. The mean record length is 118 years. The longest series length is 146 years for Foulkesmill (Longraigue), the shortest is 91 years for Portumna O.P.W. The selected stations (Table 4.2) were grouped into two categories as follows:

- Group A stations are those with continuous observations at the same site from start year to end year. Missing data were infilled using the output of the regression kriging interpolation model. The average amount of missing data across all group A stations is 4.4%. Valentia and Phoenix Park stations have 0% missing data over the period of record. The maximum amount of missing data for any station in group A is 13.6% for Kenmare.

- Group B stations represent those from the post-1940 dataset for which nearby (<15km) rescued data were available to extend records into the pre-1940 period. B stations were generated using the output of the regression kriging interpolation model.

Table 4. 2: Details of selected long-term stations including station ID, station name, location, height (m) above msl, start and end year of series, series category.

st_id	station_name	lat	lon	height	start	end	group
108	Foulkesmill (Longraigue)	52.31056	-6.76639	71	1874	2019	A
175	Phoenix Park	53.36361	-6.34972	48	1881	2019	A
417	Inagh (Mt.Callan)	52.84167	-9.23833	122	1908	2019	A
1075	Roches Point	51.79306	-8.24444	40	1883	2019	A
1275	Markree	54.175	-8.45556	34	1875	2019	A
1519	Meelick (Victoria Lock)	53.16667	-8.08056	39	1910	2019	A
1529	Drumsna (Albert Lock)	53.91111	-8.00000	45	1903	2019	A
1575	Malin Head	55.37194	-7.33917	20	1885	2019	A
2375	Belmullet	54.22750	-10.00694	9	1897	2019	A
1819	Portumna O.P.W.	53.09167	-8.19167	35	1929	2019	A
1929	Athlone O.P.W.	53.42222	-7.94167	37	1902	2019	A
2012	Cashel (Ballinamona)	52.51056	-7.92861	80	1911	2019	A
2275	Valentia Observatory	51.93833	-10.24083	24	1875	2019	A
1812	Waterford (Tycor)	52.25278	-7.13056	49	1890	2019	A
201	Glengarriff (Ilnacullin)	51.73472	-9.54583	7	1914	2019	A
603	Kenmare (Dereen)	51.76944	-9.78056	24	1912	2019	A
1923	Glenasmole D.C.W.W.	53.23889	-6.36667	158	1900	2019	A
8212	Portlaw (Mayfield II)	52.29083	-7.30083	8	1900	2019	A
675	Ballyhaise	54.05139	-7.30972	78	1900	2019	B
706	Mallow (Hazelwood)	52.19028	-8.65000	94	1900	2019	B
875	Mullingar	53.53722	-7.36222	101	1910	2019	B
944	Creeslough (Carrownamaddy)	55.13333	-7.95000	88	1908	2019	B
1338	Omeath	54.08667	-6.25639	12	1925	2019	B
1375	Dunsany	53.51583	-6.66000	83	1900	2019	B

1475	Gurteen	53.05306	-8.00861	75	1900	2019	B
2115	Hacketstown (Voc.Sch.)	52.86111	-6.55278	189	1918	2019	B
2528	Ballyforan (Bord na Mona)	53.43972	-8.30333	47	1925	2019	B
4015	Enniscorthy (Brownswood)	52.46306	-6.56083	18	1900	2019	B
4513	Kilkenny (Lavistown House)	52.63611	-7.19722	58	1900	2019	B
5131	Kilskyre (Robinstown)	53.69278	-6.96333	87	1900	2019	B
6019	Killaloe Docks	52.81	-8.44889	40	1902	2019	B
6329	Strokestown (Carrowclogher)	53.753	-8.10800	52	1908	2019	B
3310	Abbeyfeale (Caherlane)	52.35194	-9.28361	155	1925	2019	B
1175	Newport	53.92222	-9.57222	22	1910	2019	B
1433	Westport (Carrabawn)	53.7925	-9.52722	56	1909	2019	B
2227	Carndolla	53.40306	-9.01556	24	1900	2019	B

4.2.3 Homogenisation of derived station series

Relative homogenisation methods (i.e. testing with respect to a homogenous neighbouring station) are considered more robust than absolute methods, provided station records are sufficiently correlated (Gubler et al., 2017). However, the high spatial variability of daily precipitation make it difficult to find suitable reference series except in the case of parallel measurements (Wang et al., 2010). Moreover, given the sparse spatial density of the station network in the early period (Figure 4.2), the current homogeneity assessment is restricted to the application of absolute methods. We employ a two-step approach. First, for each station, breakpoints in the monthly series, derived from daily precipitation values, are detected and adjusted using the RHtests software developed by Wang and Feng (2013) at the Climate Research Division, Atmospheric Science and Technology Directorate, Environment Canada (http://etccdi.pacificclimate.org/software.shtml). Second, the output from the RHtests software is examined for further breaks by testing the homogeneity of the derived annual indices using the standard normal homogeneity test (SNHT) (Alexandersson, 1986) and the Pettitt test (Pettitt, 1979). This approach supports the selection of series for the analysis of trends.

The RHtests_dlyPrcp software package is specifically designed for homogenisation of daily precipitation data which are non-continuous, non-negative, and non-normally distributed. Breakpoints were detected in the monthly log-transformed series using the RHtests software which is based on the penalized maximal *t* test (Wang et al., 2007) and the penalized maximal *F* test (Wang, 2008b). These are embedded in a recursive testing algorithm (Wang, 2008a), with the lag-1 autocorrelation of the time series being empirically accounted for. The software facilitates metadata integration by testing both known and unknown breakpoints; hence significant breakpoints (0.05 level) were identified using a combination of the statistical methods and station reports. Specifically, the software was first run to detect all breakpoints that could be significant at the nominal level even without metadata support, referred to as Type-1 breakpoints. Next, a second function was run to detect additional breakpoints (Type-0 breakpoints) which are deemed significant only if supported by reliable metadata. For each station series, the available metadata was investigated to determine whether there was evidence at or near the identified breakpoint dates to support the shifts. Only those Type-0 breakpoints that were supported by metadata, along with all Type-1 changepoints were retained.

Before adjusting the daily series, discontinuities in the occurrence frequency of precipitation were investigated. Such discontinuities result from changes in the unit of measurement, measuring precision, and observing practices (Wang et al., 2010). For each station, the series of daily precipitation amounts greater than a given threshold ($>P_{thr}$) was examined by varying P_{thr} over a set of small values that reflect changes in the measuring precision (i.e., 0.2mm, 0.3mm, 0.5mm). This assessment revealed the presence of discontinuities in the frequency of small measured amounts in the pre-1940 record, most likely related to changes in the unit of measurement (i.e. from imperial to metric). Given that the objective of the present study is to examine trends in indices which require daily precipitation >=1mm, adjustments were applied to the series of $>P_{thr}$ daily precipitation data that was found to be free of frequency discontinuity. This P_{thr} varies across station series but does not exceed $P_{thr} = 0.5$mm. It should be noted that all $<P_{thr}$ values in the series are left unadjusted, and thus frequency discontinuities remain, affecting values between 0 and P_{thr} inclusive.

Where breaks were detected, adjustments were applied to daily station series by means of the quantile mapping (QM) algorithm in the RHtests software (Wang et al., 2010). The algorithm was applied as follows: the non-zero daily precipitation series were detrended using the linear trend estimated from a multiphase linear regression fit that accounts for the mean shifts at detected breakpoints in the data series (Wang, 2008b). For each breakpoint to be adjusted, the data in the segment immediately before and after the breakpoint is used to estimate the probability distribution function (PDF). We use all available data in a segment to estimate the PDF, as estimates of the PDF (and hence the QM adjustments) using shorter periods (e.g. < 30 years) can contain large sampling uncertainty (Wang et al., 2010). Each segment is divided into Mq quantile categories, and for each quantile, the differences between the means of the two periods is computed. A natural cubic spline is then fitted to the Mq category-mean differences between the segment to be adjusted and the base segment which is then used to estimate the QM adjustments needed to make the data series homogeneous. The choice of the number of quantile categories (Mq) used to estimate the spline to derive the QM adjustments was set at Mq = 4 ~ 16, depending on the length of the shortest segment in a given data series. We use the most recent segment as the base on the assumption that technologies and observational practices generally improve over time. We recognise that the choice of parameters is subjective and may not be optimal for all series, however, the use of detailed metadata to inform decisions increases confidence in the performance of the algorithm, while sensitivity analysis to the choice of different parameters resulted in minor changes to adjustments made.

Following Santo et al. (2014), further tests of the adjusted series were carried out by testing the homogeneity of the derived PRCPTOT and R5mm (see next section) annual indices using the standard normal absolute homogeneity test (SNHT) and the Pettitt test. The p-value of potential breakpoints were plotted to facilitate a relative assessment of breakpoints across all stations. To determine the significance of potential break points, the null hypothesis (no breakpoint detected) was tested against the alternative at the 0.05 level. As a final assessment, a visual inspection was carried out by plotting the original and the adjusted annual precipitation totals for each series found to be inhomogeneous.

4.2.4 Indices and Trend detection

The derived quality controlled, homogenised long-term daily series were assessed for evidence of trends in the characteristics of extreme precipitation. For this purpose, eleven indices defined by the CCl/ CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI, http://cccma.seos.uvic.ca/ETCCDI/) were calculated for each series and investigated for evidence of trends. The selected indices allow examination of changes in intensity, frequency, and duration of extreme precipitation events (Alexander et al., 2006). Precipitation indices are summarised in Table 4.3. For deriving indices, we define a wet day as a day with ≥1 mm precipitation, while the common 30-year base period used to define thresholds for percentile-based indices is 1961-1990.

Table 4. 3: List of extreme precipitation indices used in this study. All indices are calculated annually.

ID	Indicator name	Indicator definitions	Units
RX1day	Max 1-day precipitation amount	Annual maximum 1-day precipitation	mm
RX5day	Max 5-day precipitation amount	Annual maximum consecutive 5-day precipitation	mm
SDII	Simple daily intensity index	Ratio of annual total precipitation to number of wet days (≥1 mm)	mm/day
R5mm	Number of days when precipitation ≥5 mm	Annual count when precipitation ≥5 mm	days
R10mm	Number of heavy precipitation days	Annual count when precipitation ≥10 mm	days
R20mm	Number of very heavy precipitation days	Annual count when precipitation ≥ 20 mm	days
CDD	Consecutive dry days	Maximum number of consecutive days when precipitation <1 mm	days
CWD	Consecutive wet days	Maximum number of consecutive days when precipitation ≥1 mm	days
R95pTOT	Very wet days	Annual sum of precipitation on days when precipitation exceeds the 95th percentile of daily precipitation in the base period.	mm
R99pTOT	Extremely wet days	Annual sum of precipitation on days when precipitation exceeds the 99th percentile of daily precipitation in the base period.	mm
PRCPTOT	Annual total wet-day precipitation	Annual total precipitation from days ≥ 1 mm	mm

Indices were derived for two set periods: 1910-2019 and 1940-2019. The earlier period was selected as the longest common period for which all stations have available records, the shorter timespan was selected to represent the period of available digital records prior to our data rescue efforts. Evidence for monotonic trends was assessed using the Modified Mann-Kendall (MK) test (Hamed and Rao, 1998), a non-parametric rank-based method. The standardized MK test statistic (MKZs) follows the standard normal distribution with a mean of zero and variance of one. A positive (negative) value of MKZs indicates an increasing (decreasing) trend. Statistical significance was evaluated with probability of Type I error set at the 5 % significance level. A two tailed MK test was applied, hence the null hypothesis of no trend (increasing or decreasing) is rejected when $|MKZs| > 1.96$. The

MK test requires data to be independent (i.e. free from serial correlation) as positive serial correlation increases the likelihood of Type 1 errors or incorrect rejection of a true null hypothesis. Therefore, application of the Modified MK test includes a test to detect positive lag-1 serial correlation at the 5% level using the autocorrelation function (ACF). The existence of trend influences the correct estimate of serial correlation; therefore, original time-series are detrended to form a 'trend-removed' residual series before the ACF is applied (Hamed and Rao, 1998). The magnitude of trend (β) is estimated using the approach proposed by Theil (1950) and Sen (1968), hereinafter referred to as TSA. To facilitate a relative comparison among different sites, the approach presented by Yue and Hashino (2007) is utilised where the magnitude of trend TSA_{rel} (%) for each time-series is expressed as a percentage change over the period of record of n years relative to the mean (μ) for the period, given by:

$$\text{TSArel}\ (\%) = \left(\frac{\beta \times n}{\mu}\right) \times 100$$

For brevity, Tables 4.6 and 4.7 present the direction of change and proportion of statistically significant (5 % level) trends along with the median magnitude of trend (TSA_{rel} %) and the median number of days per year for the 11 extreme precipitation indices for both fixed periods. Full details of trend statistics calculated for each station and indicator are provided as Supplementary Information.

The persistence or dependency of trends on period of record was examined for each station by systematically reducing the start year of analysis from the whole record to a minimum of 30 years following the approach of Murphy et al. (2013) and Noone et al. (2015). For each indicator, the MKZs statistic was derived first for the full record e.g. 1890-2019, then 1891-2019, and so on until 1990-2019. MKZs values were plotted for each iteration of start years to examine the temporal evolution of trend throughout the period of record. Unlike the fixed period analysis above, the persistence of trends was evaluated for the full record available at each station.

4.3 Results

4.3.1 Homogenisation

Following application of the RHtests software, twenty stations were found to be homogenous and did not require adjustment. In total, twenty-two breaks were detected in sixteen of the 36 stations. Multiple breaks were found in five records: three in the Newport series; two each in the Malin Head, Belmullet, Glengarriff, and Drumsna series; and a single break point in each of the remaining eleven series. Approximately 80% of detected breakpoints are supported by information contained in the metadata regarding instrument and site changes. Details extracted from metadata for each identified break are presented in Table 4.4.

Table 4. 4: List of breakpoints detected by RHtests software and supporting information extracted from metadata files.

ID	Station	Breakpoint	Reason
1929	Athlone	1926	Reports of gauge leaking resulting in low readings. Recommendation for new gauge to be installed, however, no documentation to say that this was carried out. The height of the gauge above ground was changed from 2ft to 1ft in Jan 1928.
2528	Ballyforan	1976	Ballyforan station commenced observations in 1976. Record extended to 1925 using nearby station data.
675	Ballyhaise	1923	Ballyhaise station commenced observations in 1923. Record extended to 1900 using nearby station data.
2375	Belmullet	1956; 1984	New station established in September 1956. No documented reason for break detected in 1984. Significant (prob >0.99) changepoint detected in 1983 by Leahy and Kiely (2011) using the Mann-Pettitt-Whitney test on annual precipitation total over the period 1957-2008.
944	Creeslough	1928	Creeslough record was extended to 1908 using nearby station data. Dunfanaghy is among the neighbouring stations available in the pre-1940 period. Dunfanaghy was established as a rainfall station in 1880, the first location being St. Helen's (55°11'N; 7°58'W). In July 1926, the station was moved to St. Patrick's, a site which had been set up approx. a quarter of a mile from St. Helen's. The record at St. Patrick's was considered a continuation of the observations taken at St. Patrick's.
1529	Drumsna	1917; 1942	No documented reason for break detected in 1917 but this break was also detected by HOMER in previous work by Noone et al. (2015). New gauge installed in 1942. Comparison of the old site over a 13-month period showed the new gauge recording 154% of the old site.
201	Glengarriff	1931; 1942	In 1929 the height of the gauge changed from 15ft to 25ft above msl. The height of the gauge above ground also changed from 1 ft to 2ft 6in at some point during the period 1929-1936. In December 1942 the site was moved to a new site. The site was inspected in 1950 and the report states that the new site satisfies conventions.

6019	Killaloe	1959	Killaloe station was established in 1983. Record extended to 1900 using nearby station data. Killaloe OPW and Killaloe (Ballina) are among the neighbouring stations available prior to 1983. Killaloe OPW commenced in 1957 and ceased observations in 1984. Killaloe (Ballina) commenced in 1866 and ceased observations in 1959.
1575	Malin Head	1923; 1957	From 1885 readings were made at the telegraphic reporting station (Lloyds tower at a height of 230 ft). In Dec 1921 the station was moved to the coast guard station at a height of approximately 22ft above msl. New station established May 1955. Following the installation of the new site, comparison readings were taken which suggested discrepancies between readings taken at the old site and the newly established station.
706	Mallow	1993	New rain gauge installed and moved to new site due to concerns about trees that had grown high around the site.
1275	Markree	1883	The original square gauge (1 sq. yard) was positioned on top of the library. Over the period 1875-1881 a comparison 5" gauge positioned 6" above ground gave a correction multiple of 1.2045. In 1870 the diameter of the gauge is recorded as 24 inches. In 1883 the diameter of the gauge is recorded as 5 inches. No record of the gauge diameter is noted in the intervening years. In July 1884 the height of the top of the gauge above ground changed from 16.5 ft to 1 ft. In July 1884 the height of the top of the gauge above sea level changed from 148 ft to 130ft.
1519	Meelick	1944	No documented reason for break detected in 1944. Station inspection report states that no change of site took place prior to 1952. However, a report in 1942 documents concerns about over-exposure. It states that the ground around the gauge has been hallowed out saucer-like due to over-exposure. It states that the rain recorder is now accurate and recommends that a standard gauge be issued.
1175	Newport	1928; 1952; 1959	Newport TUCSON AWS was established in 2005. Prior to this, readings were taken at Newport (Furnace) which commenced observations in 1959. The record was extended to 1910 using available nearby station data.
1338	Omeath	1943	Omeath station commenced observations in March 1943. Record extended to 1925 using nearby station data.
1812	Tycor	1946	In March 1946 a rain-recorder and 5" Snowdon gauge (E.S.B. type) were installed on a new site further East (50 yds) and 10ft. lower than the existing gauge - which was over-exposed on the concrete roof of the service reservoir. Comparative readings were taken for a period of 2.5 years and the results showed that the ratio of the readings of the new gauge to those of the old gauge was 115:110. The new gauge was forthwith adopted as the official gauge.

2275	Valentia	1993	April 1993: trees removed 150m south of enclosure. May 1996: New observing system commenced.

To examine the impact of the adjustments on the station series, the original and adjusted annual precipitation totals were plotted for the 16 inhomogeneous series (Figure 4.3). The magnitude of the adjustments applied to the original data varies across station series and datum as per the QM approach, with larger (smaller) values undergoing a greater (lesser) adjustment. Table 4.5 provides details of the mean, minimum and maximum annual percentage difference between the original and the adjusted series.

Table 4. 5: Mean/min/max annual % difference between the original and the adjusted series for 16 inhomogeneous station series.

station_id	station_name	mean (%)	min (%)	max (%)
201	Glengarriff	7.4	4.8	11.5
675	Ballyhaise	8.4	0.1	9.6
706	Mallow_HW	2.7	0.2	4.4
944	Creeslough	-11.3	-12.5	1.3
1175	Newport	10.2	5.9	17.8
1275	Markree	18.4	1.9	21.2
1338	Omeath	14	3.5	16.4
1519	Meelick	11.2	3	12.5
1529	Drumsna	24.4	4.5	38.9
1575	Malin Head	17.2	8	26.6
1812	Tycor	12.8	2.1	13.9
1929	Athlone	16.8	15.4	18.4
2275	Valentia	8.8	5.5	9.3
2375	Belmullet	11.7	-9.9	25.6
2528	Ballyforan	11.6	2.3	13.1
6019	Killaloe	8.4	3.7	9.1

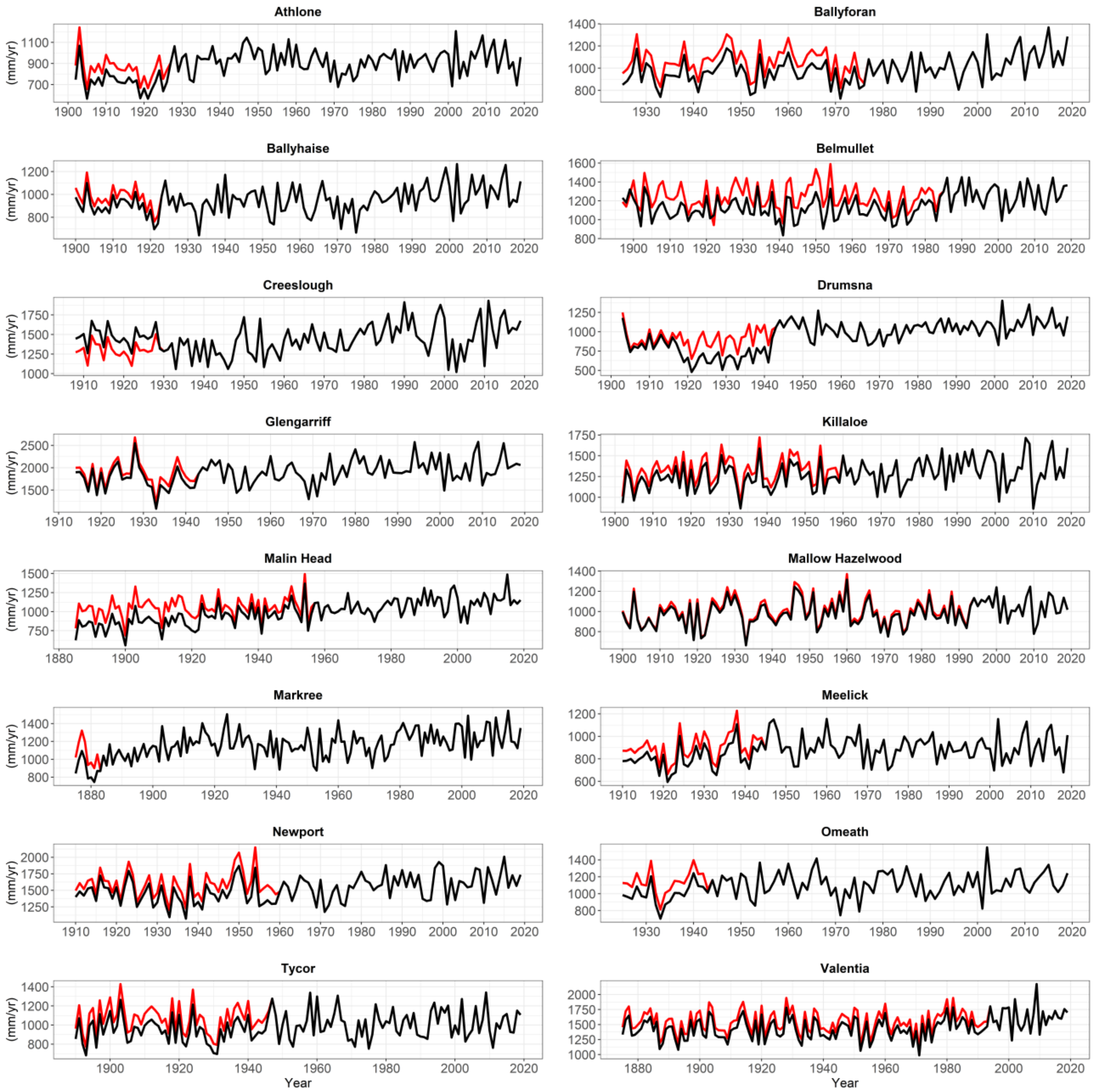


Figure 4. 3: Original (black) and adjusted (red) total annual precipitation for the 16 station series with breakpoints detected and adjusted using RHtests software. Units are millimetres per year.

The Pettitt test and the SNHT were applied to the annual series of PRCPTOT and R5mm to detect residual breakpoints following the application of the RHtests software. Across the 36 station series tested, the Pettitt test detected nine significant breakpoints in the PRCPTOT series. Five significant breakpoints were detected in the series of R5mm. The SNHT test detected eleven significant breakpoints in the PRCPTOT series, five of which coincide with breakpoints detected by the Pettitt test. Five significant breakpoints were detected in the R5mm series, four of which were also detected by the Pettitt test. Of the total 30 breakpoints detected in PRCTOT and R5mm by both tests, the timing for 21 are consistent across the network of stations. These occur in the 1970s and are consistent in timing with a shift to more positive NAO index resulting in stronger mid-latitude westerlies and increased precipitation over Ireland (Kiely, 1999). Further evidence for network wide changes during this period is provided by Coll et al. (2018) who tested the homogeneity of monthly precipitation series for Ireland using relative homogeneity methods, namely HOMER (Homogenisation Software in R) and ACMANT (Adapted Caussinus–Mestre Algorithm) programs. As would be expected in the presence of widespread natural climate variability, these breakpoints were not detected by Coll et al. 2018 using relative homogeneity methods. Given their consistency in timing across the network and their concurrent timing with a well known shift in the NAO we attribute these break points to natural climate variability.

Six stations were identified as having potential remaining discontinuities and were subsequently excluded from further analysis. These stations are Abbeyfeale (st3310), Carndolla (st2227), Westport (st1433), Creeslough (st944), Drumsna (st1529) and Kenmare (st603). The Drumsna gauge is known to have underestimated precipitation prior to 1942 when a new gauge was installed. A comparison of precipitation totals from both gauges over a 13-month period revealed that the new gauge reported 154% of the old gauge. The Drumsna series was corrected for a breakpoint detected in 1942, however, remaining discontinuities are evident in the adjusted series. The other five stations are located in the west and the northwest of the country where station density is lowest in the pre-1940 period. The Creeslough series was corrected for a breakpoint detected in 1928, however, a visual assessment of the annual series following adjustment suggests the presence of a remaining shift in the early 1960s that was not detected by RHtests. No breakpoints were detected by RHtests in the series for Abbeyfeale, Carndolla, Kenmare and Westport. However, a visual inspection of the original annual time series, the annual PRCPTOT and the R5mm time series for individual stations as well as a comparison of the series across all stations revealed large trends and potential breakpoints that were not detected by RHtests in testing the monthly series.

Following the removal of these six stations, the Pettitt test detected four significant breakpoints in the annual PRCPTOT and one significant breakpoint in the annual series of R5mm when applied to the remaining 30 station series. Figure 4.4 displays the p-values of potential break points for the Pettitt and SNHT tests. For the Pettitt test, all five breakpoints were detected in the period 1976-1979 and are consistent with the timing of other non-significant break points across the network. The SNHT detected six significant breakpoints in the annual PRCPTOT series and one significant breakpoint in the annual R5mm series. Two of these were detected in the period 1977-78. In the absence of evidence for widespread changes in measurement practices, the timing of these breakpoints suggest that a genuine climatic process was identified as inhomogeneities in the majority of series. Specifically, a shift in the North Atlantic Oscillation (NAO) to a predominantly positive phase since the mid-1970s, (Kiely, 1999).

The remaining breakpoints were detected in 2007 at Strokestown (Carrowclogher) and Inagh (Mt. Callan) and in the years 1987 and 1910 at Ballyhaise and Glenasmole, respectively. Only the break point at Glenasmole in 1910 occurs at a time when no other stations show non-significant breakpoints. It was noted during the QC process that the occurrence of heavy precipitation during December 2010 at Glenasmole resulted in large monthly and annual totals. The values were investigated and accepted based on consistency with observations recorded at neighbouring stations. Such large values can introduce spurious breakpoints in the time series. Moreover, this breakpoint is detected in both the PRCPTOT and R5mm series by the SNHT but is not detected by the Pettitt test in either series. Previous studies (e.g. Ducre-Robitaille et al., 2003; Toreti et al., 2012) have reported increases in SNHT false break detection at the beginning and the end of the series. Given the presence of this breakpoint at the beginning of the series we retain it for examining trends.

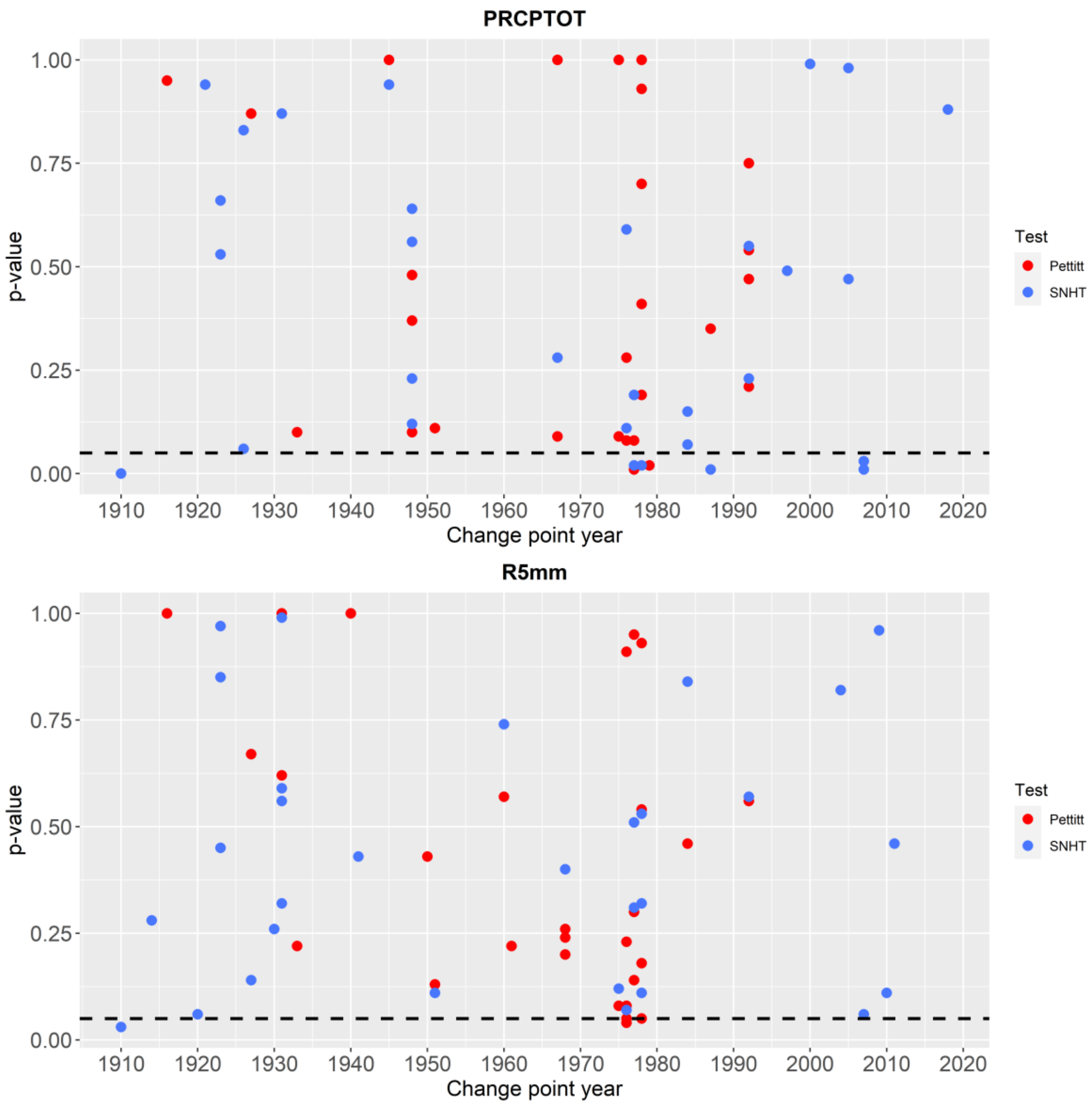


Figure 4. 4: P-values of potential breakpoints detected in the annual series of PRCPTOT (top) and R5mm (bottom) indices using the Pettitt test and the standard normal homogeneity test (SNHT). Black dashed line indicates significance at the 0.05 level.

4.3.2 Fixed period trend analysis

Trends in eleven extreme precipitation indices (see Table 4.3) were assessed for the network of 30 homogenised daily series for two fixed periods: 1910-2019 (Table 4.6) and 1940-2019 (Table 4.7). For both periods, increasing trends in extreme precipitation indices are more prevalent than decreasing, with the exception of CDD and CWD in the period 1910-2019 and CDD in the period 1940-2019.

The percentage of stations reporting increasing trends in Rx1day and Rx5day is consistent over both periods of analysis. However, there is a marked increase in the number of stations reporting statistically significant increases in Rx5day in the period 1910-2019. Of the stations reporting decreasing trends none are found to be statistically significant. There is a shift in the direction of trends at both Valentia and Glengarriff located in the southwest from a negative trend in the period 1940-2019 to a positive trend over the longer period, with a significant trend reported for Glengarriff. Large variations in trend magnitude exist across the network. Largest trends are found for Rx5day with a median increase across all stations of 6.06 percent over the period 1910-2019, increasing to 8.35 percent for 1940-2019. For the latter, the 95 percent confidence interval in trend magnitude across the network spans -6.81 to 23.77 percent. Largest trends are evident in the midlands for the period 1910-2019, extending to the southeast for the 1940-2019 period of analysis.

The percentage contribution from wet days (R95pTOT) and very wet days (R95pTOT) to total precipitation is also dominated by increasing trends. In the period 1910-2019, 83.3 percent of stations show increasing trends, 13.3 percent significant, with 76.7 percent of stations showing increasing trends in R99pTOT of which 16.7 percent are significant. For both periods, R95pTOT shows the largest median increase in trend magnitude (TSArel ~14 percent) across all stations. Again, there is large variability in trend magnitude across stations. For the period 1910-2019, for instance, R95pTOT shows a median increase of 5.07 percent, with the 95 percent confidence intervals ranging from -10.2 to 61.87 percent. Largest trends in both periods are found in the midlands and the east and southeast of the island. Significant trends reported in R95pTOT and R99pTOT at Foulkesmill (Longraigue), Portlaw (Mayfield II) and Cashel (Ballinamona) over the shorter period of analysis (1940-2019) are not observed in the longer record (1910-2019), while the longer record shows significant trends at Inagh (Mt. Callan) and Mullingar. The record for Markree station, situated in the northwest, shows the opposite. For R95pTOT, positive trends are reported in both periods of record with significant trends observed in the 1940-2019 period. For R99pTOT, negative trends are observed in both periods of record with significant trends observed in the 1940-2019 period.

CDD reveals little change, with only one significant trend reported in either period. Overall, there is a relatively even split between increasing and decreasing trends in both periods. Figures 4.5 and 4.6 indicate that decreasing trends in CDD tend to be more prevalent in the south and southwest, while increasing trends tend to be more prevalent in the midland region. In the period 1910-2019, 43.3 percent of stations report increases (10 percent significant) and 56.7 percent of stations report decreases (13.3 percent significant) in CWD. However, there is shift in the direction of trends reported in the period 1940-

2019, with 70 percent of stations reporting increasing trends in CWD, 20 percent of which are significant. The shift in direction of trend is most prevalent in the southwest of the country. Median trend in number of days per year across all stations reveals no change in either fixed period for both CDD and CWD.

Increasing trends in indices measuring precipitation threshold frequency are reported in both periods of analysis for all three indicators i.e., R5mm, R10mm, R20mm, however, the magnitude of trends are small (see Tables). The number of stations reporting increasing trends in R5mm decreases slightly in the period 1910-2019. Increasing trends in R5mm are observed at stations in the north and northwest, namely Malin Head, Belmullet and Newport which are significant over the longer period of record, whereas stations in the southwest show significant trends in the shorter period which are not observed in the period 1910-2019. The number of stations showing increasing trends in R10mm remains stable over both periods, however, a higher number of stations report significant increases over the period 1940-2019. There is an increase in the number of stations reporting increasing trends in R20mm for the period 1910-2019 compared to the period 1940-2019. The largest trends in R20mm in both periods of analysis are observed in the south and southeast of the country. No stations report statistically significant decreases in the frequency of precipitation events ≥5 mm, ≥10 mm or ≥20mm in either of the fixed periods assessed.

Both the number of stations reporting increasing trends and the median magnitude of change in intensity (SDII) is relatively consistent over the two periods of analysis, however, there is a slight increase in the number of stations reporting statistically significant increases over the period 1910-2019. These significant increasing trends in SDII are most notable in the east and southeast of the country. The magnitude of trends in intensity are also very consistent across both study periods. For the period 1910-2019 the median magnitude of change in SDII across the network is 3.84 percent, with the 95 percent confidence interval ranging from -2.02 to 9.75 percent. The number of stations reporting statistically significant increases in total annual wet-day (≥ 1 mm) precipitation (PRCPTOT) more than doubles in the period 1940-2019 compared to the period 1910-2019. There is also a shift in the direction of trend in some stations in the southeast over the shorter period of analysis. No significant decreases are reported in either period analysed. A slight decrease is noted in the median magnitude of change reported in the period 1910-2019 compared to the period 1940-2019.

Table 4. 6: Direction of change and proportion of statistically significant (5 % level) trends for 1910-2019 fixed period extreme precipitation indices. Direction and significance tested using Modified MannKendall (MKZs) and magnitude tested with the relative Theil-Sen Approach (TSArel). Magnitude of change is based on the median of the test statistics with confidence intervals given by the lower/upper bounds. Number of days is the medium numbers of days per year across all stations with confidence intervals given by the lower/upper bounds.

Indicator	Stat.	Positive(sig.) %	Negative(sig.) %	Stat.	Magnitude (CI) %	Stat.	No. of days/yr (CI)
CDD	MKZs	53.3 (3.3)	46.7 (0.0)	TSA_{rel}	0(-16.82,17.41)	TSA	0(-0.03,0.03)
CWD	MKZs	43.3 (10)	56.7 (13.3)	TSA_{rel}	0(-19.01,15.08)	TSA	0(-0.02,0.02)
PRCPTOT	MKZs	76.7 (13.3)	23.3 (0.0)	TSA_{rel}	2.79(-6.25,11.39)	TSA	NA
R95pTOT	MKZs	83.3 (13.3)	16.7 (0.0)	TSA_{rel}	13.72(-9.3,37.29)	TSA	NA
R99pTOT	MKZs	76.7 (16.7)	23.3 (0.0)	TSA_{rel}	5.07(-10.2,61.87)	TSA	NA
Rx1day	MKZs	73.3 (6.7)	26.7 (0.0)	TSA_{rel}	3.64(-11.27,19.41)	TSA	NA
Rx5day	MKZs	83.3 (26.7)	16.7 (0.0)	TSA_{rel}	6.06(-6.96,19.23)	TSA	NA
SDII	MKZs	73.3 (40)	26.7 (3.3)	TSA_{rel}	3.84(-2.02,9.75)	TSA	NA
R5mm	MKZs	66.7 (13.3)	33.3 (0.0)	TSA_{rel}	0(-8.47,11.64)	TSA	0(-0.05,0.08)
R10mm	MKZs	83.3 (10)	16.7 (0.0)	TSA_{rel}	6.22(-6.46,20.54)	TSA	0.02(-0.02,0.06)
R20mm	MKZs	83.3 (10)	16.7 (0.0)	TSA_{rel}	0(0,38.25)	TSA	0(0,0.02)

Table 4. 7: Direction of change and proportion of statistically significant (5 % level) trends for 1940-2019 fixed period extreme precipitation indices. Direction and significance tested using Modified MannKendall (MKZs) and magnitude tested with the relative Theil-Sen Approach (TSArel). Magnitude of change is based on the median of the test statistics with confidence intervals given by the lower/upper bounds. Number of days is the medium numbers of days per year across all stations with confidence intervals given by the lower/upper bounds.

Indicator	Stat.	Positive(sig.) %	Negative(sig.) %	Stat.	Magnitude (CI) %	Stat.	No. of days/yr (CI)
CDD	MKZs	40 (0.0)	60 (0.0)	TSA_{rel}	0 (-21.13,18.48)	TSA	0(-0.05,0.04)
CWD	MKZs	70 (20)	30 (0.0)	TSA_{rel}	0 (-11.05,26.25)	TSA	0(-0.02,0.04)
PRCPTOT	MKZs	76.7 (30)	23.3 (0.0)	TSA_{rel}	6.4 (-3.94,16.53)	TSA	NA
R95pTOT	MKZs	76.7 (20)	23.3 (0.0)	TSA_{rel}	13.61 (-17.05,44.69)	TSA	NA
R99pTOT	MKZs	73.3 (13.3)	26.7 (0.0)	TSA_{rel}	0 (-30.89,66.84)	TSA	NA
Rx1day	MKZs	66.7 (3.3)	33.3 (0.0)	TSA_{rel}	6.02 (-11.71,24.22)	TSA	NA
Rx5day	MKZs	83.3 (6.7)	16.7 (0.0)	TSA_{rel}	8.35 (-6.81,23.77)	TSA	NA
SDII	MKZs	73.3 (30)	26.7 (3.3)	TSA_{rel}	3.78 (-3.8,10.88)	TSA	NA
R5mm	MKZs	73.3 (13.3)	26.7 (0.0)	TSA_{rel}	5.34 (-6.64,15.82)	TSA	0.05(-0.06,0.14)
R10mm	MKZs	83.3 (20)	16.7 (0.0)	TSA_{rel}	6.09 (-9.76,22.79)	TSA	0.02(-0.03,0.09)
R20mm	MKZs	73.3 (10)	26.7 (0.0)	TSA_{rel}	0 (0,41.04)	TSA	0(0,0.03)

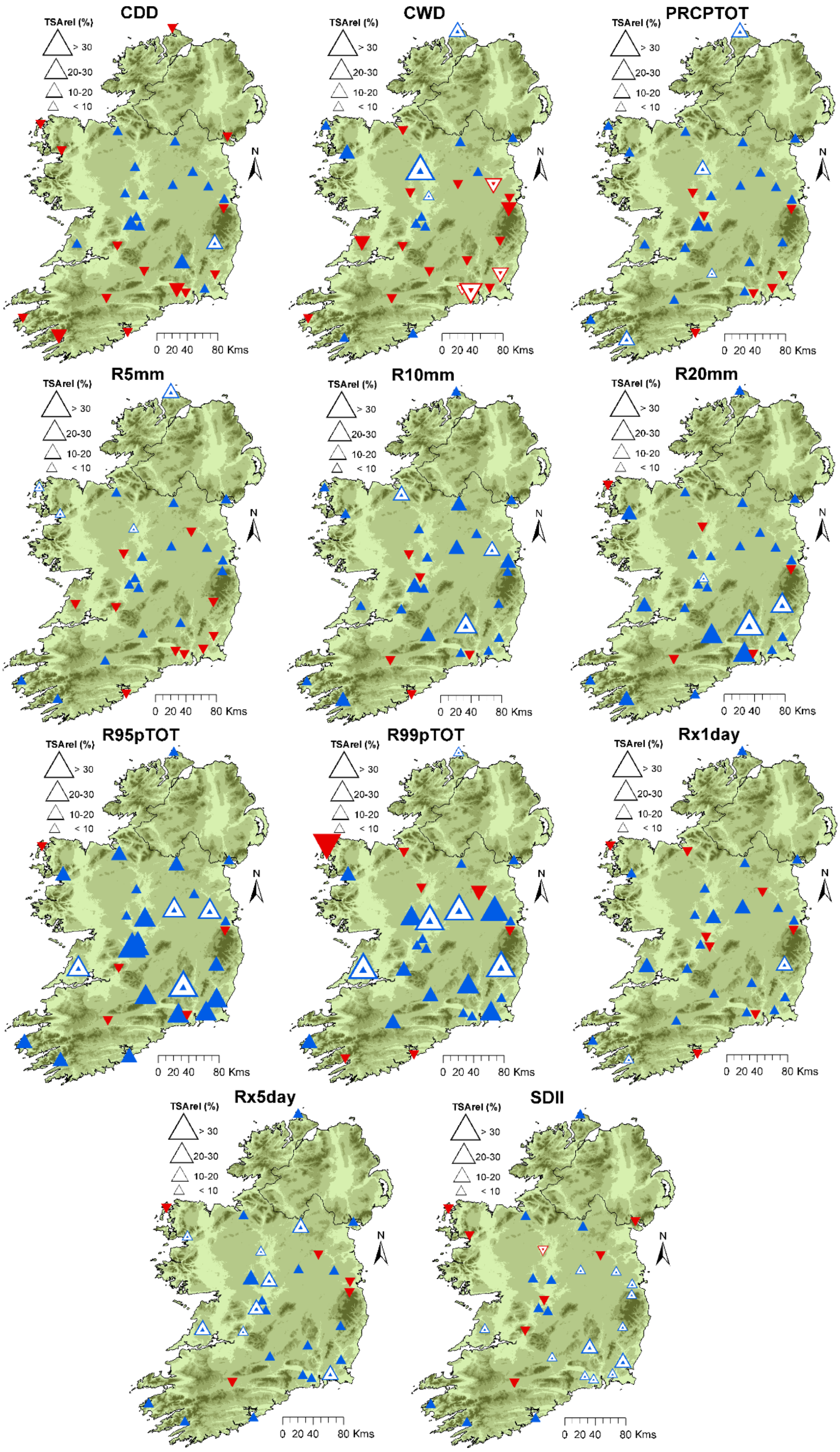


Figure 4. 5: Magnitude and direction of trends for 1910-2019 fixed period for extreme precipitation indices (see Table 4.3). Blue triangles represent increasing trends and red decreasing trends, with magnitude proportional to size. Magnitude derived using the relative Theil-Sen Approach TSArel (%). Significant trends (5 % level) shown by white triangles derived from the Modified MK (MKZs).

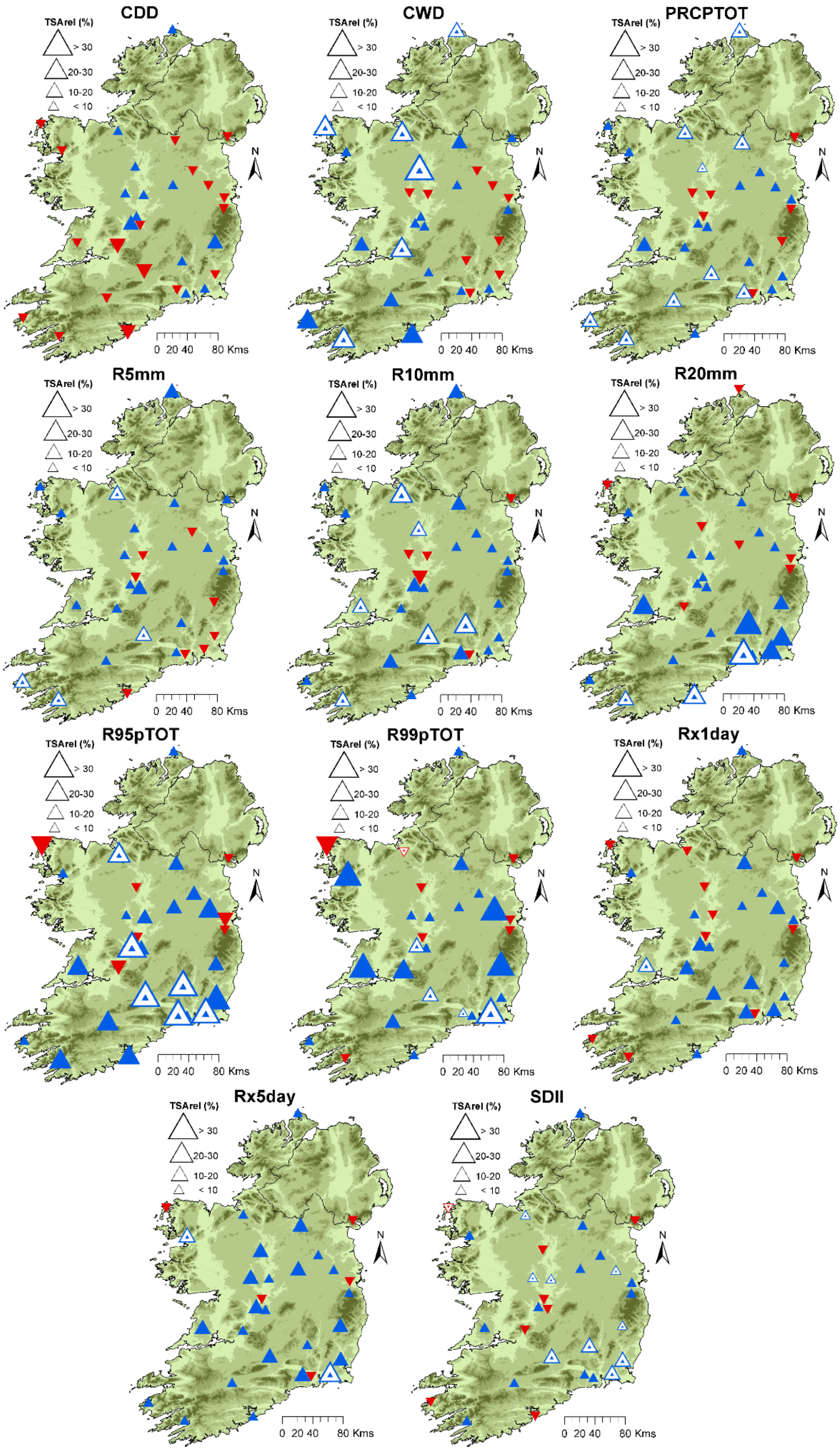


Figure 4. 6: Magnitude and direction of trends for 1940-2019 fixed period for extreme precipitation indices (see Table 4.3). Blue triangles represent increasing trends and red decreasing trends, with magnitude proportional to size. Magnitude derived using the relative Theil-Sen Approach TSArel (%). Significant trends (5 % level) shown by white triangles and derived from the Modified MK (MKZs).

4.3.3 Persistence of trends over full period of record

While the previous section highlights the predominance of increasing trends in daily precipitation indices, the dependence of results on the two different periods of record was also noted. In this section we assess the robustness of trends for different start dates over the full period of record for all 30 stations and indices. Plotting the MKZs for all stations/indices (see Figure 4.7) also allows inspection of trends in individual stations relative to the pattern of change across the network, potentially identifying stations with trends that depart from the pattern of the wider network.

CDD shows a lack of significant trends for all stations across the network for varying start years. Similarly, the majority of stations show no significant trends for varying start years for the CWD index. However, for some stations there is evidence that tests commencing before the 1900s show significant increasing trends, together with tests commencing between 1950 and 1970. Significant decreasing trends in CWD are evident for other stations for tests commencing in the early decades of the 1900s. However, these are not persistent for varying start dates. Two stations show trends in CWD that depart from the general pattern of trends across the network. For tests commencing prior to 1950, Strokestown (Carrowclogher) shows large and significant increasing trends (e.g. MKZs >4 for tests commencing in 1910). Kilskyre (Robinstown) shows large and significant decreasing trends for tests commencing after 1960. These results suggest remaining inhomogeneities at these stations affecting trends in CWD.

PRCPTOT shows a predominance of increasing trends across the full period of record for the majority of stations. Notably, there is greater variation in trend direction and magnitude for tests commencing prior to the 1900s. Persistently significant increasing trends are evident for stations/tests commencing prior to the 1970s. For tests commencing after the 1970s, the vast majority of stations show non-significant trends. The Markree series shows a notable deviation for the general pattern of trends in PRCPTOT for tests commencing before 1900, which raises suspicion about the quality of the early part of this series. Similarly, the Tycor series shows significant decreasing trends for tests commencing prior to 1905. SDII shows the largest variation in trend magnitude and direction across the network of stations, with different stations showing persistent significant increasing and decreasing trends for tests commencing throughout the record. R95pTOT and R99pTOT are dominated by increasing trends for different start years, with persistent significant increasing trends evident for some stations for tests commencing since the 1930s. Only a small number of stations show weak, non-significant decreasing trends in these indices. Both R5mm and R10mm show a predominance of increasing trends for most stations for different test periods. For tests commencing post 1970 there is a tendency for non-significant trends. For stations with records extending prior to 1900 there is greater variation across stations in trend tests commencing in the late 1800s. For R10mm Markree and Newport station series show significant increasing and decreasing trends, respectively, for tests commencing prior to 1890. While the number of stations available for comparison decreases at this time, the trends appear larger than those at other stations. Trends in Rx1day and Rx5day precipitation show coherence across the network for varying start dates. On the whole, trends in Rx5day tend to be persistently positive for the majority of stations, irrespective of start date.

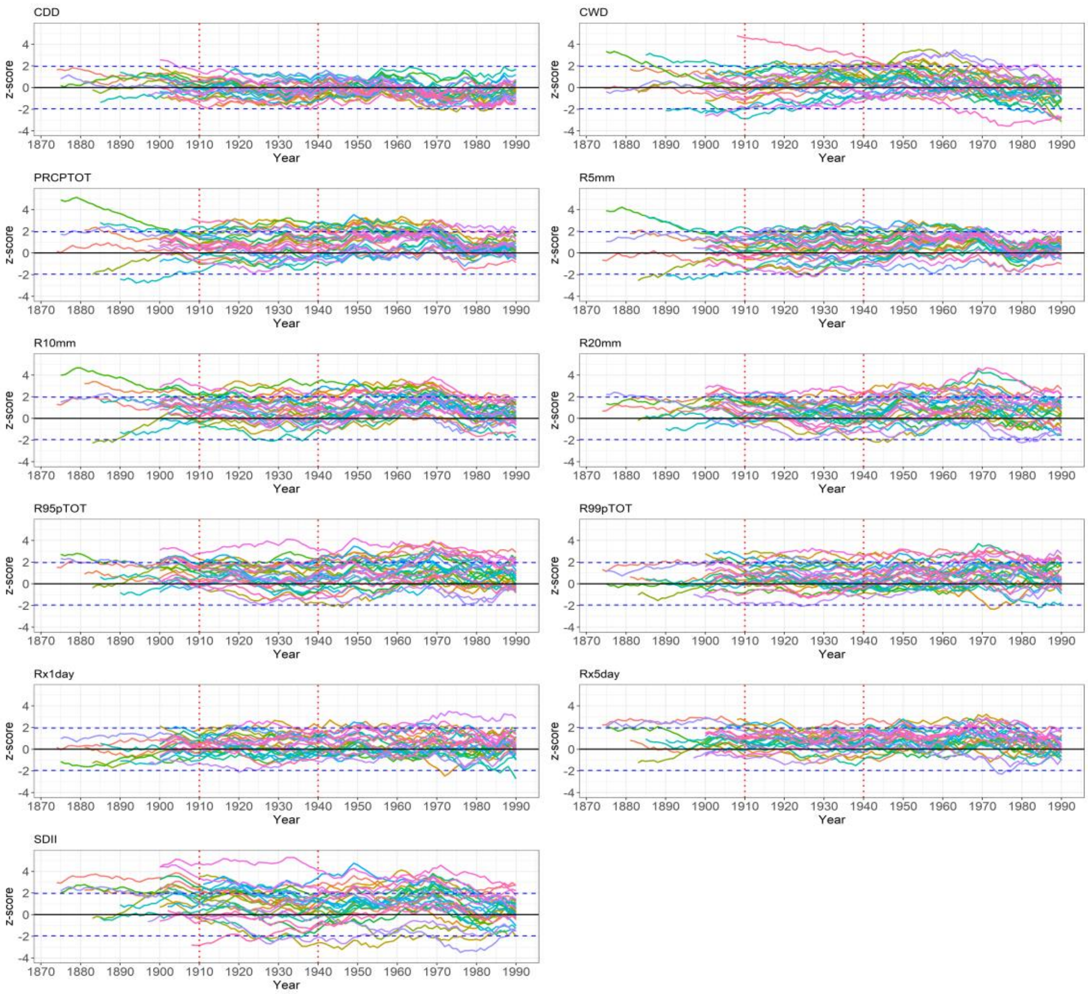


Figure 4. 7: Persistence plots for full available time-series of extreme precipitation indices. Coloured lines represent MKZs statistics for varying start years for individual stations. Dashed blue lines are the threshold for statistically significant trends at the 5 % level with MKZs above (below) these indicating significant increasing (decreasing) trends since the corresponding start date. The vertical red lines mark the start year of the 1910-2019 and 1940-2019 fixed periods.

4.4 Discussion

This study derived quality controlled long-term daily precipitation series for 36 stations across Ireland by leveraging data rescue efforts to extend existing digital records available after 1940. The average series length is 118 years, with some stations dating back to the 1870s. Following quality control and homogenisation, trends in precipitation extremes from annual ETCCDI indices were examined for 30 stations. Results indicate the predominance of increasing trends across the network. However, the direction, magnitude and significance of trends is dependent on the period of analysis, emphasising the importance of the long records developed here. The analysis of trend persistence for varying start years indicates the influential role of low frequency variability on trend results. Of particular influence has been the shift to more prevalent positive NAO conditions around the late 1970s and a greater predominance of westerly airflow, the impact of which is most evident for the CWD, PRCPTOT, R5mm and R10mm indices. The impact of this change on regional circulation has been widely recognised (e.g Hurrel, 1995; Kiely, 1999; Murphy et al., 2013).

Most previous assessments of precipitation in Ireland have been limited to the post-1940 record and have not undertaken assessment of ETCCDI indices (e.g Walsh and Dwyer, 2012; Sheridan, 2001) making direct comparison difficult. Assessments undertaken for longer records have focused on seasonal and annual totals (e.g. Butler et al., 1998; McElwain and Sweeney, 2007; Noone et al., 2015). These studies have drawn attention to the increase in annual totals, particularly for stations on the western seaboard. Our results show that of the indices examined here, PRCPTOT is most sensitive to the influence of low frequency variability and period assessed. We find the greatest magnitude of trends for indices measuring the contribution of heavy precipitation events to annual totals (i.e. R95pTOT and R99pTOT), particularly in the south and south east of the island. The largest number of significant increasing trends is reported for intensity (SDII), also in the east and south of the island.

A crucial step in assessing trends is the quality assurance of available data to ensure derived results are not an artefact of measurement practice or data processing. It is widely accepted that the reliability and accuracy of homogenisation procedures generally decrease with the increasing temporal resolution of homogenisation (Venema et al., 2012). The task is further complicated for early records as station density typically decreases with time (Hollis et al., 2019). Therefore, homogenisation of daily data is considered to be a much more challenging problem than homogenisation at monthly or annual scales (Coll et al., 2018). Relatively few homogenisation algorithms include methods for the homogenisation of daily climate series and those that do, e.g. MASH (Szentimrey, 1999, 2014) and ACMANT (Domonkos, 2015), tackle the homogenisation of daily precipitation data based on a procedure in accordance with the multiplicative (or cumulative) model that is assumed for monthly data. Moreover, the theoretical minimum number of well-correlated reference stations needed for the application of relative homogenisation methods as recommended by the World Meteorological Organisation is three, with at least four reference stations required to obtain good results in more complex situations (WMO, 2020). Obviously, this is challenging for records that extend into the 1800s, given decreasing network density.

Our study represents a relatively rare attempt to detect and correct inhomogeneities in daily precipitation data for long-term records spanning more than a century. We employed the RHtests homogenisation software as it provides one of the few methodologies for the detection and adjustments of breaks in daily precipitation data without the use of a reference series. Reeves et al. (2007) found that two-phase regression methods, as implemented in the RHtests software, had a comparable level of performance to methods such as the standard normal homogeneity test (SNHT), with the performance of the algorithm dependant on the parameters defined by the user. In total, 22 breaks were identified by RHtests software in 16 of the 36 stations, with approximately 80% of detected breakpoints supported by metadata relating to instrument and site changes and were adjusted using quantile mapping. However, subsequent assessment of residual breakpoints in PRCPTOT and R5mm using the Pettitt and SHNT tests resulted in six stations being excluded from the analysis of trends. For four of these stations (Abbeyfeale, Carndolla, Kenmare and Westport) no significant breakpoints were detected by RHtests, despite the fact that potential breakpoints documented in metadata were manually entered and tested.

For Drumnsa station, two significant residual breakpoints were detected; the first in 1917 had no associated metadata but was also detected in a previous study by Noone et al. (2015) using HOMER (Homogenisation Software in R) software. The second breakpoint detected in 1942 corresponds to the installation of a new gauge at the time. While this break was detected by RHtests and corrected, discontinuities evident in the adjusted series suggest that the adjustment applied was too modest. The final station to be excluded was

Creeslough. This series was corrected for a breakpoint detected in 1928. However, visual assessment of the annual series following adjustment suggests the presence of a remaining shift in the early 1960s that was not detected by RHtests. Given these uncertainties surrounding break detection and correction we suggest that other studies examine output from RHtests algorithms for residual breaks, as was also done by Santo et al. (2014) in their assessment of Portuguese seasonal extremes.

While the nature of the network analysed required the use of absolute homogeneity tests, our analysis built in assessments of residual breakpoints and trend tests that leverage information from across the network, thereby building confidence in our results. For instance, the plotting of p-values for residual breaks across series (Figure 4.4) allowed comparison of the timing of potential break points across the network, helping to isolate breaks unlikely to be associated with climate variability. Furthermore, the assessment of trend persistence following Noone et al. (2015) and Murphy et al. (2016) aided the visual detection of stations for which the behaviour of trends through the record deviated significantly from other stations. In this regard, cautionary flags were raised for trend in some stations/indices where remaining inhomogeneities may exist, particularly in early records. These include Strokestown, Markree and Newport, especially for trends in CWD, PRCPTOT and R5mm.

There is much scope for further work to build on the dataset and results presented. Future work will assess changes in seasonal and monthly extremes across the derived network. Moreover, Noone et al. (2015) presented a long-term monthly precipitation network for the island of Ireland. Future updates of that network should include monthly totals for the additional stations developed here. While this work has improved understanding of the changing nature of extreme precipitation on the Island of Ireland, it would not have been possible without concerted efforts at data rescue. While significant progress has been made in making pre-1940 daily precipitation available for scientific scrutiny through the PhD research of the lead author (Ryan et al., 2018; Ryan et al., 2020), much data remains to be rescued and digitised for Ireland. Ongoing data rescue initiatives at Met Éireann could be directed at filling the spatial gaps in this network and improving the density of data from early stations to both assist with understanding the spatial distribution of changes and to open the possibilities for deploying relative homogenisation methods. There is also the potential to develop gridded monthly products back to at least the start of the 20th Century (Hollis et al., 2019). Lastly, given the uncertainties associated with homogenisation of daily data, future work should also evaluate the findings presented here using other homogenisation software. Insights from initiatives such as the European Research Area for Climate Services (ERA4CS) INDECIS project that are assessing approaches to homogenisation of daily data will be critical in this regard.

Finally, with increasing concerns about the impacts of climate change, understanding changes in the characteristics of precipitation is increasingly important given that climate extremes often exhibit different behaviours than changes in average conditions. The UN Intergovernmental Panel on Climate Change (IPCC) Fifth Annual Assessment Report (AR5) states that over most of the mid-latitude land masses, extreme precipitation events are very likely to be more intense and more frequent in a warmer world (IPCC, 2014). Trenberth et al. (2003) propose a ~7% increase in water vapour content for every 1°C rise in temperature. Our findings are largely consistent with expected changes in daily precipitation extremes in a warming world. The overall tendency of increasing trends observed here suggest that heavy or extreme precipitation events are contributing significantly more to annual precipitation totals in Ireland, while precipitation is becoming more intense, particularly in the east and southeast of the country. However, potential changes in North Atlantic atmospheric circulation patterns will also have a direct impact on future precipitation, with changes in the frequency and intensity of Atlantic depressions and their associated weather fronts and convective activity important factors for extreme precipitation (Matthews et al., 2015). Therefore, it is important to further investigate the driving mechanisms behind the trends reported.

4.5 Conclusion

This study examines trends in annual precipitation for Ireland using the ETCCDI indices for a newly developed long-term daily precipitation network. Stations within the network have an average record length of 118 years, with some stations dating back to the 1870s. Our results show the predominance of increasing trends across the island, the growing contribution of heavy precipitation events to annual totals and increases in precipitation intensity, especially in the east and southeast of the island. While results are consistent with expectations in a warming world, the influence of low frequency variability is substantial. Understanding variability and change in long-term climate is dependent on data rescue activities and approaches to homogenisation of daily data. While approaches to undertake the latter are still in their infancy, this work shows the benefits of combining existing techniques to increase confidence in the derived series.

5. Discussion, conclusions, and future work

5.1 Introduction

The principal aim of this research was to develop innovative data rescue activities to transcribe and digitise archived paper records to derive a long-term daily precipitation network for Ireland. The resultant data were then quality assured and homogenised before being assessed for evidence of variability and change in precipitation extremes.

The aims were addressed by the following four thesis objectives set out in Chapter 1:

- **Thesis Objective 1**: To integrate data rescue and quality assurance with teaching and learning at Maynooth University through development of approaches for citizen science in the classroom (Chapter 2).
- **Thesis Objective 2**: To develop a long-term daily precipitation network for Ireland that takes advantage of the rich weather observing heritage on the island and to ensure the raw data and metadata is made available in a fully traceable and open way to underpin future research endeavours (Chapter 3).
- **Thesis Objective 3**: To quality control and homogenise data to facilitate the analysis of variability and change (Chapter 4).
- **Thesis Objective 4**: To conduct a statistical assessment of the derived quality assured long-term series to assess changes in the characteristics of extreme events (Chapter 4).

This final chapter summarises the main research findings, discusses limitations, and identifies areas for further research.

5.2 Summary of main research findings

5.2.1 Thesis objective 1: Integrating data rescue into the classroom (Chapter 2)

One of the challenges of this thesis, and of historical climatology in general, is the digitisation and transcription of historical climate records. Chapter 2 presents a novel approach to data rescue which was developed as an alternative, research-led approach to teaching and learning that would comply with the prerequisites of pedagogy within the university curriculum while advancing scientific understanding. The aims of the project were first, to explore the potential for integrating data rescue into the classroom, second, to examine the potential for students to produce accurate and reliable observational data, and third, to motivate students by engaging them in a practical exercise whereby their contribution has considerable value to research.

The performance of the students was evaluated using a simple error assessment which showed that the performance of the students was comparable to experienced scientists. Moreover, post-processing of the transcribed data facilitated the creation of a corrected data series.

On comparison of the double keyed data, 62% of student transcribed sheets contained no errors; thus did not require the creation of a 'corrected' data file. Further, 96% of student transcribed sheets contained less than 20 errors, minimising the amount of time spent creating the master files. These results convey the significant contribution that the integration of citizen science into the classroom offers to the study of climatology and other fields of research. Furthermore, the students demonstrated that their performance was comparable to that of experienced professionals, with cumulative errors across all sheets transcribed by the students and Met Éireann revealing a percentage error of ~1% and 1.75%, respectively. Upon investigation of the source of errors it was noted that different approaches to the transcription process i.e., row-based or column-based approach had an impact on the number of errors produced within individual sheets. Specifically, the row-based approach employed by Met Éireann resulted in the majority of incorrect sheets falling into the lowest error category (i.e. ≤5 errors), while the column-based approach adopted by the students produced a greater number of sheets containing errors in the intermediate categories resulting from the incorrect placement of a data value which then propagated down throughout the monthly column. While the values were correct, differences were flagged due to the data values being input to the incorrect date/cell.

Student engagement was evident throughout the course of the project and through the results of the assignment evaluation. In particular, this was highlighted by the detection of errors in monthly totals as calculated by the original observer. This was also a significant learning opportunity to highlight the imperfect nature of data to the students and to initiate discussions regarding quality control measures and the importance of metadata.

Supports implemented to assist students were fundamental to the success of the project. The management of these resources demanded a significant investment of time, in particular the online discussion forum which fielded over 500 questions from students during the course of the first round of the project. Nevertheless, effective management of the project and the necessary supports facilitated the development of the ‘corrected’ data series by notably reducing the amount of time required to carry out post-processing of the data.

Following the success of the first initial experiment two further iterations were carried out. Issues identified in the first round of the project were addressed with the development of additional supports to facilitate the transcription of the data and inevitably reduce the margin of error further. These included i) the development of a comprehensive metadata file whereby students could access information relating to, for example, station name and location, length of record and observer details. This reduced the potential for inconsistencies in the collection of metadata; thus increasing confidence in the product and ii) delineating the margins of the column and rows in the template which led to a reduction in errors associated with the incorrect placement of data values, especially differences arising from the different approaches adopted by the students to carry out the transcription process, noted as the greatest source of error in the first round of the project. Finally, students involved in the data rescue project were given the opportunity to take a research methods module in which they carried out statistical analysis of trends in daily rainfall on the raw data with a view to developing students’ analytical skills.

This research has resulted in i) excellent learning outcomes for students relating to the processes and methods involved in historical climatology; ii) a highly successful means of transcribing paper based data; iii) large volumes of rainfall records that were previously not available to scientific scrutiny being made accessible. To my knowledge, this was the first attempt to demonstrate the role that students can play in improving our understanding of past climate through data rescue activities.

The innovative nature of the data rescue in the classroom initiative has been recognised nationally and internationally. The project has won notable commendation from the World Meteorological Organisation (WMO) who have highlighted the work as an example of best practice in data rescue (see following link for details: https://library.wmo.int/index.php?lvl=notice_display&id=19894#.YDPYzej7T-g). The approach has also inspired interest at other Universities, with input and advice given to efforts at data rescue in the classroom at other universities in Ireland, the U.K. and Sweden. The approach has also been awarded Maynooth University’s Research Impact Award

2018 and the Department of Housing, Planning, Community and Local Government's Réalta Award 2018 for innovative collaboration between Met Éireann and Maynooth to the benefit of public services in Ireland.

5.2.2 Thesis objective 2: Production of the raw datasets and associated metadata (Chapter 3)

Despite recent efforts to increase access to data resources, limits to the accessibility of such records continues to impede progress in the development of robust climate analyses. Over the course of the three data rescue projects undertaken by Maynooth University students, 3616 station years of daily precipitation data (~1.32 million daily values) and associated metadata were transcribed for stations across Ireland. Post-processing of the transcribed data involved a comparison of the double-keyed data and the application of quality assurance tests to check for transcription errors. Furthermore, a significant amount of metadata was extracted from the original records and transcribed into individual station files. Station details including location, record length, unit of measurement and percentage missing were also compiled for presentation in the publication.

In addition to the transcribed rainfall observations, meteorological observations from eight long-term stations have been scanned and integrated into Met Éireann's digital database using the Metis EDS professional digital scanner that was purchased by Maynooth University during the course of this work. Individual pages were scanned as high quality tif files and categorised using the standard naming convention adopted by Met Éireann. The scanned images were then uploaded to Met Éireann's database where they will be made available for research purposes.

The main objective of the work presented in Chapter 3 was to make the data readily updatable and widely available to national and international researchers to progress science and benefit society at large. The data presented in the publication are the raw data and associated metadata. It is envisaged that by presenting the data in its original state it can be easily integrated into current international data rescue initiatives, e.g. Copernicus Climate Change Service Global Land and Marine Observations Database, and that future research will have recourse to the raw data. Moreover, details of the data rescue process and dataset development provided in the publications that are presented in this thesis provide a fully traceable dataset.

The data are freely available from the edepositIreland data centre (**http://hdl.handle.net/2262/91347**). The dataset comprises daily rainfall data for 114 stations at various locations throughout Ireland for varying time periods. These data are the only daily rainfall observations available for Ireland prior to 1941. Individual station folders contain two files: a data file in ASCII format and a corresponding metadata text file. The derived data will also be hosted by Met Éireann and routinely updated and widely disseminated.

The data have already been utilised by Met Éireann's Flood Forecasting Division as well as a varied range of ongoing research projects. These include:

- Research undertaken by a PhD student in the University of East Anglia on the impact of weather variability on crop production.
- Research undertaken at the University of Reading which utilised the imaged meteorological records to reconstruct the 'Ulysses Storm' which hit the British-Irish Isles on the 27th February 1903 (https://weatherrescue.wordpress.com/2017/10/12/february-1903-the-ulysses-storm/)
- Research for an ongoing Irish based art project to ascertain the daily weather patterns for some specific locations in Ireland for the period 14th August to 29th September 1937. The data is being used to facilitate a more detailed mapping of the travel experience of an influential French artist, Antonin Artaud whose visit to Ireland proved to be highly impactful on the remainder of his life and his later artistic work.
- The data is also feeding into ongoing PhD research at ICARUS with Paul O'Connor using the data to reconstruct river flows for Irish catchments, while Prof. Murphy is using the data to understand historical drought and the impacts of arterial drainage in the Boyne catchment. Moreover, work on Seasonal Hydrological Forecasting by Prof. Murphy, funded by Science Foundation Ireland, is using the long records developed to identify analogue conditions for ensemble forecasting techniques. The increased sample provided by data rescue activities increases potential to provide usable seasonal forecasts for water managers.

The varied nature of the above research highlights the value of the rescued data across a range of disciplines. In addition to the scientific benefit, unlocking historical data contributes culturally significant material to understanding Irish environmental history and acknowledges and brings back to life the dedicated efforts of many hundreds of weather observers.

5.2.3 Thesis objective 3: Quality control and homogenisation of the data (Chapter 4)

The lack of long-term quality controlled daily rainfall data for Ireland was highlighted in Chapter 1. Utilising the rescued data described in previous sections, Chapter 4 describes the application of quality control and homogenisation methods and the development of long-term daily rainfall series for 36 stations across Ireland. Post 1940 daily rainfall observations extracted from Met Éireann's database had previously undergone quality control tests including spatial techniques, comparison with daily radar accumulations, and an automated system for the re-distribution of cumulative daily values. Due to the sparse density of available stations in the early record it was necessary to develop additional tests that could be used to quality control the historical data. Following guidelines provided by the World Meteorological Organisation (WMO) a set of statistical tests that were applicable to this type of data were developed. The final set of QC checks include (i) basic integrity checks, (ii) tolerance checks (iii) temporal consistency checks and (iv) spatial consistency checks.

Following the application of QC methods, long-term series were created using the method developed at Met Éireann for deriving gridded datasets, with the approach also used to infill and extend station series using rescued data. First a regression-kriging interpolation model was used to generate monthly values for all stations in the database. For stations that had gaps in data, the monthly total was estimated using nearby station data. Next, for missing daily values, inverse distance weighted interpolation was used to disaggregate monthly interpolated values to provide a complete series of daily values. Finally, stations for analysis were selected. Given the lower station density in the pre-1940 period, only sites with neighbouring station data available were selected when developing long-term series. In total 36 stations were identified based on record length, continuity, completeness and proximity of pre-1940 station locations.

Following a review of the available methods, RHtests was deemed the most suitable method for homogenisation. Furthermore, it is the method recommended by ETCCDI for use in the Rclimdex software http://etccdi.pacificclimate.org/software.shtml. Breakpoints were detected in the monthly log-transformed series and adjustments were applied to the non-zero daily precipitation series by means of a Quantile Mapping (QM) approach using the underlying distribution of the data to determine the magnitude of the adjustment.

In total, 22 breaks were identified by RHtests software in 16 of the 36 stations, with approximately 80% of detected breakpoints supported by metadata relating to instrument and site changes and were adjusted using quantile mapping. However, a subsequent assessment of the annual series of PRCPTOT and R5mm using the Pettitt and SHNT tests identified a number of stations with potential breakpoints remaining in the series which resulted in six stations being excluded from further analysis. The detection of residual breaks following application of RHtests software highlights the challenges of homogenising daily data, with a key recommendation from this work being the need to integrate multiple homogenisation tests.

5.2.4 Thesis objective 4: Analysis of trends in long-term extreme precipitation indices in Ireland (Chapter 4)

From a societal perspective, changes in the frequency or magnitude of extreme events tend to have the greatest impact. Events of this nature typically arise from short duration events; analysis of such events requires an evaluation of daily indices, for instance, changes in the number of wet days or number of days during which rainfall amounts exceed a certain threshold. In Ireland, analysis of daily precipitation has largely been limited to the post 1940 period, for which digitised observational and gridded precipitation products are available. Moreover, no previous analysis has assessed long-term changes in precipitation extremes for Ireland using the suite of indices defined by the CCl/ CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI). Chapter 4 addresses this research gap by presenting a statistical assessment of the derived quality assured long-term series to assess changes in the characteristics of extreme events at 30 stations across Ireland. Quality assurance of the rescued data was carried out before selected long-term stations were tested for homogeneity using RHtests software. Eleven extreme precipitation indicators were calculated on an annual basis for two fixed periods, 1910-2019 and 1940-2019, and analysed to determine spatial and temporal trends in the frequency, intensity and magnitude of observed precipitation.

Evidence for monotonic trends was assessed for two fixed periods using the Modified Mann-Kendall (MK) test (Hamed and Rao, 1998), a non-parametric rank-based method. Trend magnitudes were also estimated from the Theil-Sen approach (TSA) in order to corroborate and map trends detected by the MK method. To facilitate a relative comparison, the trends for each time-series were expressed as a percentage over the period of record of n years relative to the mean for the period (TSArel (%)). Making full use of available record length, dependency of trends on period of record was examined by systematically reducing the start year of analysis from the whole record to a minimum of 30 years. This approach also allows further evaluation of series homogeneity by facilitating comparison of the consistency of trends across the network.

Coincident with global trends, results indicate the predominance of increasing trends in extreme rainfall indicators across Ireland. However, the direction, magnitude and significance of trends is dependent on the period of analysis, emphasising the importance of the long records developed here. The analysis of trend persistence for varying start years indicates the influential role of low frequency variability on trend results. Of particular influence has been the widely noted shift to more prevalent positive NAO conditions around the late 1970s. Of the indices examined here, PRCPTOT is most sensitive to the influence of low frequency variability and period assessed, while the largest increasing trends were found for indices measuring the contribution of heavy precipitation events to annual totals (i.e. R95pTOT and R99pTOT), particularly in the south and south east of the island. These findings call into question the dominant narrative of climate change in Ireland of the west getting wetter. Findings presented here suggest that increasing trends in annual totals in the west are an artefact of the period of record analysed. Rather the most significant changes, that show evidence of persistence throughout the long records analysed, are changes in extreme precipitation in the south and southeast.

5.3 Limitations and priorities for future work

5.3.1 Data Rescue

Long-term quality assured records are critical, because without them, we cannot understand the scale and pace of climate change (WMO, 2020). Long records can be used to monitor climate variability and change, support effective climate risk management and improve future climate projections (Thorne et al., 2018). Yet every year millions of meteorological observations are being lost due to deterioration of raw data sheets (Brönniman et al., 2018). While the data presented in this thesis mark a significant and innovative effort to progress data rescue efforts for Ireland, a considerable amount of data has not yet been digitized and exist only in hard copy format. The priority for Met Éireann over the coming years is to catalogue and image all historical records currently held in the archives for integration into the climate repository, an effort significantly enhanced by the development of procedures and resources (including a digital scanner) as part of this work. However, the most significant barrier to progressing data rescue activities is the lack of human resources to carry out the work (Wilkinson et al., 2019). The novel approach to data rescue developed as part of this thesis as a collaboration between Maynooth University and Met Éireann has significantly advanced data rescue efforts for Ireland. In addition, the performance of students in the transcription of meteorological data from images of historical records is comparable to other large-scale data rescue projects (e.g. Hawkins et al., 2019; Mateus et al., 2020). Given the benefits associated with integrating data rescue activities into the classroom for both students and researchers, as well as the accuracy students have shown in such activities, it is likely that such initiatives will play an increasingly important role in future data rescue activities (Brönniman et al., 2018).

Maynooth University is keen to facilitate future iterations of the data rescue project as part of the undergraduate Geography module GY313 Climate Change. The amount of data and associated metadata that could potentially be transcribed during one iteration of the project is approximately 800 station years of daily observations. This would require a commitment of one full-time member of Met Éireann's team for a period of ~16-20 weeks per iteration of the project. For comparison, utilising in-house staff, it would take a commitment of two full-time staff members ~1 year to match this output. For quality control purposes, the data is double-keyed, hence the requirement for two members of staff.

There is also potential to extend the data rescue project through the development of a web-based platform. Similar online projects have been successfully developed for weather records in the UK (Rainfall Rescue, Hawkins et al., 2019) and Canada (DRAW, Park et al., 2018). Such an application would facilitate the extension of the data rescue project to other teaching programmes (e.g. second-level education and further education initiatives) and significantly advance contributions from citizen scientists. Slonosky et al. (2019) and Wilkinson et al. (2019) provide recommendations on how to engage volunteers as citizen scientists and highlight the importance of good resources to avoid data quality problems, particularly in the case of non-expert participants. The lessons learnt during the course of this thesis, for instance, the development of project resources e.g. transcription templates, video tutorials and reference metadata files will be fundamental to the development of such a project for Ireland.

Collaboration with other organisations both nationally and internationally is also important to progress data rescue activities in Ireland. Involvement in such partnerships has many benefits such as: (1) helping to identify raw sources of data in archives (especially those at greatest risk of being lost), (2) sourcing already digitised datasets which are not in the databases, (3) speeding up the process of scanning and digitising historical data, (4) ensuring that all quality-controlled digitised data makes its way into shared databases, and (5) demonstrating the value of historical data for answering important scientific questions. Following the success of the global ACRE initiative (http://met-acre.net/), led by Prof. Rob Allan at the UK Met Office, the 1st ACRE British & Irish Isles workshop was held on 11th-12th February 2020 at the University of Reading. The aim of the workshop was to discuss prospects for recovering historical weather data which is currently undigitised, with a focus on the UK and Ireland. The 2nd meeting is being co-ordinated by the University of Birmingham in 2021. As part of my new role in Met Éireann I will be attending these meetings to promote collaborations and knowledge-sharing between interested parties.

Finally, whilst Met Éireann has a large repository of paper-based records it is important to continue searching for other historical weather records. Some of these records are known but missing whilst others are still undiscovered. In recent years, various weather records have been recovered following public appeals. For example, the complete climate records from Headfort House, Kells, Co. Meath which date back to the late 1800s were recovered in 2020 as well as the Kilkenny Castle series for the period 1884-1933. Daily rainfall observations recorded in the diaries were subsequently transcribed as part of this research.

5.3.2 Developing long-term gridded datasets

Met Éireann have previously produced gridded datasets for monthly and daily rainfall for the period 1941-2016 (Walsh, 2016). With the availability of the rescued data there is now potential to extend these gridded datasets back to the late 19th Century. This process is more difficult for daily grids given the, at times, limited spatial coverage of stations in particular regions (e.g. west of Ireland) in the pre-1940 dataset. However, the production of monthly gridded datasets extending to at least the early 20th Century should be an important and achievable next step from this research.

The UK Rainfall Rescue project run by Prof. Ed Hawkins at the University of Reading has recently digitised all the Met Office ten-year rainfall sheets and is currently working to extend the Met Office monthly gridded dataset back to 1836. Many Irish stations (ROI & NI) were included in this project and are now available. Moreover, given that there is so much information to share across lines on a map it would be important to tie together the UK and Ireland data in the same gridded dataset. Indeed, sharing of expertise, resources and datasets between our islands is key to progressing understanding of historical climatology. In my role in Met Éireann I am excited to continue progression of this important work.

The gridding programmes developed by Met Éireann were used in this thesis to facilitate the development of complete daily series required in the calculation of extreme indices. For interpolation of precipitation, geostatistical methods such as the kriging spatial interpolation method used in this thesis are usually preferrable (Szentimrey et al., 2007; Tao et al., 2009). This method performed well in a previous assessment by Walsh (2016) using data from the Irish precipitation network. However, due to the lower station density in the pre-1940 period, the selection of stations for analysis of trends was limited to sites for which continuous station series exist for the period of record or where neighbouring station data were available in the pre-1940 period. Gridded datasets based only on instrumental observation networks depend on the spatiotemporal distribution, quality, and length of records (Sun et al., 2014). Therefore, the challenge in trying to represent the highly variable nature of rainfall in the development of gridded datasets is to overcome the lack of density of observations in both time and space (Gofa et al., 2019). This obstacle, particularly acute for daily data, could be significantly alleviated through additional data rescue efforts.

5.3.3 Homogenisation

The effectiveness of homogenisation methods depends on the type, temporal resolution and spatial variability of the climatic variable (Ribeiro et al., 2015). Uncertainties in homogenising and applying corrections to daily precipitation data are well documented (e.g. Auer et al., 2005; Venema et al., 2012; Coll et al., 2019). The high spatial variability and non-continuity of daily precipitation mean that true climate fluctuations may be interpreted as change points and removed from the time series as inhomogeneities (Ribeiro et al., 2015). Moreover, the magnitude of inhomogeneities may differ with varying weather situations (Nemec et al., 2013).

Relative homogenisation methods are considered more robust than absolute methods when suitable references series are available because the climatic variation that is common for the study region is not detected as an inhomogeneity in the differences between the candidate and nearby stations (Domonkos, 2013). However, many of the available methods aim to detect inhomogeneities in the time series at the monthly or annual scales but do not provide techniques for the adjustment of breaks in the daily time series. Those that do use a procedure in accordance with the multiplicative (or cumulative) model that is assumed for monthly data and employ techniques to pattern scale the monthly adjustment to daily data e.g. MASH (Szentimrey, 1999, 2014) and ACMANT (Domonkos, 2015). Such methods are based on the assumption of normality in the monthly or annual series for which a correction in the mean or central tendency is considered sufficient, but do not apply corrections to higher order moments which would be necessary in the correction of daily data (Della-Marta and Wanner, 2006).

A review of current literature during the course of this research revealed that limited examples of applications of homogenisation methods at a daily time scale exist. In particular, for previous research utilising breakpoint detection methods for daily precipitation data, the majority of studies (e.g. Alexander et al., 2006; Caesar et al., 2011; Shrestha et al., 2016; Sun et al., 2016) do not attempt to adjust the data series but rather exclude data series that are deemed inhomogeneous or only consider the longest homogeneous period in the analysis. As such, this thesis provides a rare example of the application of homogenisation software to daily precipitation data. While the nature of the network analysed required the use of absolute homogeneity tests, further assessment of residual breakpoints and trend tests facilitated a relative assessment of breakpoints by comparing trends across the network,. For instance, the plotting of p-values for residual breaks across series allowed comparison of the timing of potential break points, helping to isolate breaks unlikely to be associated with natural climate variability. Furthermore, the assessment of trend persistence following Noone et al. (2015) and Murphy et al. (2016) aided the visual detection of stations for which the behaviour of trends through the record deviated significantly from other stations. While cautionary flags were raised for some stations/indices the use of multiple tests builds confidence in the results.

As homogenisation methods for the detection and correction of breakpoints in daily series develop there will be opportunities to assess the strengths and weaknesses of different approaches. Such assessments require access to the raw data and associated metadata as presented in Chapter 3. Moreover, in the absence of advancements in the development of methods for the statistical homogenisation of daily climate data, improving data availability through data rescue initiatives becomes increasingly important. In such cases, improving network density and access to station metadata can help to improve the performance of existing homogenisation methods.

5.3.4 Future dataset use

There is much potential for further work to build on the dataset and results presented here. While beyond the scope of this research, future work will assess trends in seasonal and monthly extremes for Ireland. This will address a significant research gap in understanding the temporal distribution of extremes across the network. The derived data will also be used to expand the long-term monthly precipitation network for the island of Ireland in future updates (IoI, Noone et al., 2015) and to develop gridded monthly precipitation datasets.

Potential changes in North Atlantic atmospheric circulation patterns will have a direct impact on future precipitation, with changes in the frequency and intensity of Atlantic depressions and their associated weather fronts and convective activity important factors for extreme precipitation (Matthews et al., 2015). Results show evidence of an increase in annual totals associated with a widely noted shift to more prevalent positive NAO conditions around the late 1970s (e.g Hurrel, 1995; Kiely, 1999; Murphy et al., 2013). Previous studies found that the influence of the NAO is more prominent for winter, spring and autumn, with a negative correlation during the summer, following a NAO positive winter in Ireland (Kiely, 1999; Wang et al., 2006; Leahy and Kiely, 2011; Noone et al., 2015). Notably, changes in the contribution of extremes to annual totals appear to be less influenced by low frequency variability (see. Figure 4.7).

Other atmospheric pressure patterns can modulate and cause temporal non-stationarities in the relationships between the NAO index and European climate variability, particularly the East Atlantic (EA) and the Scandinavian (SCA) patterns. The East Atlantic (EA) pattern was first described by Wallace and Gutzler (1981) as anomalously high 500 mb height anomalies over the subtropical North Atlantic and eastern Europe when in positive mode. The Scandinavian (SCA) pattern, which corresponds to the Eurasia-1 pattern (Barnston and Livezey, 1987) is characterised by positive sea-level pressure (SLP) anomalies over Scandinavia and a more diffuse negative SLP centre over Greenland.

Comas-Bru and McDermott (2014) attribute much of the twentieth century multidecadal variability in the relationship between the North Atlantic Oscillation (NAO) and winter climate over the North Atlantic–European region to the combined effects of the NAO and either the East Atlantic pattern (EA) or the Scandinavian pattern (SCA). The study demonstrates how different NAO-EA and NAO-SCA combinations influence winter climatic conditions (temperature and precipitation) as a consequence of NAO dipole migrations. Findings suggest that the interannual variability in the north–south winter precipitation gradient in the UK appears to reflect the migration of the NAO dipole linked to linear combinations of the NAO and the EA. Furthermore, Hall and Hanna (2018) show that although the phase and magnitude of the NAO influences total UK rainfall, there are regional variations which it does not explain. The study examines potential links between North Atlantic atmospheric circulation indices and precipitation and temperature over the UK and show that in 2013/2014, precipitation anomalies in southern England can be attributed to the positive EA phase. These findings highlight the importance of considering secondary and tertiary atmospheric circulation indices when exploring regional and seasonal variations in precipitation extremes for Ireland. Future work to assess trends in seasonal and monthly extreme rainfall events will make it possible to assess the seasonal influence of the combined effects of the NAO and either the East Atlantic pattern (EA) or the Scandinavian pattern (SCA) on extremes for Ireland over longer timescales.

Furthermore, the Atlantic Multidecadal Oscillation (AMO) is linked to changes in weather patterns on both sides of the Atlantic Ocean and is an important feature in the variation of North Atlantic and Irish sea surface temperature records. Casanueva et al. (2014) show that the Atlantic Multidecadal Oscillation (AMO) index is strongly positively correlated with extreme precipitation (RR95pTOT) across all seasons in Europe, while McCarthy et al. (2015) highlight the relationship between the AMO and summer precipitation in Ireland. However, these studies are limited by the period of analysis (post 1940) as the AMO has been shown to have a periodicity of ~ 65-80 years (Enfield et al., 2001). Climate variability patterns occur in time scales that range from seasonal to centennial and are caused by both system´s internally generated variability and external forcings. Recently, Mann et al. (2021) examined the Atlantic Multidecadal Oscillation forcings through an ensemble of climate models and argue that volcanic forcing is the major influence on this variation mode and that the role of internal climate variability in twentieth century trends is limited (Haustein et al., 2019). As such, an improved understanding of the influence of the AMO on climate variability could help mitigate uncertainties. Future work with the long records produced here may be used to assess the stationarity of relationships between Irish precipitation and the AMO.

Finally, the availability of this data now makes it possible to investigate significant weather events documented in the historical records. For instance, the extreme rainfall event which occurred on the 25th August 1905. This rainfall event which resulted in severe and widespread flooding in the east of Ireland, was noted in the observations at several stations around the Dublin region. Notably, the event occurred exactly 81 years to the date before Hurricane Charley (1986) and 115 years to the date before Storm Francis (2020). Understanding how such events have changed over time can better inform estimates of the present-day risk of such events.

5.4 Concluding remarks

This research has made a number of important contributions. Data rescue is transforming understanding of historical climate, climate variability and change and makes a major contribution to understanding the nature of extremes and their impact on society. Despite its value, progress on data rescue remains slow, in large part due to the tedious, time consuming and painstaking workflow involved. This work has developed a highly innovative approach to data rescue in the classroom, creating multiple benefits for the pedagogical outcomes for students and progress in data rescue activities. This has significantly advanced data rescue efforts for Ireland with the approach also being adopted in other countries/universities. The thesis presents 3616 station years of daily precipitation data (~1.32 million daily values) and associated metadata for stations across Ireland that were previously unavailable to scientific scrutiny. The raw data and metadata are freely available through the edepositIreland data centre (http://hdl.handle.net/2262/91347). The derived data will also be hosted by Met Éireann and routinely updated and widely disseminated. This includes access to both the raw and quality assured datasets. Details of the data rescue process and dataset development are presented in publications to provide a fully traceable dataset, in line with best international practice. The data is already enabling diverse research within and beyond climatology and academia. The novel approach to data rescue developed as part of this thesis in a collaboration between Maynooth University and Met Éireann provides a framework for both Irish and global data rescue activities with huge potential for further expansion.

This work also makes a significant contribution to the homogenisation of historical daily precipitation records. This is a difficult task. The deployment of RHtests software shows the ability of the approach to detect discontinuities present in metadata and to sensibly correct daily series. However, the approach is not perfect. In particular, through application of additional break point tests after application of RHtests it was found that corrections applied in some stations may be too little. The study-design integrates additional checks throughout the process of homogenisation and trend assessment to ensure that any remaining issues around homogenisation are identified. This tiered approach of using multiple tests, together with the information content of the network in terms of timing of breaks and coherence of trends will be of value to others attempting to homogenise historical daily rainfall data. Up to now the dominant approach in the literature has been to discard daily data prior to identified break points, undermining significant efforts in data rescue. There can be a misconception that homogenisation is final. This is not the case and as further developments are made in approaches to homogenisation and additional data rescued to increase the spatial density of early records, the approaches developed here can be used as a benchmark against which to evaluate progress. A key contribution to this end has been the presentation of the raw data so that future researchers can go direct to source rather than trying to unpick adjustments to data through previous homogenisation attempts.

Finally, the analysis of trends presented marks the first assessment of the suite of indices defined by the CCl/ CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI) for Ireland. Results clearly show the large influence of low frequency variability for certain indicators (e.g. PRCPTOT). Climate variability will remain a dominant driver of change in Irish precipitation and the long records developed here will facilitate increased insight into the relationship between key modes of variability and precipitation characteristics. A key finding from this analysis is the persistent increasing trends in precipitation extremes in the southeast of the island, an insight that will be of interest to flood managers in developing adaptive responses.

References

Aguilar, E., Auer, I., Brunet, M., Peterson, T. C. and Wieringa, J. (2003) Guidelines on climate metadata and homogenization. World Meteorological Organization, WMO-TD No. 1186, WCDMP No. 53, Geneva, Switzerland, 55 pp.

Aguilar, E. and Prohom, M. (2011) RClimDex-extraQC (EXTRAQC Quality Control Software). User Manual, Centre for Climate Change, University Rovira i Virgili, Tarragona, Spain. http://www.c3.urv.cat/data/manual/Manual_rclim dex_extraQC.r.pdf (accessed 14th November 2020).

Aguilar, E. (2020) INDECIS Quality Control Software and Manual: INQC, beta version. Work Package 3 Deliverable 3.1.a. Centre for Climate Change (C3), Rovira i Virgili University (URV), Vila-seca – Spain.

Alexander, L.V., Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Klein Tank, A.M.G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., RupaKumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D.B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., Vazquez-Aguirre, J.L. (2006) Global observed changes in daily climate extremes of temperature and precipitation. J Geophys Res 111: D05109. https://doi.org/10.1029/2005jd006290.

Alexander, L.V., Fowler, H.J., Bador, M., Behrangi, A., Donat, M.G., Dunn, R., Funk, C., Goldie, J., Lewis, E., Rogé, M., Seneviratne, S.I., and Venugopal, V. (2019) On the use of indices to study extreme precipitation on sub-daily and daily timescales. Environmental Research Letters, 14(12), 125008.

Alexandersson, H. (1986) A homogeneity test applied to precipitation data. Journal of Climatology, 6, 661–675. doi:10.1002/joc.3370060607.

Alexandersson, H. and Moberg, A. (1997) Homogenization of Swedish temperature data. Part 1: homogeneity test for linear trends. International Journal of Climatology, 17 (1997), pp. 25-34

Ali, H., Fowler, H. J., and Mishra, V. (2018) Global observational evidence of strong linkage between dew point temperature and precipitation 15 extremes, Geophysical Research Letters, 45, 12–320.

Allan, R., Brohan, P., Compo, G.P., Stone, R., Luterbacher, J. and Brönnimann, S. (2011) The international atmospheric circulation reconstructions over the earth (ACRE) initiative. Bull. Am. Meteorol. Soc., 92, 1421–1425, doi:10.1175/2011BAMS3218.1.

Allan, R., Endfield, G., Damodaran, V., Adamson, G., Hannaford, M., Carroll, F., Macdonald, N., Groom, N., Jones, J., Williamson, F., Hendy, E., Holper, P., Arroyo-Mora, P., Hughes, L., Bickers, R. and Bliuc, A.-M. (2016) Toward integrated historical climate research: the example of Atmospheric Circulation Reconstructions over the Earth. Wiley Interdisciplinary Reviews: Climate Change, vol. 7, no. 2, pp. 164-174.

Ashcroft, L., Coll, J.R., Gilabert, A., Domonkos, P., Brunet, M., and Aguilar, E. (2018) A rescued dataset of sub-daily meteorological observations for Europe and the southern Mediterranean region, 1877–2012. Earth System Science Data, 10 (3), 1613– 1635. https://doi.org/10.5194/essd-10-1613-2018

Auer, I., Böhm, R., Jurković , A., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Brunetti, M., Nanni, T., Maugeri, M., Briffa, K., Jones, P., Efthymiadis, D., Mestre, O., Moisselin, JM., Begert, M., Brazdil, R., Bochníček, O., Cegnar, T., Gajić-Čapka, M., Zaninović, K., Majstorović, Ž., Szalai, S., Szentimery, T., Mercalli, L. (2005) A new instrumental precipitation dataset for the Greater Alpine Region for the period 1800–2002. Int. J. Climatol. 25(2): 139–166. https://doi.org/10.1002/joc.1135.

Auer, I., Böhm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P., Efthymiadis, D., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisselin, J.M., Begert, M., Müller-Westermeier, G., Kveton, V., Bochnicek, O., Stastny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic-Capka, M., Zaninovic, K., Majstorovic, Z. and Nieplova, E. (2007) HISTALP: historical instrumental climatological surface time series of the greater Alpine region. International Journal of Climatology, 27, 17–46.

Barnston AG, Livezey RE. (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. Mon. Weather Rev. 115 1083–1126.

Beaulieu, C., Ouarda, T., Seidou, O. (2010) A Bayesian normal homogeneity test for the detection of artificial discontinuities in climate series. International Journal of Climatology 30: 2342– 2357, DOI: 10.1002/joc.2056.

Begert M, Schlegel, T. and Kirchhofer W. (2005) Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. International Journal of Climatology, 25: 65–80, DOI: 10.1002/joc.1118.

Blenkinsop, S., Chan, S.C., Kendon, E.J., Roberts, N.M. and Fowler, H.J. (2015) Temperature influences on intense UK hourly precipitation and dependency on large-scale circulation, Environ. Res. Lett., 10, 054021, doi:10.1088/1748-9326/10/5.

Bonney, R., Shirk, J.L., Phillips, T.B., Wiggins, A., Ballard, H.L., Miller-Rushing, A.J., Parrish, J.K. (2014) Next steps for citizen science, Science, 343, pp. 1436-1437, 10.1126/science.1251554

Brönnimann, S., Brugnara, Y., Allan, R. J., Brunet, M., Compo, G. P., Crouthamel, R. I., Jones, P.D., Jourdain, S., Luterbacher, J., Siegmund, P., Valente, M.A. and Wilkinson, C.W. (2018) A roadmap to climate data rescue services. Geoscience Data Journal, 5, 28–39. https://doi.org/ 10.1002/gdj3.56

Brönnimann, S., Annis, J., Dann, W., Ewan, T., Grant, A.N., Griesser, T., et al. (2006) A guide for digitising manuscript climate data. Climate of the Past Discussions, 2, 191–207. https://doi.org/10.5194/cpd-2-191-2006

Brönnimann, S., Allan, R., Ashcroft, L., Baer, S., Barriendos, M., Brázdil, R. et al. (2019) Unlocking pre-1850 instrumental meteorological records: A global inventory. Bulletin of the American Meteorological Society. 100: ES389–ES413.

Brunet, M., Gilabert, A., Jones, P., and Efthymiadis, D. (2014) A historical surface climate dataset from station observations in Mediterranean North Africa and Middle East areas, Geosci. Data J., 1, 121–128, https://doi.org/10.1002/gdj3.12

Brunet, M., and Jones, P. (2011) Data rescue initiatives: Bringing historical climate data into the 21st century. Climate Res., 47, 29–40, https://doi.org/10 .3354/cr00960.

Butler, C.J., Coughlin, A.D.S. and Fee, D.T. (1998) Precipitation at Armagh Observatory 1838–1997. Biol. Environ. 98B(2): 123–140.

Butler, C.J., García-Suárez, A.M., Coughin, A.D.S. and Cardwell, D. (2005) Meteorological data recorded at Armagh Observatory: Vol. II – Daily, Mean Monthly, Seasonal and Annual, Maximum and Minimum Temperatures, 1844–2004. Armagh Observatory Climate Series.

Butler, C.J., Garcia-Suarez, A. and Palle, E. (2007) Trends and cycles in long Irish meteorological series. Biology and Environment: Proceedings of the Royal Irish Academy, 107B, 157– 165.

Casanueva, A., Rodríguez-Puebla, C., Frías, M. D. and González-Reviriego, N. (2014) Variability of extreme precipitation over Europe and its relationships with teleconnection patterns, Hydrol. Earth Syst. Sci., 18(2), 709–725, doi:10.5194/hess-18-709-2014.

Caesar, J., Alexander, L.V., Trewin, B., Tse-ring, K., Sorany, L., Vuniyayawa, V., Keosavang, N., Shimana, A., Htay, M.M., Karmacharya, J., et al. (2005) Changes in temperature and precipitation extremes over the Indo-Pacific region from 1971 to 2005. *Int J Climatol* 2011, **31**: 791– 801. doi:10.1002/joc.2118.

Chen, M. and Xie, P. (2008), Quality control of daily precipitation report at NOAA/CPC, AMS 12th conferences on IOAS-AOLS, New Orleans, LA.

Coleman, K., Murdoch, J., Rayback, S., Seidl, A., Wallin, K. (2017) Students' understanding of sustainability and climate change across linked service-learning courses, J. Geosci. Education. 65: 158–167.

Coll, J.R., van der Schrier, G., Aguilar, E., Rasol, D., Coscarelli, R., and Bishop, A. (2019) Daily rescued meteorological observations across Europe (1917-1990), PANGAEA. https://doi.pangaea.de/10.1594/PANGAEA.896957

Coll, J., Curley, C., Walsh, S. and Sweeney, J. (2018) HOMERUN: Relative Homogenisation of the Irish Precipitation Network. Environmental Protection Agency, Johnstown Castle Estate: Wexford, Ireland Report No. 242. EPA RESEARCH PROGRAMME 2014–2020.

Coll, J., Domonkos, P., Guijarro, J., Curley, M., Rustemeier, E., Aguilar, E., Walsh, S. and Sweeney, J. (2020) Application of homogenization methods for Ireland's monthly precipitation records: comparison of break detection results. International Journal of Climatology. 2020;1–20. http://doi.org/10.1002/joc.6575.

Comas-Bru L, McDermott F. (2014). Impact of the EA and SCA patterns on the European twentieth century NAO-winter climate relationship. Q. J. R. Meteorol. Soc. 140: 354–363. https://doi.org/10.1002/qj.2158.

Cornes, R. C., Jones, P. D., Briffa, K. R., and Osborn, T. J. (2012) A daily series of mean sea-level pressure for London, 1692–2007, International Journal of Climatology, 32, 641–656. https://doi.org/10.1002/joc.2301

Craddock, J.M. (1979) Methods of comparing annual rainfall records for climatic purposes. Weather, 34, 332–346.

Creilson, J.K., Pippin, M.R., Henderson, B.L. et al. (2008) Surface ozone measured at GLOBE schools in The Czech Republic: a demonstration of the importance of student contribution to the larger science picture. Bull. Am. Meteorol. Soc. 89(4): 505–514.

Dai, A., (2013) The influence of the inter-decadal Pacific Oscillation on U.S. precipitation during 1923–2010. Climate Dyn., 41, 633–646, doi:10.1007/s00382-012-1446-5.

Daly, C., Gibson, W.P., Taylor, G.H., Doggett, M.K. and Smith, J.I. (2007) Observer bias in daily precipitation measurements at U.S. Cooperative Network stations. Bull. Amer. Meteor. Soc., 88, 899–912.

De Jongh, I.L.M., Verhoest, N.E.C. and De Troch, F.P. (2006) Analysis of 105-year time series of precipitation observed at Uccle, Belgium. Int. J. Climatol. 26: 2023–2039.

Della-Marta, P. M. and Wanner, H., (2006) A method of homogenizing the extremes and mean of daily temperature measurements. J. Climate, 19, 4179–4197.

Dixon, F.E. (1987) Early Irish weather records. In Shields, L., eds. The Irish Meteorological Service: The First Fifty Years, 1936-1986. Dublin, Ireland: Met Éireann.

Domonkos, P. (2015) Homogenization of precipitation time series with ACMANT. Theoretical and Applied Climatology, 122, 303–314. http://doi.org/10.1007/s00704-014-1298-5.

Domonkos, P. (2013) Efficiencies of inhomogeneity-detection algorithms: comparison of different detection methods and efficiency measures. Journal of Climatology, 2013, 390945.

Domonkos, P. (2011) Adapted Caussinus-Mestre algorithm for networks of temperature series (ACMANT). International Journal of Geosciences, 2, 293–309.

Donat, M., Alexander, L., Yang, H., Durre, I., Vose, R., Dunn, R., Willett, K., Aguilar, E., Brunet, M., Caesar, J., et al. (2013) Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, Journal of Geophysical Research: Atmospheres, 118, 2098–2118.

Donat, M.G., Lowry, A.L., Alexander, L.V., O'Gorman, P.A. and Maher, N. (2016) More extreme precipitation in the world's dry and wet regions. Nat. Climate Change, 6, 508–513, https://doi.org/10.1038/nclimate2941.

Ducre-Robitaille, J.F., Vincent, L.A., and Boulet, G. (2003) Comparison of techniques for detection of discontinuities in temperature series. International Journal of Climatology 23: 1087–1101.

Dunn, R.J.H., Alexander, L.V., Donat, M.G., Zhang, X., Bador, M., Herold, N. et al. (2020) Development of an updated global land in-situ-based dataset of temperature and precipitation extremes: HadEX3 J. Geophys. Res. Atmos., 125, 10.1029/2019JD032263

Durre, I., Menne, M.J. Gleason, B.E., Houston T.G. and Vose R.S. (2010) Comprehensive automated quality assurance of daily surface observations. Journal of Applied Meteorology and Climatology, 49: 1615-1633.

Eden, P., (2009) Traditional weather observing in the UK: An historical overview. Weather, 64, 239–245. doi:10.1002/wea.469

Enfield, D. B., Mestas-Nuñez, A. M. and Trimble, P. J. (2001) The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S., Geophys. Res. Lett., 28(10), 2077–2080, doi:10.1029/2000GL012745.

Eischeid, J.K., Baker, C.B., Karl, T.R. and Diaz, H.F. (1995) The quality control of long-term climatological data using objective data analysis. Journal of Applied Meteorology 34: 2787–2795.

Santo, F.E., Ramos, A.M., de Lima, M.I.P. and Trigo, R.M. (2014). Seasonal changes in daily precipitation extremes in mainland Portugal from 1941 to 2007. Reg Environ Change 14, 1765–1788.

Fowler, H.J. and Kilsby, C.G. (2003a) Implications of changes in seasonal and annual extreme rainfall. Geophysical Research Letters, 30(13), 1720. https://doi.org/10.1029/2003GL017327.

Fowler, H.J. and Kilsby, C.G. (2003b) A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. International Journal of Climatology, 23(11), 1313–1334.

Freitas, L., Gonzalez Pereira, M., Caramelo, L., Mendes, M. and Nunes, L.F. (2013) Homogeneity of monthly air temperature in Portugal with HOMER and MASH. Quarterly Journal of the Hungarian Meteorological Service, 117, 69–90.

Glasspoole, J. (1952) The British Rainfall Organisation. Weather, 7 (3), 85–87. doi:10.1002/j.1477-8696.1952.tb01449.x

Gofa, F., Mamara, A., Anadranistakis, M., and Flocas, H. (2019) Developing Gridded Climate Data Sets of Precipitation for Greece Based on Homogenized Time Series. Climate, 7, 68. https://doi.org/10.3390/cli7050068

Green, J.S., Klatt, M., Morrissey, M. and Postawko, S. (2008) The Comprehensive Pacific Rainfall Database. Journal of Atmospheric and Oceanic Technology, 25, 71–82.

Gu, G. and Adler, R.F. (2015) Spatial patterns of global precipitation change and variability during 1901–2010. J Clim 28:4431–4453. doi:10.1175/JCLI-D-14-00201.1

Gubler, S., Hunziker, S., Begert, M., Croci-Maspoli, M., Konzelmann, T., Brönnimann, S., Schwierz, C., Oria, C. and Rosas, G. (2017) The influence of station density on climate data homogenization using HOMER. International Journal of Climatology, 37(13), 4670–4683.

Guijarro, J.A. (2018) Homogenization of climatic series with Climatol. http://www.Climatol.eu/homog_Climatol-en.pdf.

Hall, R.J. and Hanna, E. (2018) North Atlantic circulation indices: links with summer and winter UK temperature and precipitation and implications for seasonal forecasting. International Journal of Climatology, 38, e660.#x2013;e677. https://doi.org/10.1002/joc.5398.

Hamada, A., Arakawa, O. and Yatagai, A. (2011) An Automated Quality Control Method for Daily Rain-gauge Data. Global Environmental Research 15/2011: 183-192.

Hamed, K.H. and Rao, A.R., (1998) A modified Mann-Kendall trend test for autocorrelated data. Journal of Hydrology, 204, 182–196.

Harnik, P. G., and Ross, R.M. (2003) Assessing Data Accuracy When Involving Students in Authentic Paleontological Research. J. Geosci. Educ., 51, 76–84, doi:10.5408/1089-9995-51.1.76. http://nagt-jge.org/doi/10.5408/1089-9995-51.1.76.

Haustein, K., Otto, F. E. L., Venema, V., Jacobs, P., Cowtan, K., Hausfather, Z., et al. (2019). A limited role for unforced internal variability in twentieth-century warming. Journal of Climate, 32, 4893–4917. https://doi.org/10.1175/JCLI-D-18-0555.1

Hawkins, E., Burt S., Thomas S., Lockwood M., Brohan P., Richardson H. and Roy M. (2019) Hourly weather observations from the Scottish Highlands (1883-1904) rescued by volunteer citizen scientists. Geoscience Data Journal, 6, 2, (160-173). https://doi.org/10.1002/gdj3.79

Hegerl, G. C., Black, E., Allan, R. P., Ingram, W. J., Polson, D., Trenberth, K. E., Chadwick, R. S., Arkin, P. A., Sarojini, B. B., Becker, A., Dai, A., Durack, P. J., Easterling, D., Fowler, H. J., Kendon, E. J., Huffman, G. J., Liu, C., Marsh, R., New, M., Osborn, T. J., Skliris, N., Stott, P. A., Vidale, P.-L., Wijffels, S. E., Wilcox, L. J., Willett, K. M. and Zhang, X. (2015) Challenges in Quantifying Changes in the Global Water Cycle, Bull. Amer. Meteor. Soc., 96(7), 1097–1115, doi:10.1175/BAMS-D-13-00212.1.

Hollis, D., McCarthy, M.P., Kendon, M., Legg, T. and Simpson, I. (2019) HadUK-Grid – a new UK dataset of gridded climate observations. Geoscience Data Journal. 6: 151–159. https://doi.org/10.1002/gdj3.78

HOME (2011) Monitoring Progress Report 03/05/2007–01/06/2011. HOME – Advances in Homogenisation Methods of Climate Series: An Integrated Approach (COST Action ES0601), Chair of the Management Committee of the Action, Submitted to the Relevant Domain Committee.

Hunziker, S., Brönnimann, S., Calle, J., Moreno, I., Andrade, M., Ticona, L., Huerta, A., and Lavado-Casimiro, W. (2018) Effects of undetected data quality issues on climatological analyses, Clim. Past, 14, 1–20, https://doi.org/10.5194/cp-14-1-2018.

Hurrell, J. W. (1995) Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation, Science, 269(5224), 676–679, doi:10.1126/science.269.5224.676.

Iles, C. and Hegerl, G. (2017) Role of the North Atlantic Oscillation in decadal temperature trends. Environ Res Lett 12:114010, 10.1088/1748-9326/aa9152/pdf

IPCC (2014) Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Jones, M.R., Fowler, H.J., Kilsby, C.G. and Blenkinsop, S. (2013) An assessment of changes in seasonal and annual extreme rainfall in the UK between 1961 and 2009. Int J Climatol 33: 1178−1194

Kaspar, F., Tinz, B., Mächel, H., and Gates, L. (2015) Data rescue of national and international meteorological observations at Deutscher Wetterdienst, Adv. Sci. Res., 12, 57–61, https://doi.org/10.5194/asr-12-57-2015.

Kendon, E.J., Roberts, N.M., Fowler, H.J., Roberts, M.J., Chan, S.C., and Senior, C.A. (2014) Heavier summer downpours with climate change revealed by weather forecast resolution model, Nature Climate Change, 4, 570–576.

Khan, N., Shahid, S., Chung, E-S., Behlil, F., and Darwish, M.S.J. (2020) Spatiotemporal changes in precipitation extremes in the arid province of Pakistan with removal of the influence of natural climate variability. Theor Appl Climatol 142:1447–1462. https://doi.org/10.1007/s00704-020-03389-9

Kiely, G. (1999) Climate change in Ireland from precipitation and streamflow observations, Advances in Water Resources, 23(2), 141–151, doi:10.1016/S0309-1708(99)00018-4.

Killick, E. (2016) Benchmarking the Performance of Homogenisation Algorithms on Daily Temperature Data. PhD thesis. University of Exeter. http://hdl.handle.net/10871/23095

Klein Tank, A.M.G., Zwiers, F.W., and Zhang, X. (2009) Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation. WCDMP No 72, WMO/TD No 1500. World Meteorological Organization, Geneva, Switzerland.

Kosmala, M., Wiggins, A., Swanson, A. and Simmons, B. (2016) Assessing data quality in citizen science. Front. Ecol. Environ., 14, 551–560, doi:10.1002/fee.1436.

Kunkel, K.E., Easterling, D.R., Redmond, K., Hubbard, K., Andsager, K., Kruk, M. and Spinar, M. (2005) Quality control of pre-1948 cooperative observer network data. Journal of Atmospheric and Oceanic Technology, 22, 1691–1705.

Leahy, P. G. and Kiely, G. (2011) Short Duration Rainfall Extremes in Ireland: Influence of Climatic Variability, Water Resour Manage, 25(3), 987–1003, doi:10.1007/s11269-010- 9737-2.

Lenderink, G. and van Meijgaard, E. (2008) Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geosci.* 1, 511–514.

Lindau, R., Venema, V., (2016) The uncertainty of break positions detected by homogenization algorithms in climate records. Int. J. Climatol. http://dx.doi.org/ 10.1002/joc.4366.

Llabrés-Brustenga, A., Rius, A., Rodríguez-Solà, R. et al. (2019) Quality control process of the daily rainfall series available in Catalonia from 1855 to the present. Theor Appl Climatol 137, 2715–2729, https://doi.org/10.1007/s00704-019-02772-5.

Mann, M. E., Steinman, B. A., Brouillette, D. J., and Miller, S. K. (2021). Multidecadal climate oscillations during the past millennium driven by volcanic forcing. Science, 371(6533), 1014–1019. https://doi.org/10.1126/science.abc5810.

Maraun, D., Osborn, T. J. and Gillett, N. P. (2008) United Kingdom daily precipitation intensity: improved early data, error estimates and an update from 2000 to 2006, Int. J. Climatol., 28(6), 833–842, doi:10.1002/joc.1672.

Maraun, D., Osborn, T. J., and Rust, H. W. (2011) The influence of synoptic airflow on UK daily precipitation extremes. Part I: Observed spatio-temporal relationships, Clim. Dynam., 36, 261-275, doi:10.1007/s00382-009-0710-9.

Mateus, C., Potito, A., and Curley, M. (2019) Reconstruction of a long-term historical daily maximum and minimum air temperature network dataset for Ireland (1831-1968). Geoscience Data Journal.

Mateus, C., Potito, A., Curley, M. (2020). Engaging secondary school students in climate data rescue through service-learning partnerships. Weather. https://doi.org/10.1002/wea.3841.

Mateus, C (2021) Searching for historical meteorological observations in the Island of Ireland. Weather. https://doi.org/10.1002/wea.3887

Matthews, T., Murphy, C., Wilby, R.L. and Harrigan, S. (2014) Stormiest winter on record for Ireland and UK. Nature Climate Change, 4, pp. 738-740. https://doi.org/10.1038/nclimate2336

Matthews, T., Murphy, C., Wilby, R.L., and Harrigan, S. (2015) A cyclone climatology of the British-Irish Isles 1871–2012. Int. J. Climatol., doi: 10.1002/joc.4425.

McCarthy, G.D., Gleeson, E. and Walsh S. (2015) The influence of ocean variations on the climate of Ireland. Weather, 70, pp. 242-245, doi/full/10.1002/wea.2543.

McElwain, L. and Sweeney, J. (2007) Key Meteorological Indicators of Climate Change in Ireland. Environmental Protection Agency, Johnstown Castle Estate: Wexford, Ireland.

McKeown, M., Potito, A. and Hickey, K. (2012) The long-term temperature record from Markree Observatory, County Sligo, from 1842 to 2011. Irish Geography, 45(3), 257–282.

Mestre, O., Domonkos, P., Picard, F., Auer, I., Robin, S., Lebarbier, E., Boehm, R., Aguilar, E., Guijarro, J., Vertachnik, G., Klancar, M., Dubuisson, B. and Stepanek, P. (2013) HOMER: a homogenization software—methods and applications. Idöjárás. Quarterly Journal of the Hungarian Meteorological Service, 117(1), 47–67.

Min, S-K., Zhang, X., Zwiers, F. W. and Hegerl, G. C. (2011) Human contribution to more-intense precipitation extremes. Nature 470, 378–381.

Mitchell, N., Triska, M., Liberatore, A., Ashcroft, L., Weatherill, R. and Longnecker, N. (2017) Benefits and challenges of incorporating citizen science into university education. PLoS ONE 12(11): e0186285. https://doi.org/10.1371/journal.pone.0186285

Moberg, A., Jones, P.D., Lister, D., Walther, A., Brunet, M., Jacobeit, J., Alexander, L.V., Della-Marta, P.M., Luterbacher, J., Yiou, P., Chen, D., Klein Tank, A.M.G., Saladié, O., Sigró, J., Aguilar, E., Alexandersson, H., Almarza, C., Auer, I., Barriendos, M., Begert, M., Bergström, H., Böhm, R., Butler, C.J., Caesar, J., Drebs, A., Founda, D., Gerstengarbe, F-W., Micela, G., Maugeri, M., Österle, H., Pandzic, K., Petrakis, M., Srnec, L., Tolasz, R., Tuomenvirta, H., Werner, P.C., Linderholm, H., Philipp, A., Wanner, H. and Xoplaki, E. (2006) Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901–2000. Journal of Geophysical Research 111: D22, 106. DOI: 10.1029/2006JD007103.

Murphy, C., Harrigan, S., Hall, J. and Wilby, R. L. (2013) Climate-driven trends in mean and high flows from a network of reference stations in Ireland, Hydrological Sciences Journal, 58(4), 755–772, doi:10.1080/02626667.2013.782407.

Murphy, C., Noone, S., Duffy, C., Broderick, C., Matthews, T. and Wilby R.L. (2017) Irish droughts in newspaper archives: rediscovering forgotten hazards? Weather 72 (6), 151-155

Murphy, C., Broderick, C., Burt, T. P., Curley, M., Duffy, C., Hall, J., Harrigan, S., Matthews, T. K. R., Macdonald, N., McCarthy, G., McCarthy, M. P., Mullan, D., Noone, S., Osborn, T. J., Ryan, C., Sweeney, J., Thorne, P. W., Walsh, S., and Wilby, R. L. (2018) A 305-year continuous monthly rainfall series for the island of Ireland (1711–2016). Climate of the Past, 14, 413–440. https://doi.org/10.5194/cp-14-413-2018.

Murphy, C., Wilby, R.L., Matthews, T.K.R., Thorne, P., Broderick, C., Fealy, R., Hall, J., Harrigan, S., Jones, P., McCarthy, G., MacDonald, N., Noone, S. and Ryan, C. (2019) Multi-century trends to wetter winters and drier summers in the England and Wales precipitation series explained by observational and sampling bias in early records. International Journal of Climatology, https://doi.org/10.1002/joc.6208.

Nemec, J., Gruber, C., Chimani, B. and Auer, I. (2013), Trends in extreme temperature indices in Austria based on a new homogenized dataset, Int. J. Climatol., doi:10.1002/joc.3532.

Noone, S., Murphy, C., Coll, J., Matthews, T., Mullan, D., Wilby, R.L. and Walsh, S. (2015) Homogenization and analysis of an expanded long-term monthly rainfall network for the Island of Ireland (1850–2010). International Journal of Climatology, 36, 2837–2853. https://doi.org/10.1002/joc.4522.

Noone, S., Broderick, C., Duffy, C., Matthews, T., Wilby, R. L., and Murphy, C. (2017) A 250-year drought catalogue for the island of Ireland (1765–2015). International Journal of Climatology, 37, 239–254. https://doi.org/10.1002/joc.4999.

O'Gorman, P. A. (2015) Precipitation extremes under climate change. Curr. Clim. Change Rep. 1, 49–59.

Osborn, T.J., Hulme, M., Jones, P.D. and Basnett, T.A. (2000) Observed trends in the daily intensity of United Kingdom precipitation. International Journal of Climatology, 20(4), 347– 364.

Pall, P., Allen, M. R., & Stone, D. A. (2007). Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO2 warming. Climate Dynamics, 28(4), 351–363. https://doi.org/10.1007/s00382-006-0180-2.

Papalexiou, S. M. and Montanari, A. (2019) Global and regional increase of precipitation extremes under global warming. Water Resources Research, 2018WR024067.

Park, E.G., Burr, G., Slonosky, V., Sieber, R. and Podolsky, L. (2018), "Data rescue archive weather (DRAW): Preserving the complexity of historical climate data", Journal of Documentation, Vol. 74 No. 4, pp. 763-780. https://doi.org/10.1108/JD-10-2017-0150

Peterson, T.C., Easterling, D.R., Karl, T.R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salingeri, J., Førland, E.J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., Parker, D. (1998) Homogeneity adjustments of in situ atmospheric climate data: A review. Int. J. Climatol., 18, 1493–1517.

Peterson, T.C., and Manton, M.J. (2008), Monitoring changes in climate extremes: A tale of international collaboration, Bulletin of the American Meteorological Society, 89, 1266–1271. DOI:10.1175/2008BAMS2501.1.

Pettitt, A.N. (1979) A non-parametric approach to the change-point detection. Applied Statistics., 28 (2), pp. 126-135.

Pfister, L., Hupfer, F., Brugnara, Y., Munz, L., Villiger, L., Meyer, L., Schwander, M., Isotta, F. A., Rohr, C., and Brönnimann, S. (2019) Early instrumental meteorological measurements in Switzerland, Clim. Past, 15, 1345–1361, https://doi.org/10.5194/cp-15-1345-2019.

Plummer, N., Allsopp, T. and Lopez, J.A. (2003) Guidelines on climate observations networks and systems. WCDMP No 53, WMO/TD No 1186. World Meteorological Organization, Geneva, Switzerland

Reeves, J., Chen, J., Wang, X.L., Lund, R. and Lu, Q. (2007), A review and comparison of changepoint detection techniques for climate data, J. Appl. Meteorol. Climatol., 46, 900–915.

Rennie J.J., Lawrimore, J.H., Gleason, B., Thorne, P.W., Morice, C.M., Menne, M.J., Williams, C.N., Gambi de Almeida, W., Christy, J., Flannery, M., Ishihara, K., Kamiguchi, K., KleinTank, A.M.G., Mhanda, A., Lister, D., Razuvaev, V., Renom, M., Rusticucci, M., Tandy, J., Worley, S., Venema, V., Angel, W., Brunet, M., Dattore, B., Diamond, H., Lazzara, M., Le Blancq, F., Luterbacher, J., Mächel, H., Revadekar, J., Vose, R., and Yin, X. (2014) The international surface temperature initiative global land surface databank: monthly temperature data release description and methods, Geosci. Data J., 1, 75–102, doi:10.1002/gdj3.8.

Rhines, A., Tingley, M.P., McKinnon, K.A. and Huybers, P. (2015) Decoding the precision of historical temperature observations. Q. J. R. Meteorol. Soc. 141: 2923–2933, doi: 10.1002/qj.2612.

Ribeiro, S., Caineta, J., and Costa, A.C.C. (2016) Review and discussion of homogenisation methods for climate data. Phys. Chem. Earth 94: 167–179. https://doi.org/10.1016/j.pce.2015.08.007.

Ryan, C., Duffy, C., Broderick, C., Thorne, P.W., Curley, M., Walsh, S., Daly, C., Treanor, M. and Murphy, C. (2018) Integrating data rescue into the classroom. Bulletin of the American Meteorological Society, 99 (9), 1757–1764. https://doi.org/10.1175/BAMS-D17-0147.1

Ryan, C., Murphy, C., McGovern, R., Curley, M., Walsh, S. and 476 students (2020) Ireland's pre-1940 daily rainfall records. Geoscience Data Journal, 1–13. https://doi.org/10.1002/gdj3.1

Santo, F.E., Ramos, A.M., de Lima, M.I.P. and Trigo, R.M. (2014). Seasonal changes in daily precipitation extremes in mainland Portugal from 1941 to 2007. Reg Environ Change 14, 1765–1788.

Santos, M. and Fragoso, M. (2013) Precipitation variability in Northern Portugal: data homogeneity assessment and trends in extreme precipitation indices. Atmos. Res., 131 (2013), pp. 34-45, 10.1016/j.atmosres.2013.04.008

Sen, P. K. (1968) Estimates of the regression coefficient based on Kendall's tau, J. Am. Stat. Assoc., 63, 1379-1389.

Seneviratne, S.I., Nicholls, N., Easterling, D., Goodess, C.M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C. and Zhang, X. (2012) Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. A Special Report of Working Groups I and II of the Intergovernmental

Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 109-230.

Shah, H.R. and Martinez, L.R. (2016) Current approaches in implementing citizen science in the classroom, Journal of Microbiology & Biology Education, 17, https://doi.org/10.1128/jmbe.v17i1.1032.

Sheridan, T. (2001) Analysis of trends at some Irish rainfall stations. Dublin, Ireland: Met Éireann Technical Report 59.

Shrestha, A.B., Bajracharya, S.R., Sharma, A.R. et al. (2016) Observed trends and changes in daily temperature and precipitation extremes over the Koshi river basin 1975–2010. Int. J. Climatol., 10.1002/joc.476.

Simpson, I.R. and Jones, P.D. (2014) Analysis of UK precipitation extremes derived from Met Office gridded data. International Journal of Climatology, 34(7), 2438– 2449.

Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., Kelder, T., Kowal, K., Lees, T., Matthews, T., Murphy, C., and Wilby, R.L. (2020) Nonstationary weather and water extremes: a review of methods for their detection, attribution, and management, Hydrol. Earth Syst. Sci. Discuss. [preprint], https://doi.org/10.5194/hess-2020-576.

Slonosky, V., Sieber, R., Burr, G., Podolsky, L., Smith, R., Bartlett, M. et al. (2019) From books to bytes: A new data rescue tool. Geoscience Data Journal. 6:58-73. https://doi.org/10.1002/gdj3.62.

Štěpánek, P., Zahradníček, P. and Farda, A. (2013) Experiences with data quality control and homogenization of daily records of various meteorological elements in the Czech Republic in the period 1961–2010. Q. J. Hung. Meteorol. Serv. 117: 123– 141.

Stickler, A., Brönnimann, S., Valente, M.A., Bethke, J., Sterin, A., Jourdain, S. Roucaute, E. Vasquez, M.V., Reyes, D.A., Allan, R. and Dee, D. (2014) ERA-CLIM: Historical Surface and Upper-Air Data for Future Reanalyses. B. Amer. Meteorol. Soc., 95, 1419–1430.

Sun, Q., Miao, C., Duan, Q., Kong, D., Ye, A., Di, Z., et al. (2014). Would the ‘real’ observed dataset stand up? A critical examination of eight observed gridded climate datasets for China. Environ. Res. Lett. 9, 015001. doi:10.1088/1748-9326/9/1/015001

Sun, Q., Zhang, X., Zwiers, F., Westra, S., and Alexander, L. V. (2020) A global, continental and regional analysis of changes in extreme precipita20 tion, Journal of Climate, pp. 1–52.

Szentimrey, T. (1999) Multiple analysis of series for homogenization (MASH). In: 2nd Seminar for Homogenization of Surface Climatological Data. Geneva: WMO, pp. 27–46, WCDMP-41.

Szentimrey, T., Bihari, Z. and Szalai, S. (2007) Comparison of geostatistical and meteorological interpolation methods (What is What?). In Spatial Interpolation for Climate Data: The Use of GIS in Climatology and Meteorology;ISTE Ltd.: Newport Beach, CA, USA; pp. 45–56.

Szentimrey, T. (2011) Theoretical questions of daily data homogenisation and saily benchmark results for MASH. In: Proceedings of the 7th Seminar for Homogenization and Quality Control in Climatological Databases. Geneva: WMO, pp. 45–55, WCDMP-85.

Szentimrey, T. (2014) Manual of homogenization software MASHv3.03. In. Hungarian Meteorological Service, Budapest, Hungary.

Tao, T., Chocat, B., Liu, S. and Xin, K. (2009) Uncertainty Analysis of Interpolation Methods in Rainfall Spatial Distribution–A Case of Small Catchment in Lyon. J. Water Resour. Prot., 1, 136–144.

Theil, H. (1950) A Rank-invariant Method of Linear and Polynomial Regression Analysis, I., Nederlands Akad. Wetensch. Proc., 53, 386-392.

Thorne, P., Allan, R. J., Ashcroft, L., Brohan, P., Dunn, R. H., Menne, M., Pearce, P., Picas, J., Willett, K., Benoy, M., et al. (2018) Toward an integrated set of surface meteorological observations for climate science and applications, Bulletin of the American Meteorological Society, 98, 2689–2702.

Toreti, A., Kuglitsch, F.G., Xoplaki, E., and Luterbacher, J. (2012) A novel approach for the detection of inhomogeneities affecting climate time series. Journal of Applied Meteorology and Climatology, 51: 317–326, DOI: 10.1175/JAMC-D-10-05033.1.

Toreti, A., Kuglitsch, F.G., Xoplaki, E., Della-Marta, P.M., Aguilar, E., Prohom, M. and Luterbacher, J. (2011), A note on the use of the standard normal homogeneity test to detect inhomogeneities in climatic time series, Int. J. Climatol., 31, 630–632, doi:10.1002/joc.2088.

Treanor, M., Keane, A. and Lonergan, T. (2011) Meteorological Observations in Ireland 1855-1976: an archival Finding Aid. Technical Note 6. Met Éireann.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B. (2003) The changing character of precipitation, B. Am. Meteorol. Soc., 84, 1205–1218.

Trenberth, K. E. (2011) Changes in precipitation with climate change, Clim Res, 47(1-2), 123– 138. https://doi.org/10.3354/cr00953.

Trewin, B.C. (2010) Exposure, instrumentation and observing practice effects on land temperature measurements. Wiley Interdisciplinary Reviews: Climate Change, 1: 490–506.

Tsidu, G.M. (2012) High-resolution monthly rainfall database for Ethiopia: homogenization, reconstruction, and gridding. Journal of Climate, 25(24), 8422–8443. https://doi.org/10.1175/JCLI-D-12-00027.1.

Tye, M.R. (2015) Understanding the risks from extreme rainfall. Proceedings of the Institution of Civil Engineers: Forensic Engineering, 168(2), 71– 80.

van den Besselaar E.J.M., Klein Tank A.M.G. and Buishand, T.A. (2013) Trends in European precipitation extremes over 1951-2010. Int J Climatol 33:2682–2689.

van der Velde, T., Milton, D.A., Lawson, T.J., Wilcox, C., Lansdell, M., Davis, G. Perkins, G. and Hardesty, B.D. (2017) Comparison of marine debris data collected by researchers and citizen scientists: Is citizen science data worth the effort? Biol. Conserv., 208, 127–138.

Venema, V., Trewin, B., Wang, X., Szentimrey, T., Lakatos, M., Aguilar, E., Auer, I., Guijarro, J. A., Menne, M., Oria, C., Louamba, W.S.R.L., Rasul, G. (2018) Guidance on the homogenization of climate station data. WMO, Geneva. https://doi.org/10.31223/osf.io/8qzrf

Venema, V.K.C., Mestre, O., Aguilar, E., Auer, I., Guijarro, J.A., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P., Viarre, J., Müller-Westermeier, G., Lakatos, M., Williams, C.N., Menne, M.J., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andresen, L., Acquaotta, F., Fratianni, S., Cheval, S., Klancar, M., Brunetti, M., Gruber, C., Prohom Duran, M., Likso, T., Esteban, P., Brandsma, T. (2012) Benchmarking homogenization algorithms for monthly data. Clim. Past 8(1), 89–115.

Vertačnik, G., Dolinar, M., Bertalanič, R., Klančar, M., Dvoršek, D. and Nadbath, M. (2015) Ensemble homogenization of Slovenian monthly air temperature series. International Journal of Climatology, 35, 4015–4026. https://doi.org/10.1002/joc.4265.

Vitone, T., Stofer, K.A., Steininger, M.S. Hulcr, J., Dunn, R. and Lucky, A. (2016) School of ants goes to college: Integrating citizen science into the general education classroom increases engagement with science. J. Sci. Commun., 15, 1–24.

Von Neumann, J. (1941) Distribution of the ratio of the mean square successive difference to the variance. Ann. Math. Stat., 13, pp. 367-395.

Wallace, J.M. and Gutzler, D.S. (1981) Teleconnections in the geopotential height field during the Northern Hemisphere winter. Mon. Weather Rev. 109 784–812.

Walsh, S. (2016) Long-term rainfall averages for Ireland. Met Éireann Climatological Note 15. Dublin.

Walsh, S. and Dwyer, N. (2012) Rainfall. In: The Status of Ireland’s Climate, 2012. Dwyer N (ed). CCRP Report No. 26. EPA, Wexford, Ireland.

Walsh, S. (2013) Setting the scene: the climate of Ireland 1900-2012. In Gleeson, E., McGrath, R. and Treanor, M., eds. Ireland’s climate: the road ahead. Dublin, Ireland: Met Éireann. Available at: http://hdl.handle.net/2262/71304

Wang, G., Wang, D., Trenberth, K. E., Erfanian, A., Yu, M., Bosilovich, M. G., and Parr, D. T. (2017). The peak structure and future changes of the relationships between extreme precipitation and temperature. Nature Climate Change, 7(4), 268–274. https://doi.org/10.1038/ nclimate3239

Wang, X. L., Chen, H., Wu, Y. Feng, Y. and Pu, Q. (2010) New techniques for detection and adjustment of shifts in daily precipitation data series. J. Appl. Meteor. Climatol., 49 (No. 12), 2416-2436. DOI: 10.1175/2010JAMC2376.1

Wang, X. L. (2008a) Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal t or F test. J. Appl. Meteor. Climatol., 47, 2423-2444.

Wang, X. L. (2008b) Penalized maximal F-test for detecting undocumented mean-shifts without trend-change. J. Atmos. Oceanic Tech., 25 (No. 3), 368-384. doi:10.1175/2007/JTECHA982.1.

Wang, X. L., Wen, Q.H. and Wu, Y. (2007) Penalized maximal t test for detecting undocumented mean change in climate data series. J. Appl. Meteor. Climatol., 46 (No. 6), 916-931. DOI:10.1175/JAM2504.1

Wang, X. L. and Feng, Y. (2013) RHtests_dlyPrcp User Manual. Climate Research Division, Atmospheric Science and Technology Directorate, Science and Technology Branch, Environment Canada. 17 pp. [Available online at http://etccdi.pacificclimate.org/software.shtml]

Wang, X.L. and Feng, Y. (2010) RHtestsV3 user manual. Climate Research Division, Science and Technology Branch, Environment Canada, Toronto, ON, Canada. [Available at http:// cccma.seos.uvic.ca/ETCCDMI/RHtest/RHtestsV3_UserManual. doc.]

Westerberg I, Walther A, Guerrero J-L, Coello Z, Halldin S, Xu C-Y, Chen D, and Lundin L-C. (2009) Precipitation data in a mountainous catchment in Honduras: quality assessment and spatiotemporal characteristics. Theor. Appl. Climatol. 101: 381–396, doi: 10.1007/ s00704-009-0222-x

Westra, S., Alexander, L. V. and Zwiers, F. W. (2013) Global increasing trends in annual maximum daily precipitation. J. Clim. 26, 3904–3918.

Wijngaard, J.B., Klein Tank, A.M.G. and Konnen, G.P. (2003) Homogeneity of 20th century European daily temperature and precipitation series. International Journal of Climatology, 23, 679–692.

Wilby, R.L., Clifford, N.J., De Luca, P., Harrigan, S., Hillier, J.K., Hodgkins, R., Johnson, M.F., Matthews, T.K., Murphy, C., Noone, S.J., Parry, S., Prudhomme, C., Rice, S., Slater, L., Smith, K. and Wood, P. (2017) The 'dirty dozen'of freshwater science: detecting then reconciling hydrological data biases and errors, Wiley Interdisciplinary Reviews (WIREs) Water, 4 (3), e1209.

Wilby, R.L., Noone, S., Murphy, C., Matthews, T., Harrigan, S. and Broderick, C. (2016) An evaluation of persistent meteorological drought using a homogeneous Island of Ireland precipitation network. International Journal of Climatology, 36, 2854– 2865. https://doi.org/10.1002/joc.4523

Wilkinson, C., Brönnimann, S., Jourdain, S., Roucaute, E., Crouthamel, R. and IEDRO Team; Brohan, P., Valente, A., Brugnara, Y., Brunet, M., Compo, G.P. and Gilabert, A. (2019) Best Practice Guidelines for Climate Data Rescue v1, of the Copernicus Climate Change Service Data Rescue Service [Ref: C3S_DC3S311a_Lot1.3.4.1_2019_v1 - contract: 2019/C3S_311a_Lot1_Met Office/SC2]. DOI: 10.24381/x9rn-mp92, https://doi.org/10.24381/x9rn-mp92

World Meteorological Organization (1986) Guidelines on the Quality Control of Surface Climatological Data. WMO No. 111. World Meteorological Organization, Geneva, Switzerland.

World Meteorological Organization (2011) Guide to Climatological Practices (3rd edn). WMO No. 100. World Meteorological Organization, Geneva, Switzerland.

World Meteorological Organization (2016) Guidelines on Best Practices for Climate Data Rescue 2016 (WMO-No. 1182). Geneva.

World Meteorological Organization (2020) Guidelines on climate metadata and homogenization (WMO/TD-No. 1186, WCDMP No. 53). Geneva.

Yosef, Y., Aguilar, E. and Alpert, P. (2019) Changes in extreme temperature and precipitation indices: Using an innovative daily homogenized database in Israel. Int J Climatol, 1–24. https://doi.org/10.1002/joc.6125

Yue, S. and Hashino, M. (2007) Long term trends of annual and monthly precipitation in Japan. Journal of the American Water Resources Association, Vol 39 (3), 587-596.

Zhang, X. and Yang, F. (2004) R Climdex (1.0) User Manual. Climate Research Branch Environment Canada Downsview, Ontario Canada.

Zolina, O. (2012) Changes in intense precipitation in Europe. In: Kundzewicz, Z.W. (ed) Changes in flood risk in Europe, Special Publication No. 10. IAHS Press, Wallingford, pp 96–120. https://doi.org/10.1201/b12348-8

Zwiers, F. W., Alexander, L.V., Hegerl, G.C., Knutson, T.R, Kossin, J. et al. (2012). Challenges in estimating and understanding recent changes in the frequency and intensity of extreme climate and weather events. In: Climate Science for Serving Society: Research, Modelling and Prediction Priorities. Asrar, G.R. and Hurrell, J.W., Eds. Springer.

Appendix I: Authorship Declaration Forms

Maynooth University
Department of Geography

This form is to accompany the submission of a PhD that contains research reported in published or unpublished work. **Please include one copy of this form for each co-authored work.** This form along with the published work should, under normal circumstances, appear in an Appendix.

Authorship Declaration Form

Publication Details:

Thesis Chapter/Pages	Chapter 2
Publication title	Integrating data rescue into the classroom
Publication status	Published
Type of publication	Journal article
Publication citation	Ryan, C., Duffy, C., Broderick, C., Thorne, P.W., Curley, M., Walsh, S., Treanor, M., Daly, C. and Murphy, C.. (2018) Integrating data rescue into the classroom. *Bulletin of the American Meteorological Society*, 99(9), 1757– 1764. https://doi.org/10.1175/BAMS-D17-0147.1

Nature/extent of my contribution to the work detailed above is as follows:

Nature/Extent of Contribution	**Lead author**
Design and management of data rescue project Station selection – selected data to be transcribed based on record length, continuity and spatial distribution of station network. Transcription of historical records. Comparison of double keyed data. Metadata collection Post-processing of transcribed data. Evaluation of student experience. Write paper for publication.	☒ Yes ☐ No

The following co-authors contributed to the work (all contributing co-authors):

Name	**Nature of contribution**
Conor Murphy	Proof-reading, minor edits to the original text
Peter Thorne	Proof-reading, ad hoc advice, hosted ACRE meeting that inspired the effort
Mary Curley	Proof-reading, ad hoc advice, minor edits to text
Seamus Walsh	Proof-reading, ad hoc advice, minor edits to text
Conor Daly	Supply of source imagery. Design consultation on input forms. Ongoing consultancy during transcription process.
Catriona Duffy	Assisted with post-processing of transcribed data
Ciaran Broderick	Assisted with post-processing of transcribed data
Mairead Treanor	Guest speaker at Maynooth University, advice on library collections.

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-author's contribution to this work

	Name	Signature	Date
Student	**Ciara Ryan**		27/01/2021
Co-author 1	**Conor Murphy**		27/01/2021
Co-author 2	**Mary Curley**		22/02/2021
Co-author 3	**Seamus Walsh**		20/02/2021
Co-author 4	**Conor Daly**		01/02/2021
Co-author 5	**Peter Thorne**		28/01/2021
Co-author 6	**Ciaran Broderick**		28/01/2021
Co-author 7	**Catriona Duffy**		28/01/2021
Co-author 8	**Mairéad Treanor**		02/02/2021

Maynooth University
Department of Geography

This form is to accompany the submission of a PhD that contains research reported in published or unpublished work. **Please include one copy of this form for each co-authored work.** This form along with the published work should, under normal circumstances, appear in an Appendix.

Authorship Declaration Form

Publication Details:

Thesis Chapter/Pages	Chapter 3
Publication title	Ireland's pre 1940 daily rainfall records
Publication status	Published
Type of publication	Journal article
Publication citation	Ryan, C., Murphy, C., McGovern, R., Curley, M., Walsh, S. and 476 students (2020) Ireland's pre-1940 daily rainfall records. *Geoscience Data Journal*, 1–13. https://doi.org/10.1002/gdj3.1

Nature/extent of my contribution to the work detailed above is as follows:

Nature/Extent of Contribution	**Lead author**
Imaging and transcription of historical records. Post-processing of datasets. Collection of metadata from historical records. Created individual station folders containing data file and metadata file per station. Created files/table of station metadata. Error checking of transcribed data Write paper for publication.	☒ Yes ☐ No

The following co-authors contributed to the work (all contributing co-authors):

Name	**Nature of contribution**
Conor Murphy	Supervisor; Proof-reading, minor edits to the original text.
Rhonda McGovern	Assisted in the scanning and transcription of historical precipitation records.
Mary Curley	Proof-reading, ad hoc advice, minor edits to text
Seamus Walsh	Proof-reading, ad hoc advice, minor edits to text

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-author's contribution to this work

	Name	Signature	Date
Student	Ciara Ryan	Ciara Ryan.	27/01/2021
Co-author 1	Conor Murphy	Conor Murphy	27/01/2021
Co-author 2	Rhonda McGovern	Rhonda McGovern	08/01/2020
Co-author 3	Mary Curley	Mary Curley	22/02/2021
Co-author 4	Seamus Walsh	Seamus Walsh	20/02/2021

Maynooth University
Department of Geography

This form is to accompany the submission of a PhD that contains research reported in published or unpublished work. **Please include one copy of this form for each co-authored work.** This form along with the published work should, under normal circumstances, appear in an Appendix.

Authorship Declaration Form

Publication Details:

Thesis Chapter/Pages	Chapter 4
Publication title	Long-term trends in extreme precipitation indices in Ireland
Publication status	Article submitted for review
Type of publication	Journal article
Publication citation	Ryan, C., Curley, M., Walsh, S. and Murphy, C.. (2021) Long-term trends in extreme precipitation indices in Ireland. International Journal of Climatology.

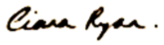
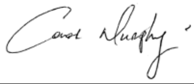
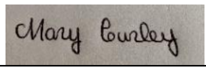
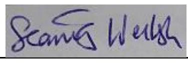
Nature/extent of my contribution to the work detailed above is as follows:

Nature/Extent of Contribution	Lead author
Application of quality control methods. Derived long-term station series. Application of homogenisation methods. Calculation of ETCCDI indices. Analysis of trends. Presentation of results. Write paper for publication.	☒ Yes ☐ No

The following co-authors contributed to the work (all contributing co-authors):

Name	Nature of contribution
Conor Murphy	Supervisor; Proof-reading, minor edits to the original text
Mary Curley	Proof-reading, ad hoc advice, minor edits to the original text
Seamus Walsh	Proof-reading, ad hoc advice, minor edits to text

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-author's contribution to this work

	Name	Signature	Date
Student	Ciara Ryan		27/01/2021
Co-author 1	Conor Murphy		27/01/2021
Co-author 2	Mary Curley		22/02/2021
Co-author 3	Seamus Walsh		20/02/2021

IN BOX
INSIGHTS and INNOVATIONS

Integrating Data Rescue into the Classroom

Ciara Ryan, Catriona Duffy, Ciaran Broderick, Peter W. Thorne, Mary Curley, Séamus Walsh, Conor Daly, Mairéad Treanor, and Conor Murphy

Historical observations are fundamental for advancing understanding of past changes in climate. Fundamentally, what you do not observe, you cannot understand. Globally, however, the temporal extent of digitized records to this day remains grossly incomplete, with many records available in image or hardcopy formats only (Allan et al. 2011; Brunet and Jones 2011). The task of digitizing these data is daunting, both logistically and in terms of the sheer volume of records that still require rescue.

The capacity to extend current observational data holdings is largely dependent on the resources available to carry out the digitization and transcription process. Some records have been, and continue to be, rescued via professional keying, but this practice is far from sufficient to key all known records. Engagement of nonexperts or "citizen scientists" on a voluntary basis has become increasingly integral to the rescue and refinement of observational data across multiple scientific disciplines (Bonney et al. 2014). The success of ongoing citizen science applications—for example, OldWeather.org (www.oldweather.org/) and Data. Rescue@Home (www.data-rescue-at-home.org)—underscores the potential of crowdsourcing as a data rescue strategy.

To date, however, the potential of university students to engage in data rescue activities for credit has not been systematically explored. Students represent a pool of interested, qualified, and instructor-guided talent that could, if harnessed, significantly accelerate progress in historical data rescue. Student–scientist partnerships such as The Global Learning and Observations to Benefit the Environment (GLOBE) program are collaborations that engage students, teachers, and scientists in authentic research projects delivering context-rich, hands-on approaches to science (Mitchell et al. 2017; Allan et al. 2011). Students gain important insights into processes underpinning research, while scientists gain access to otherwise unobtainable information (Vitone et al. 2016; Harnik and Ross 2003).

The present study describes a novel research-led accredited assignment as part of the geography program at Maynooth University in Ireland, in which final-year undergraduate students successfully transcribed more than 1,300 station years of daily precipitation data and associated metadata for the period 1860–1939. The student-rescued data relate to pre-1940 (post-1940 is already digitized) daily rainfall observations from 43 stations across Ireland. Most records commence in the early 1900s; however, a few extend into the 1800s. The earliest record transcribed commences in 1866, with the shortest in 1927. In this article we explore i) the potential for integrating data-rescue activities into the classroom, ii) the ability of students to produce reliable transcriptions, and iii) the learning outcomes for students. We provide an overview of the assignment, its management, and the student aids implemented. Details of the workflow employed to evaluate student performance and facilitate the creation of a "corrected" data series are provided. In the final section, we discuss the student's perceptions of

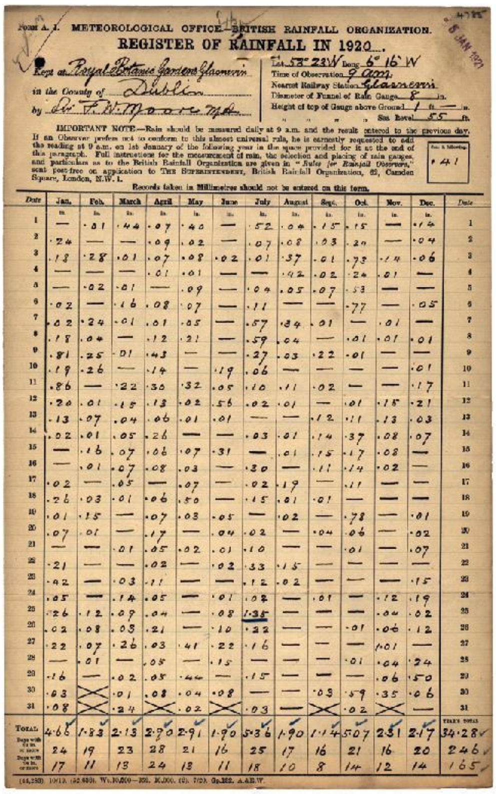
AFFILIATIONS: Ryan—Irish Climate Analysis and Research Units, Department of Geography, Maynooth University, Maynooth, Co. Kildare, and Climatology and Observations Division, Met Éireann, Dublin, Ireland; Duffy, Broderick, Thorne, and Murphy—Irish Climate Analysis and Research Units, Department of Geography, Maynooth University, Maynooth, Co. Kildare, Ireland; Curley, Walsh, and Daly—Climatology and Observations Division, Met Éireann, Dublin, Ireland; Treanor—Library, Met Éireann, and Oireachtas Library, Houses of the Oireachtas, Dublin, Ireland
CORRESPONDING AUTHOR: Ciara Ryan, ciarapryan@gmail.com

The abstract for this article can be found in this issue, following the table of contents.
DOI:10.1175/BAMS-D-17-0147.1
A supplement to this article is available online (10.1175/BAMS-D-17-0147.2).

the project and reflect upon learning outcomes. Our experience demonstrates the substantial potential to extend such approaches to other universities, thus providing an enriched learning experience for the students and a lasting legacy to the scientific community. Project resources, including transcription templates, MATLAB code, and student aids, are provided as online supplemental material.

OVERVIEW OF ASSIGNMENT AND STUDENT AIDS. The assignment was set in the historical climatology component of the final-year climate change course that focuses on the importance of long-term records for understanding past and contemporary climate change. The course takes place in semester 1 (September–December) and is a large class with 142 undergraduate students enrolled during the 2016–17 academic year. The assignment, which accounted for 33% of the course marks, was designed as a research-led teaching experiment to reflect on the importance of historical climatology, to provide insights into the process of data rescue, and to explore the power of the crowd in rescuing and transcribing meteorological data. As a lead-in to the assignment, students were given a guest lecture by staff from Met Éireann to convey the scientific and cultural importance of the data they would be working with. To coincide with the transcription component of the coursework, students were required to submit a written reflection on the importance of historical climatology, utilizing material and publications presented in the module. Learning outcomes were designed to provide students with the following:

- firsthand experience in working with historical climate observations;
- a critical appreciation of the processes involved in data rescue, digitization, and quality-assurance procedures that are essential to understanding past climate variability and change; and
- firsthand experience with the powerful contribution that citizen science can make to the study of climatology, geography, and other disciplines.


FORM A. 1. METEOROLOGICAL OFFICE—BRITISH RAINFALL ORGANIZATION.

REGISTER OF RAINFALL IN 1920.

Kept at Royal Botanic Gardens Glasnevin
in the County of Dublin
by Sir F. W. Moore MA

Lat. 53° 23' N Long. 6° 16' W
Time of Observation 9 am
Nearest Railway Station Glasnevin
Diameter of Funnel of Rain Gauge 8 in.
Height of top of Gauge above Ground 1 ft. — in.
" " " " Sea Level 55 ft.

IMPORTANT NOTE—Rain should be measured daily at 9 a.m. and the result entered to the previous day. If an Observer prefers not to conform to this almost universal rule, he is earnestly requested to add the reading at 9 a.m. on 1st January of the following year in the space provided for it at the end of this paragraph. Full instructions for the measurement of rain, the selection and placing of rain gauges, and particulars as to the British Rainfall Organization are given in "Rules for Rainfall Observers," sent post-free on application to The Superintendent, British Rainfall Organization, 62, Camden Square, London, N.W. 1.

·41

Records taken in Millimetres should not be entered on this form.

Date	Jan.	Feb.	March	April	May	June	July	August	Sept.	Oct.	Nov.	Dec.	Date
	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	
1	—	·81	·44	·07	·40	—	·52	·04	·15	·15	—	·14	1
2	·24	—	—	·09	·02	—	·07	·08	·03	·20	—	·04	2
3	·18	·28	·01	·07	·08	·02	·01	·37	·01	·73	·14	·06	3
4	—	—	—	·01	·01	—	—	·42	·02	·24	·01	—	4
5	—	·02	·01	—	·09	—	·04	·05	·07	·53	—	—	5
6	·02	—	·16	·08	·07	—	·11	—	—	·77	—	·05	6
7	·02	·24	·01	·01	·05	—	·57	·34	·01	—	·01	—	7
8	·18	·04	—	·12	·21	—	·59	·04	—	·01	·01	·01	8
9	·81	·25	·01	·43	—	—	·27	·03	·22	·01	—	—	9
10	·19	·26	—	·14	—	·19	·06	—	—	—	—	·01	10
11	·86	—	·22	·30	·32	·05	·10	·11	·02	—	—	·17	11
12	·20	·01	·15	·13	·02	·56	·02	·01	—	·01	·15	·21	12
13	·13	·07	·04	·06	·01	·01	—	—	·12	·11	·13	·03	13
14	·02	·01	·05	·26	—	—	·03	·01	·14	·37	·08	·07	14
15	—	·16	·07	·06	·07	·31	—	·01	·15	·17	·08	—	15
16	—	·01	·07	·08	·03	—	·30	—	·11	·14	·02	—	16
17	·02	—	·05	—	·07	—	·02	·19	—	·11	—	—	17
18	·26	·03	·01	·06	·50	—	·15	·01	·01	—	—	—	18
19	·01	·15	—	·07	·03	·05	—	·02	—	·78	—	·01	19
20	·07	·01	—	·17	—	·04	·02	—	·04	·06	—	·02	20
21	—	—	·01	·05	·02	·01	·10	—	—	·01	—	·07	21
22	·21	—	—	·02	—	·02	·33	·15	—	—	—	—	22
23	·42	—	·03	·11	—	—	·12	·02	—	—	—	·15	23
24	·05	—	·14	·05	—	·01	·02	—	·01	—	·12	·19	24
25	·26	·12	·09	·04	—	·08	1·35	—	—	—	·04	·02	25
26	·02	·08	·03	·21	—	·10	·22	—	—	·01	·06	·12	26
27	·22	·07	·26	·03	·41	·22	·16	—	—	—	1·01	—	27
28	—	·01	—	·05	—	·15	—	—	—	·01	·04	·24	28
29	·16	—	·02	·05	·44	—	·15	—	—	—	·06	·50	29
30	·03	X	·01	·08	·04	·08	—	—	·03	·59	·35	·06	30
31	·08	X	·24	X	·02	X	·03	—	X	·02	X	—	31
TOTAL	4·66	1·83	2·13	2·70	2·91	1·90	5·36	1·90	1·14	5·07	2·31	2·17	YEAR'S TOTAL 34·28
Days with ·01 in. or more	24	19	23	28	21	16	25	17	16	21	16	20	246
Days with ·04 in. or more	17	11	13	24	13	11	18	10	8	14	12	14	165

Fig. 1. Example of an annual rainfall sheet distributed to students.

Scans of more than 1,300 pre-1940 annual rainfall sheets (e.g., Fig. 1) containing daily precipitation values and associated station metadata, together with templates used in transcribing the data (Fig. 2), were provided by Met Éireann. Of the sheets provided, 274 had been previously transcribed (keyed once) by Met Éireann. These data were used as a benchmark for student performance.

Upon receipt of the digital images, the file format was converted from a Portable Network Graphics (PNG) file type to a Joint Photographic Experts Group (JPEG) format, reducing file size from 17 Mb to ~1.4Mb, without any apparent loss of resolution. This facilitated the distribution of images to students and avoided potential logistical issues concerning students accessing large files. Each student was assigned 18 annual rainfall sheets to transcribe, and each sheet was assigned twice (i.e., to two different students). Such double-key data entry is a widely used method of quality control to detect incorrectly keyed information. For the purpose of this project, double-key entry was necessary for the allocation of student grades, as a component of the grade was based on transcription accuracy, which could only

Metadata		Day	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Month	M_total	M_rainday	M_wetday	QCmonth
image_no	1163	1		0.01	0.44	0.07	0.4		0.52	0.04	0.15	0.15		0.14	1	4.66	24	17	4.66
station	Royal Botanic Ga	2	0.24			0.09	0.02		0.07	0.08	0.03	0.24		0.04	2	1.83	19	11	1.83
county	Dublin	3	0.18	0.28	0.01	0.07	0.08	0.02	0.01	0.37	0.01	0.73	0.14	0.06	3	2.13	23	13	2.13
observer	Sir F.W.Moore	4				0.01	0.01			0.42	0.02	0.24	0.01		4	2.9	28	24	2.9
lat_degree	53	5		0.02	0.01		0.09		0.04	0.05	0.07	0.53			5	2.91	21	13	2.91
lat_min	23	6	0.02		0.16	0.08	0.07		0.11			0.77		0.05	6	1.9	16	11	1.71
lon_degree	6	7	0.02	0.24	0.01	0.01	0.05		0.57	0.34	0.01		0.01		7	5.36	25	18	5.49
lon_min	16	8	0.18	0.04		0.12	0.21		0.59	0.04		0.01	0.01	0.01	8	1.9	17	10	1.96
year	1920	9	0.81	0.25	0.01	0.43			0.27	0.03	0.22	0.01			9	1.14	16	8	1.14
start_month	1	10	0.19	0.26		0.14			0.19	0.06				0.01	10	5.07	21	14	5.07
start_day	1	11	0.86		0.22	0.3	0.32	0.05	0.1	0.11	0.02			0.17	11	2.31	16	12	2.31
funnel_diameter_	8	12	0.2	0.01	0.15	0.13	0.02	0.56	0.02	0.01		0.01	0.15	0.21	12	2.17	20	14	2.17
gauge_height_ft	1	13	0.13	0.07	0.04	0.06	0.01	0.01			0.12	0.11	0.13	0.03					
gauge_height_inc	0	14	0.02	0.01	0.05	0.26			0.03	0.01	0.14	0.37	0.08	0.07					
gauge_pattern		15		0.16	0.07	0.06	0.07	0.31		0.01	0.15	0.17	0.08						
railway	Glasnevin	16	0.02	0.01	0.07	0.08	0.03		0.3		0.11	0.14	0.02						
railway_distance		17	0.26		0.05		0.07		0.02	0.19		0.11							
elevation_msl	55ft	18	0.01	0.03	0.01	0.06	0.5		0.15	0.01	0.01			0.01					
jan_1_following	0.41	19	0.07	0.15		0.07	0.03	0.05		0.02		0.78		0.02					
year_total	34.28	20		0.01		0.17		0.04	0.02		0.04	0.06		0.07					
year_rainday	246	21			0.01	0.05	0.02	0.01	0.1			0.01							
year_wetday	165	22	0.21			0.02		0.02	0.33	0.15				0.15					
time_of_observati	9:00 AM	23	0.42		0.03	0.11			0.12	0.02				0.19					
comments_sheet		24	0.05		0.14	0.05		0.01	0.02		0.01		0.12	0.02					
observer_notes		25	0.26	0.12	0.09	0.04		0.08	1.35				0.04	0.12					
		26	0.02	0.08	0.03	0.21		0.1	0.22			0.01	0.06						
		27	0.22	0.07	0.26	0.03	0.41	0.22	0.16				1.01	0.24					
		28		0.01		0.05		0.15				0.01	0.04	0.5					
		29	0.16		0.02	0.05	0.44		0.15				0.06	0.06					
		30	0.03		0.01	0.08	0.04	0.08			0.03	0.59	0.35						
		31	0.08		0.24		0.02		0.03			0.02							

FIG. 2. Template distributed to students for transcribing the data. Columns D–O contain the 31 × 12 array of values extracted for comparison against the corresponding file. Differences detected by the MATLAB code between the two corresponding files are highlighted in yellow. Columns Q–S contain respective values for monthly totals, total number of rain days per month, and total number of wet days per month as recorded by the observer. Column T (red font) shows the quality check used by the students to compare automatically generated monthly totals to the monthly totals recorded by the observer and input to columns D–O.

be ascertained through comparison. The assignment was set in an hour-long lecture that outlined both the assignment and the rationale. Individual directories containing 18 randomly selected annual rainfall sheets were created and distributed to each student via Dropbox along with a Microsoft Excel template for keying the data. Students were provided a link from which they downloaded their personalized directory and performed the transcription component of the coursework. Students were given five weeks in which to complete and submit their transcribed data.

Several additional student aids were implemented:

- A simple video tutorial was produced to describe the different sections of the annual rainfall sheets and to demonstrate the transcription process, and posted to Moodle, the university's online learning platform.
- An automated quality-assurance check was integrated into the Excel template to generate a monthly total based on the daily values transcribed. Comparison against the monthly totals recorded on the original sheet provided students a basic quality check to ensure transcribed monthly totals were consistent with the originals (Fig. 2).
- An online discussion forum was set up on the Moodle course page through which teaching staff could address queries raised by students. Students were invited to post questions relating to, for example, differences identified in monthly totals and difficulties interpreting metadata, or to clarify illegible values.
- An in-class check-in clinic was organized at the midway point of the assignment to highlight frequently asked questions from the online forum and to allow students the opportunity to raise questions in class.

Once the assignment was complete, students compiled the transcribed files into a compressed zipped directory, maintaining a consistent file-naming convention, and then uploaded the files to the online course portal. Developing a file-naming convention that uniquely identifies the image and student (e.g., image number_student number.xlsx) was essential to data postprocessing. This allowed for a comparison between equivalent sheets based on image number

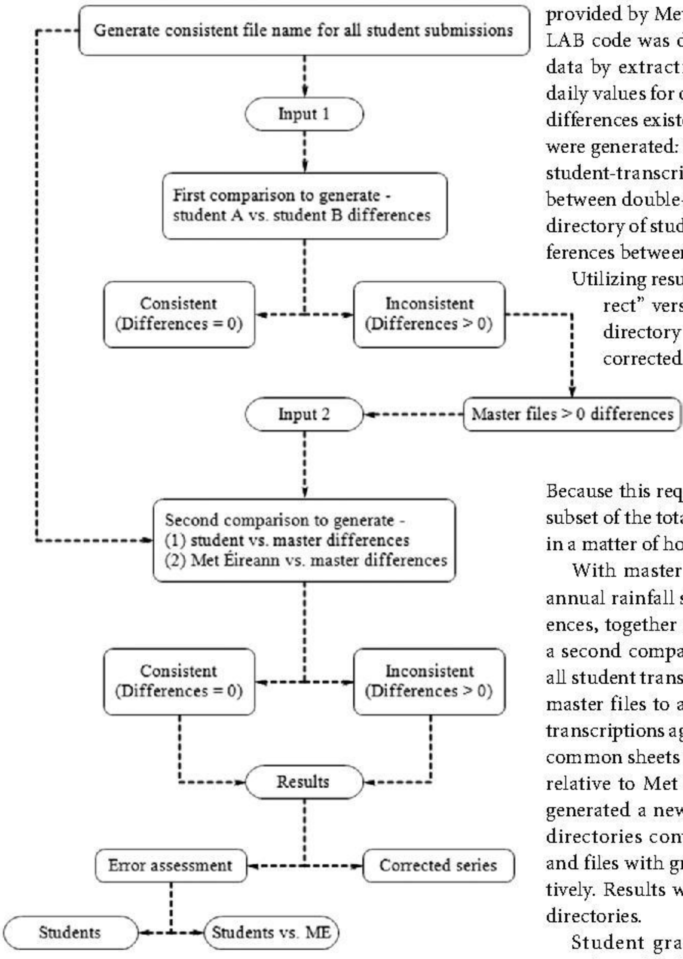


Fig. 3. Flowchart outlining the workflow from receiving student submissions to final output and assessment of errors.

and assignment of grades to students based on an evaluation of their performance.

EVALUATING STUDENT PERFORMANCE. Figure 3 provides an overview of the workflow implemented once student transcriptions were received. Having verified that consistent naming conventions existed for all student files, an initial comparison of the double-keyed data was performed. Student files corresponding to the 274 previously transcribed files provided by Met Éireann were also included. MATLAB code was developed to compare double-keyed data by extracting the 31 × 12 array representing daily values for one year and to highlight cells where differences existed (Fig. 2). From this, two directories were generated: i) a consistent directory comprising student-transcribed files indicating zero differences between double-keyed sheets and ii) an inconsistent directory of student-transcribed files containing differences between double-keyed sheets.

Utilizing results from the first comparison, a "correct" version of each file in the inconsistent directory was created by teaching staff. These corrected or "master" files were manually created by examining highlighted differences against the original scanned images and then adjusting transcribed values accordingly. Because this required consideration of solely a small subset of the total values transcribed, it was achieved in a matter of hours by two members of staff.

With master files now available for all original annual rainfall sheets (i.e., those showing no differences, together with the manually corrected files), a second comparison was performed to evaluate i) all student transcriptions against the corresponding master files to assess student errors and ii) student transcriptions against Met Éireann transcriptions for common sheets to benchmark student performance relative to Met Éireann. This second comparison generated a new set of consistent and inconsistent directories containing files with zero differences and files with greater than zero differences, respectively. Results were derived from these two output directories.

Student grades were assigned based on the number of differences within any transcribed file. A file containing zero differences obtained full marks; for files showing differences, marks were deducted for each incorrect entry. Figure 4a displays the frequency (%) of student submissions being less than or equal to x, where x is the number of errors per transcribed file (bounded by 0 and 365 for all years, except leap years when 366 are possible), together with a bar graph (Fig. 4b) categorizing the total number of incorrect files by actual number of errors per transcribed sheet. When compared against the corrected master files, 62% of student-transcribed files showed no errors. In 96% of student-transcribed files, fewer than 5% of data entries were incorrect (i.e., 96% of files had

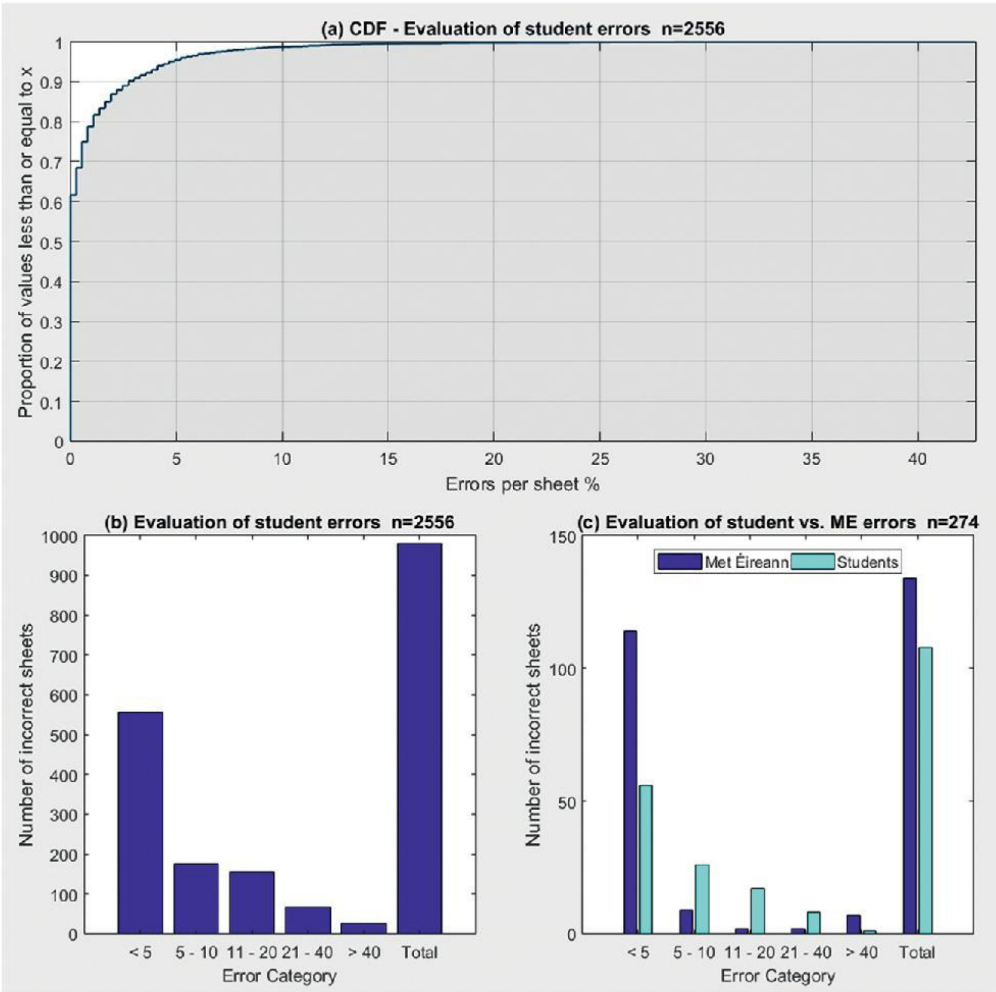


Fig. 4. (a) Empirical cumulative distribution function (CDF) showing the frequency (%) of student submissions (*n* = 2556) being less than or equal to *x*, where *x* is the percentage of errors per sheet. (b) Total number of incorrect student sheets by number of errors per sheet. (c) Comparison of errors by error category for sheets transcribed by both students and Met Éireann (*n* = 274).

fewer than 20 errors). A review of all incorrect files reveals that 57% of the transcribed files containing errors had fewer than 5 errors, 90% had fewer than 20 errors, and only 3% had greater than 40 errors. Cumulative error across all 2,556 files transcribed by students reveals a percentage error of less than 1%. In the unlikely event that both students were wrong (i.e., that they produced identical errors), the errors would not be detected in the comparison. Such errors would, however, be identified during the subsequent application of more comprehensive quality-assurance techniques.

BENCHMARKING STUDENTS AGAINST PROFESSIONALS. A final assessment was carried out utilizing the 274 sheets transcribed by both Met

Éireann and the students. This provided a benchmark against which student performance could be evaluated. Figure 4c displays the proportion of incorrect files by error category. While the students have a smaller number of incorrect files overall (39% for students compared with 49% for Met Éireann), the majority of Met Éireann's incorrect files lie in the lowest error-propensity category. Of the incorrect files from students, 52% contained fewer than 5 errors and 47% contained 6–40 errors, with the remaining 1% of incorrect files showing more than 40 errors. For Met Éireann, 85% of incorrect files contained fewer than 5 errors; however, Met Éireann also had a greater number of incorrect files with more than 40 errors.

Upon further investigation, it was noted that different approaches to the transcription process (i.e., row based vs column based) had an impact on the number of errors produced within individual files. Specifically, the row-based approach employed by Met Éireann resulted in the majority of incorrect files falling into the lowest error category (i.e., ≤5 errors). Alternatively, the column-based approach adopted by the students produced a greater number of files containing errors in the intermediate categories (i.e., 6–10, 11–20, and 21–40). Utilizing a column-based approach, the incorrect placement of a data value occasionally propagated down the entire monthly column. While the individual values were correct, differences were flagged due to the data values being input to the incorrect date/cell. An example of errors arising from the two different approaches is highlighted in Fig. 2.

Table 1. List of statements provided to students to assess learning outcomes and the extent to which proportions of the class (%) agreed or disagreed with these statements (n = 62).

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Total
The assignment provided insights into the process of data rescue.	58	40	0	2	0	100
I now have an appreciation of the role of historical data in climate research.	55	39	5	2	0	100
Supports provided were sufficient to aid completion of the assignment.	66	29	2	3	0	100
The process of working with the data/images was clear and easy to follow.	41	41	10	7	2	100
The assignment provided insights into the power of citizen science.	44	35	16	5	0	100
I would prefer to participate in CA like this over more traditional assignments.	60	20	10	7	3	100
The workload was appropriate to the level of credit given.	45	32	11	8	3	100
Overall, I found the assignment to be a valuable learning experience.	44	44	6	6	0	100
I was more motivated than is usual for me in doing this assignment.	35	23	39	3	0	100
I feel that I have contributed to understanding Irish rainfall trends.	44	47	10	0	0	100
I think that future students would be pleased to participate in a similar assignment.	52	37	5	5	2	100
I think that working with data is a useful skill to have gained.	52	44	3	2	0	100
I did not gain any valuable skills from this assignment.	0	2	11	34	53	100

EVALUATING THE STUDENT EXPERIENCE. On completion of the project, feedback was provided to the students. First, we explained how the evaluation was performed and what the results told us about the accuracy with which the students had transcribed the data. Second, data transcribed by the students for a local station were collated and presented to the class to exemplify their contribution to understanding changes in historical rainfall. We discussed trends and notable events present in the early record and how it commensurately nuanced our understanding of long-term local climate. To investigate the students' perspectives of the project and reflect upon their perceptions of learning outcomes, a formal assessment was conducted via anonymous questionnaire completed in class in the final week. Table 1 presents the list of statements provided to the students and the extent to which proportions of the class who responded agreed or disagreed with each of these statements.

The response was positive across all aspects of the assignment. The majority of students (>90%) gained insights into the process of data rescue and an appreciation of the role of historical data in climate research. The assignment afforded students a firsthand experience in working with raw historical climate data and dataset development—in particular, processes behind the cataloguing and imaging of historical rainfall sheets, the importance of double entry to avoid gross errors, and the value of metadata recorded by the original observers. More than 90% of participants stated that they could see value in their contribution (extending available digital records) to understanding Irish rainfall trends. Presentation of the long-term time series at the end of the project supported this consensus. Notably, 80% of respondents stated that they would prefer to participate in class assignments (CA) like this over other, more traditional, assignments. Students were given the opportunity to continue working with the data, but in smaller groups, in the second semester as part of a research methods workshop whereby they could conduct statistical analysis on the transcribed data and present findings in individual research reports.

Students were asked to identify which student aids they found most useful in completing the assignment. The most popular were the online discussion forum and the video tutorial, at 44% and 36%, respectively. The management of these resources demanded a significant investment of time by teaching staff, particularly the online discussion forum, which received more than 500 queries from students. Nevertheless, effective management of the project and student aids facilitated the development of the corrected dataset by notably reducing the propensity for errors and the amount of time required to carry out postprocessing of the data. While questions have been raised over the accuracy and reliability of citizen science–produced datasets, a number of projects have demonstrated the potential for enhancing data quality through practical management (van der Velde et al. 2017; Kosmala et al. 2016). Despite the success of this initial experiment, a number of issues arose that require consideration in future iterations:

- Given that the misplacement of daily values, rather than incorrect entry of data, was the main source of error, simple additions to the Excel template used to transcribe the data (e.g., delineating the margins of the columns and rows) could significantly reduce the propensity for errors. Additionally, adopting the row-based approach used by Met Éireann to transcribe the data could further reduce errors associated with the incorrect placement of data values.
- Reducing the number of sheets allocated to the students could potentially increase student motivation and further reduce the number of errors.
- Having successfully integrated the transcription component into the coursework, the next iteration of the assignment will build in basic quality assurance and statistical analysis of the resulting dataset. Development of such skills will further deepen the research experience offered to students.

SUMMARY. We have outlined an initial successful attempt to integrate data rescue into the classroom using a research-led approach to teaching and learning that complies with the prerequisites of pedagogy within the university curriculum. Additionally, the project motivated students by engaging them in a practical exercise whereby their contribution adds considerable value to research. Such initiatives promote the development of mutualistic collaborations between national meteorological services and higher-level institutions with the paired objectives of accelerating science. At the same time, students gain a firsthand understanding of the processes that underpin data rescue and research.

Having established confidence in the transcription, the next steps in developing a long-term daily rainfall network for Ireland will involve i) repeated iterations of this assignment with subsequent classes, ii) the application of comprehensive quality-assurance and homogeneity techniques, and iii) analysis of the derived long-term record to assess changes in the characteristics of extreme rainfall events. Metadata transcribed by the students will be systematically extracted and catalogued to facilitate the process. A final objective is to make the data widely available to national and international researchers. To this end, the data shall be shared with the recently awarded Copernicus Climate Change Service Global Land and Marine Observations Database service led by coauthor Peter W. Thorne.

It is hoped that the framework outlined in this paper may be integrated into teaching programs across other universities and used to highlight the importance of historical meteorological sources to students and to encourage their involvement in data-rescue efforts. We will provide the developed student aids to help realize this objective. These materials are available via both online supplemental material and at www.maynoothuniversity.ie/icarus/data-rescue-classroom. The latter shall be updated as we make modifications in subsequent years.

ACKNOWLEDGMENTS. The authors thank and commend Maynooth University (GY313: Climate Change) students who participated in the project. Ciara Ryan gratefully acknowledges funding provided by the Maynooth University John and Pat Hume Doctoral Scholarship and the Irish Research Council Employment Based Scholarship. Conor Murphy acknowledges support from the Irish Environmental Protection Agency project 2014-CCRP-MS.16.

FOR FURTHER READING

Allan, R., P. Brohan, G. P. Compo, R. Stone, J. Luterbacher, and S. Brönnimann, 2011: The International Atmospheric Circulation Reconstructions over the Earth (ACRE) initiative. *Bull. Amer. Meteor. Soc.*, **92,** 1421–1425, https://doi.org/10.1175/2011BAMS3218.1.

Bonney, R., J. L. Shirk, T. B. Phillips, A. Wiggins, H. L. Ballard, A. J. Miller-Rushing, and J. K. Parrish, 2014: Next steps for citizen science. *Science,* **343,** 1436–1437, https://doi.org/10.1126/science.1251554.

Brunet, M., and P. Jones, 2011: Data rescue initiatives: Bringing historical climate data into the 21st century. *Climate Res.*, **47,** 29–40, https://doi.org/10.3354/cr00960.

Harnik, P. G., and R. M. Ross, 2003: Assessing data accuracy when involving students in authentic paleontological research. *J. Geosci. Educ.*, **51,** 76–84, https://doi.org/10.5408/1089-9995-51.1.76.

Kosmala, M., A. Wiggins, A. Swanson, and B. Simmons, 2016: Assessing data quality in citizen science. *Front. Ecol. Environ.*, **14,** 551–560, https://doi.org/10.1002/fee.1436.

Mitchell, N., M. Triska, A. Liberatore, L. Ashcroft, R. Weatherill, and N. Longnecker, 2017: Benefits and challenges of incorporating citizen science into university education. *PLOS ONE*, **12,** e0186285, https://doi.org/10.1371/journal.pone.0186285.

van der Velde, T., D. A. Milton, T. J. Lawson, C. Wilcox, M. Lansdell, G. Davis, G. Perkins, and B. D. Hardesty, 2017: Comparison of marine debris data collected by researchers and citizen scientists: Is citizen science data worth the effort? *Biol. Conserv.*, **208,** 127–138, https://doi.org/10.1016/j.biocon.2016.05.025.

Vitone, T., K. A. Stofer, M. S. Steininger, J. Hulcr, R. Dunn, and A. Lucky, 2016: School of ants goes to college: Integrating citizen science into the general education classroom increases engagement with science. *J. Sci. Commun.*, **15,** 1–24.

Received: 22 January 2020 | Revised: 25 April 2020 | Accepted: 8 July 2020

DOI: 10.1002/gdj3.103

DATA PAPER

Geoscience Data Journal RMetS WILEY

Ireland's pre-1940 daily rainfall records

Ciara Ryan[1,2] | Conor Murphy[1] | Rhonda McGovern[1] | Mary Curley[2] | Séamus Walsh[2] | 476 students

[1]Irish Climate Analysis and Research Units, Department of Geography, Maynooth University, Maynooth, Co., Kildare, Ireland

[2]Climate and Observations Division, Met Éireann, Dublin 9, Ireland

Correspondence
Ciara Ryan, Irish Climate Analysis and Research Units, Department of Geography, Maynooth University, Maynooth, Co. Kildare, Ireland.
Email: ciara.ryan@mu.ie

Funding information
This article was partly funded by a grant from the Irish Research Council (IRC) as part of the IRC Employment Based PhD Programme (EBP/2017/441). CM was funded by the Irish Research Council COALESCE scheme (Grant COALESCE/2019/43). Additional funding was provided by the Maynooth University Research Incentivisation Scheme (RIS) and the Irish Climate Analysis and Research UnitS (ICARUS) for the purchase of the Metis digital scanner.

Abstract

This article presents daily rainfall data and metadata for Ireland transcribed from historical manuscript and printed copies of rainfall registers located in Met Éireann's archives. To facilitate the transcription of rainfall observations from paper records, the historical manuscripts were scanned and integrated into Met Éireann's digital archives. The transcription from digital image to data format was undertaken in collaboration with students at Maynooth University as part of a novel crowdsourcing initiative to integrate data rescue activities into the classroom. In total, 3,616 station years of rainfall data (~1.32 million daily values) were transcribed. The data, which was double keyed, have undergone basic quality assurance to check for transcription errors and the resultant raw data and associated metadata are presented here. Ongoing work involves the application of further quality assurance and homogenization techniques to develop a long-term, quality assured daily rainfall network for Ireland.

KEYWORDS

data rescue, historical data, Ireland, precipitation

Dataset

Identifier: http://hdl.handle.net/2262/91347

Creator: C. Ryan, C. Murphy, R. McGovern, M. Curley, S. Walsh and 476 students

Title: Ireland's pre-1940 daily rainfall records

Publisher: edeposit Ireland

Publication year: 2020

Resource type: Dataset

Version: 1.0

1 | INTRODUCTION

There is an increasing demand for climatological information to support scientific understanding of the impacts of climate change. Accurate and reliable long-term meteorological records are essential for understanding climate trends and variability, evaluating reanalysis products and climate models and climate risk management (e.g. Matthews *et al*., 2014; Wilby *et al*., 2016; Noone *et al*., 2017; Murphy *et al*., 2019). Consequently, much work is being carried out internationally to rescue historical climate data and develop long-term climate series (e.g. Cornes *et al*., 2012; Allan *et al*., 2016; Ashcroft *et al*., 2018; Hawkins *et al*., 2019; Coll *et al*., 2019).

There is a long history of meteorological observation in Ireland. The earliest known weather observations were recorded at Kilkenny for the Duke of Ormond by John Kevan in 1682. From 1684 to 1686, William Molyneux and George Ashe, under the auspices of the Dublin Philosophical Society, documented weather at Trinity College, Dublin. Records resumed here in 1708 to 1709 when Samuel Molyneux documented observations of temperature, rainfall and pressure. In 1787, Richard Kirwan established the first Irish meteorological station at Cavendish Row, Dublin. These records are the earliest surviving observations compiled with the aid of accurate instruments. Kirwan maintained the observations until 1808, by which time regular meteorological observations had commenced at the National Botanic Gardens (Dixon, 1987). Observations from Armagh Observatory date from 1796 to present, providing the longest continuous series from any single site on the island (Butler *et al*., 2007). In 1829, the Ordnance Survey Office located in Phoenix Park, Dublin began recording systematic meteorological measurements providing continuous readings up to the present day.

In 1859, George James Symons, working with a network of voluntary observers, set up a system for gathering and publishing rainfall records from across Britain and Ireland. The data from Symons' observational network were published in British Rainfall (formerly English Rainfall), which for the year 1860 reports that there were then 168 stations in the network. No observations were included from sites in Scotland or Ireland, and only three from Wales (Glasspoole, 1952). The earliest available observations taken at Irish sites are from 1864. In 1900, the network founded by Symons became known as the British Rainfall Organisation and was later (in 1919) formally transferred to the aegis of the British Meteorological Office. At this time, there were approximately 5,000 observers contributing to the rainfall network (Glasspoole, 1952).

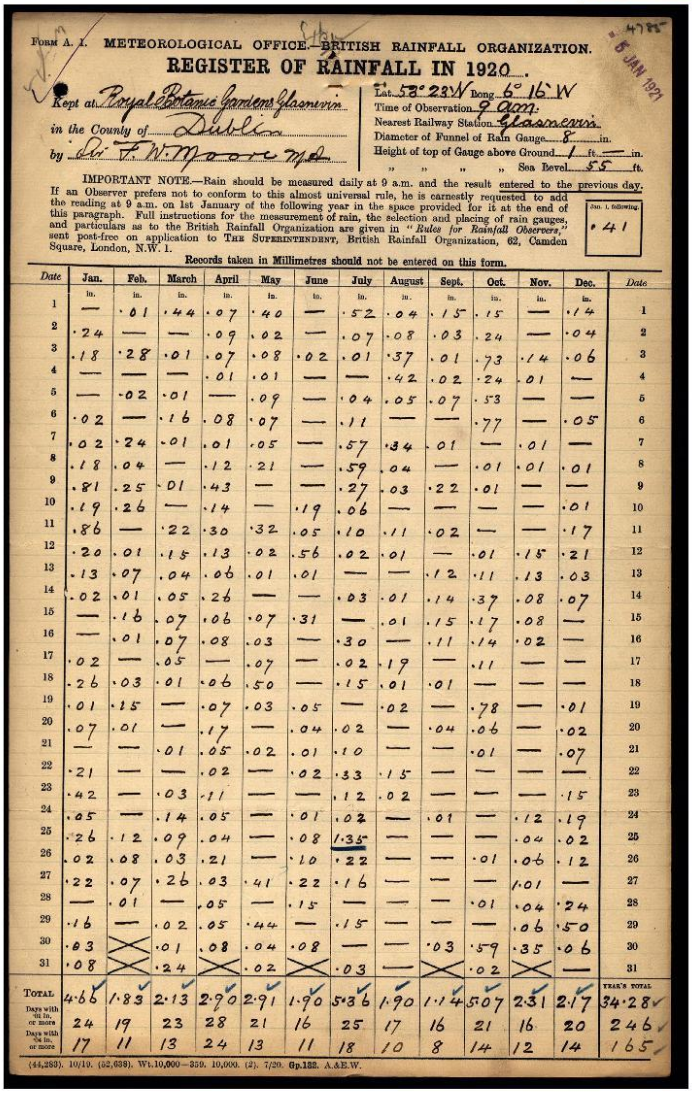
At the end of the 19th century, under the direction of the British Meteorological Department, led by Admiral Robert Fitzroy of the Royal Navy, an operational network of observing stations was established. In 1868, the Met Office equipped seven observatories, including two in Ireland, with autographic instruments, providing the first continuous record of a variety of meteorological parameters at a selection of locations (Eden, 2009). The records, along with the Phoenix Park records, which were maintained by the Royal Engineers, represent the early efforts of the British Meteorological Office to establish a network of meteorological stations across Britain and Ireland.

The Irish Meteorological Service, later Met Éireann, was established in December 1936 and subsequently took over responsibility for the network of Irish stations from the British Meteorological Office. The registers that pre-date the establishment of the Irish Meteorological Service were later transferred from the British Meteorological Office to the current service in Ireland. The substantial paper records are carefully preserved in Met Éireann's archives, but until now have largely remained in paper format for years prior to 1941.

Met Éireann maintains the National Climate Database. The database comprises observations received by Met Éireann from the current network of staffed, synoptic, climatological and rainfall stations. Observations are quality controlled and archived in Met Éireann's database. Daily observations of rainfall have been digitized back to 1941 and temperature to 1961 (Walsh, 2013). Work to digitize and transcribe long-term daily minimum and maximum air temperatures is currently underway at the National University of Ireland (NUI), Galway (Mateus *et al*., 2020). Other long-term series include work by Murphy *et al*. (2018) to compile a 305-year continuous monthly rainfall series for the island of Ireland (1711–2016) utilizing data from the UK and Ireland, including previously unpublished work from the 1970s by the British Meteorological Office. The comprehensive monthly series builds on the work of Noone *et al*. (2015) which developed a monthly rainfall series for 25 stations throughout Ireland for the period 1850–2015. However, until now there has not been a concerted effort to construct a long-term daily rainfall series for Ireland using historical records.

This paper describes recent work, undertaken as part of the PhD research of the lead author, to digitize and transcribe historical daily rainfall records from Met Éireann's extensive archive collection. The main objective of this work was to create a digital archive of the paper records of Ireland's longest meteorological stations, and from these imaged records and additional rainfall registers to extend the availability of long-term daily rainfall data prior to 1941. The data presented here are the raw data and associated metadata. It is envisaged that by presenting the data in its original state it can be easily integrated into current international data rescue initiatives, for example Copernicus Climate Change Service Global Land and Marine Observations Database, and that future research will have recourse to the raw data.

The remainder of the paper is organized as follows: Section 2 provides a description of the data. First, we discuss the digitization of historical records from Met Éireann's archives;



FORM A. 1. METEOROLOGICAL OFFICE.—BRITISH RAINFALL ORGANIZATION.

REGISTER OF RAINFALL IN 1920.

5 JAN 1921

4785

Kept at Royal Botanic Gardens Glasnevin

in the County of Dublin

by Sir F. W. Moore M.A.

Lat. 53° 23' N Long. 6° 16' W

Time of Observation 9 am.

Nearest Railway Station Glasnevin

Diameter of Funnel of Rain Gauge 8 in.

Height of top of Gauge above Ground 1 ft. — in.

" " " " Sea Level 55 ft.

IMPORTANT NOTE.—Rain should be measured daily at 9 a.m. and the result entered to the previous day. If an Observer prefers not to conform to this almost universal rule, he is earnestly requested to add the reading at 9 a.m. on 1st January of the following year in the space provided for it at the end of this paragraph. Full instructions for the measurement of rain, the selection and placing of rain gauges, and particulars as to the British Rainfall Organization are given in "*Rules for Rainfall Observers*," sent post-free on application to THE SUPERINTENDENT, British Rainfall Organization, 62, Camden Square, London, N.W. 1.

Jan. 1, following: ·41

Records taken in Millimetres should not be entered on this form.

Date	Jan.	Feb.	March	April	May	June	July	August	Sept.	Oct.	Nov.	Dec.	Date
	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	in.	
1	—	·01	·44	·07	·40	—	·52	·04	·15	·15	—	·14	1
2	·24	—	—	·09	·02	—	·07	·08	·03	·24	—	·04	2
3	·18	·28	·01	·07	·08	·02	·01	·37	·01	·73	·14	·06	3
4	—	—	—	·01	·01	—	—	·42	·02	·24	·01	—	4
5	—	·02	·01	—	·09	—	·04	·05	·07	·53	—	—	5
6	·02	—	·16	·08	·07	—	·11	—	—	·77	—	·05	6
7	·02	·24	·01	·01	·05	—	·57	·34	·01	—	·01	—	7
8	·18	·04	—	·12	·21	—	·59	·04	—	·01	·01	·01	8
9	·81	·25	·01	·43	—	—	·27	·03	·22	·01	—	—	9
10	·19	·26	—	·14	—	·19	·06	—	—	—	—	·01	10
11	·86	—	·22	·30	·32	·05	·10	·11	·02	—	—	·17	11
12	·20	·01	·15	·13	·02	·56	·02	·01	—	·01	·15	·21	12
13	·13	·07	·04	·06	·01	·01	—	—	·12	·11	·13	·03	13
14	·02	·01	·05	·26	—	—	·03	·01	·14	·37	·08	·07	14
15	—	·16	·07	·06	·07	·31	—	·01	·15	·17	·08	—	15
16	—	·01	·07	·08	·03	—	·30	—	·11	·14	·02	—	16
17	·02	—	·05	—	·07	—	·02	·19	—	·11	—	—	17
18	·26	·03	·01	·06	·50	—	·15	·01	·01	—	—	—	18
19	·01	·15	—	·07	·03	·05	—	·02	—	·78	—	·01	19
20	·07	·01	—	·17	—	·04	·02	—	·04	·06	—	·02	20
21	—	—	·01	·05	·02	·01	·10	—	—	·01	—	·07	21
22	·21	—	—	·02	—	·02	·33	·15	—	—	—	—	22
23	·42	—	·03	·11	—	—	·12	·02	—	—	—	·15	23
24	·05	—	·14	·05	—	·01	·02	—	·01	—	·12	·19	24
25	·26	·12	·09	·04	—	·08	1·35	—	—	—	·04	·02	25
26	·02	·08	·03	·21	—	·10	·22	—	—	·01	·06	·12	26
27	·22	·07	·26	·03	·41	·22	·16	—	—	—	1·01	—	27
28	—	·01	—	·05	—	·15	—	—	—	·01	·04	·24	28
29	·16	—	·02	·05	·44	—	·15	—	—	—	·06	·50	29
30	·03	X	·01	·08	·04	·08	—	—	·03	·59	·35	·06	30
31	·08	X	·24	X	·02	X	·03	—	X	·02	X	—	31
TOTAL	4·66	1·83	2·13	2·90	2·91	1·90	5·36	1·90	1·14	5·07	2·31	2·17	YEAR'S TOTAL 34·28
Days with ·01 in. or more	24	19	23	28	21	16	25	17	16	21	16	20	246
Days with ·04 in. or more	17	11	13	24	13	11	18	10	8	14	12	14	165

(44,283). 10/19. (52,638). Wt.10,000—359. 10,000. (2). 7/20. Gp.132. A.&E.W.

FIGURE 1 Sample rainfall register. The annual forms report daily rainfall observations and station metadata for various locations throughout Ireland

then, we outline the data transcription process and provide details of the stations transcribed, followed by a description of the metadata that was collected during the transcription process. In Section 3, we describe the process of checking for errors in the transcribed data. Section 4, provides details of dataset use, while Section 5 concludes with a proposed pathway to complete the imaging and transcription of stations and meteorological data that remain on paper.

2 | DESCRIPTION OF THE DATA

2.1 | Data sources

The daily rainfall data were largely taken from annual Rainfall Registers and Meteorological Registers held in Met Éireann's archives. These registers are described in detail below. In addition, digital images of station records for Kells

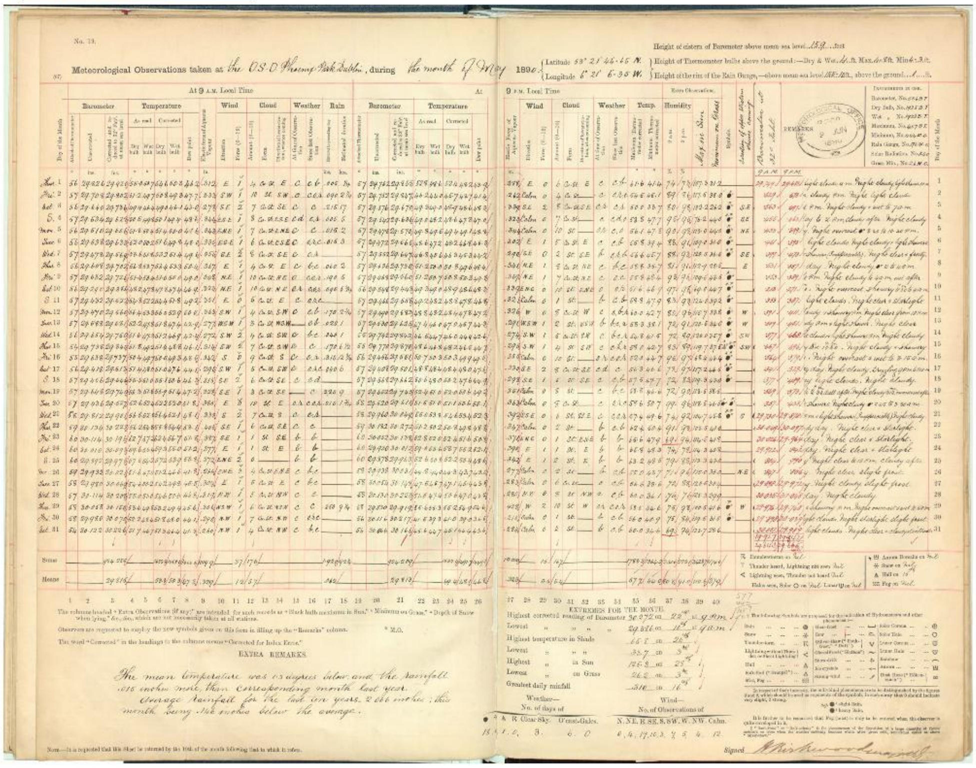


FIGURE 2 Sample meteorological register. The monthly forms report daily observations for multiple parameters including temperature, pressure, precipitation, sunshine, wind and other parameters

(Headfort) and the National University of Ireland (NUI), Galway were obtained from the National Library of Ireland and NUIG, respectively.

2.2 | Data imaging

2.2.1 | Rainfall registers

Previous work by Met Éireann focused on imaging the single-sheet annual rainfall registers collected from the early rainfall network, specifically the original handwritten rainfall observations that were taken by the volunteers who worked as part of Symons' rainfall network in Ireland (Treanor *et al.*, 2011). In 2014, the Met Éireann Library received funding to collate, catalogue and preserve the registers in this collection. The records are known as the 'Rainfall Registers' and contain daily rainfall data and station metadata from various locations throughout Ireland. Readings were taken once a day and recorded on a standard form issued by Symons and later the British Rainfall Organisation (Figure 1). The collection includes readings from every county in the Republic of Ireland and intermittent observations taken in Antrim, Armagh, Derry, Down, Fermanagh and Tyrone in Northern Ireland. There are over 700 stations in total, which return data for varying time periods between 1864 and 1940.

2.2.2 | Meteorological registers

Non-digitized records include a collection of meteorological registers consisting of 224 bound volumes, 128 folders and 353 folios of station records covering the period 1855–1976. The collection includes observations from Ireland's longest operating stations, many of which are still operational. The registers have been arranged by date and region and catalogued by their location in the archives at Met Éireann (Treanor *et al.*, 2011). They provide a daily record of weather parameters at various locations throughout Ireland covering the mid-19th century until the latter part of the 20th century.

Observers recorded data from a selection of instruments that measured temperature, pressure, precipitation, sunshine and other fields. Readings were taken at least once a day and recorded on a British Meteorological Office issued broadsheet form that was designed to contain a month's data (Figure 2). At the end of the month, the forms were posted back to the British Meteorological Office for centralized calculations and corrections.

A Metis EDS professional digital scanner (Figure 3) was purchased to progress the digitization of the historical meteorological registers. To date, meteorological observations from eight long-term stations have been scanned and integrated into the digital database (Table 1). Work is ongoing to digitize the remaining manuscripts. Individual pages were scanned as high quality tif files and categorized using the standard naming convention adopted by Met Éireann. The scanned images were then uploaded to Met Éireann's database where they will be made available for research purposes.

2.3 | Data transcription

Daily rainfall observed at 114 sites throughout Ireland was transcribed as part of this work (Figure 4). Stations were selected based on record length, continuity and spatial distribution. Additional stations that could potentially be used to infill gaps in long-term data series were also included. The transcription from paper and digital image format to digital numerical format was largely undertaken by final year Geography students at Maynooth University as part of a novel crowdsourcing initiative to integrate data rescue activities into the classroom. Ryan *et al*. (2018) presented an innovative approach to data rescue by developing a research-led project to engage students in data rescue tasks for credit. The study explored (a) the potential for integrating data rescue activities into the classroom, (b) the ability of students to produce reliable transcriptions and (c) the achieved learning outcomes for students. The work was facilitated by the provision of student aids including written guidelines, transcription templates with an automated quality-assurance check, a video tutorial, in-class workshops and an online discussion forum. An evaluation of learning outcomes and student's perceptions of the project demonstrated a positive educational experience. Following the success of the initial

FIGURE 3 Metis EDS Gamma professional digital scanner used to image historical meteorological registers held in Met Éireann

TABLE 1 Meteorological registers imaged and integrated into Met Éireann's digital archives

Station	Start year	End year
Birr	1873	1951
Blacksod	1884	1956
Fitzwilliam Square	1869	1935
Malin Head	1885	1955
Markree	1869	1968
Phoenix Park	1866	1959
Roches Point	1873	1956
Valentia	1873	1950

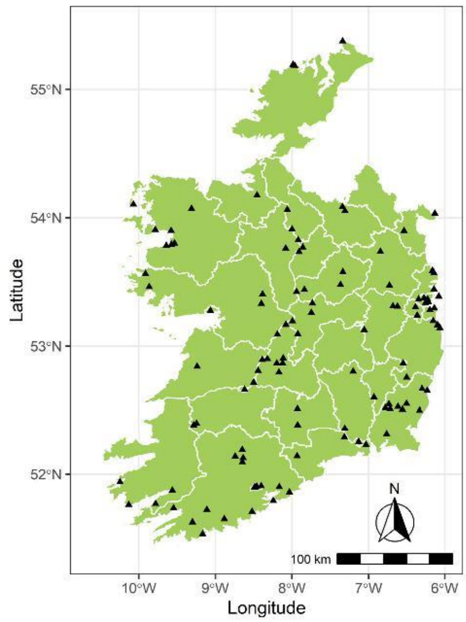


FIGURE 4 Location of stations for which daily rainfall data was transcribed as part of this work. Individual folders containing a data file and metadata file are available for each station

TABLE 2 Station details for data transcribed as part of this work

Station	Lat	Lon	County	Start year	Start month	Start day	End year	End month	End day	% NA	Units
Abbeyfeale (Presbytery)	522,300	91,700	Limerick	1936	1	1	1940	12	31	5.6	Inch
Abbeyfeale (Springmount)	522,350	91,440	Limerick	1925	1	1	1940	12	31	3.2	Inch
Ahascragh (Clonbrock)	532,410	82,300	Galway	1876	1	1	1940	12	31	14.1	Inch
Ardee (Lisrenny)	535,350	63,220	Louth	1886	1	1	1913	7	31	0.0	Inch
Ardgillan	533,509	60,942	Dublin	1894	1	1	1938	2	28	2.6	Inch
Athlone (Twyford)	532,620	75,020	Westmeath	1872	1	1	1940	12	31	14.6	Inch
Athlone O.P.W.	532,520	75,630	Westmeath	1902	1	1	1940	12	31	2.7	Inch
Ballinacurra	515,400	81,000	Cork	1904	4	1	1959	11	30	0.0	IN/MM
Ballycumber (Belair)	532,000	74,400	Offaly	1923	1	1	1940	12	31	8.5	Inch
Ballyhaise (Agr.Coll.)	540,305	71,835	Cavan	1923	1	1	1943	12	31	31.0	Inch
Ballynahinch castle	532,740	95,150	Galway	1913	1	1	1940	12	31	0.1	Inch
Banagher (Canal HSE.)	531,140	75,940	Offaly	1893	1	1	1940	12	31	46.1	Inch
Belgard castle	531,810	62,320	Dublin	1914	4	1	1940	12	31	8.2	Inch
Birr castle	530,540	75,530	Offaly	1875	1	1	1951	12	31	2.0	IN/MM
Blacksod point S.W.S.	540,610	100,410	Mayo	1884	10	1	1956	10	31	1.6	IN/MM
Borris	523,610	65,550	Carlow	1,890	1	1	1895	12	31	19.6	Inch
Bray (Fassaroe)	531,146	60,900	Wicklow	1871	1	1	1916	12	31	0.0	Inch
Caheragh	513,732	91,755	Cork	1910	9	1	1935	12	31	0.3	Inch
Cahir abbey	522,300	75,540	Tipperary	1924	1	1	1940	12	31	11.8	Inch
Cashel (Ballinamona)	523,038	75,543	Tipperary	1911	1	1	1940	12	31	0.0	Inch
Castlecomer HSE	524,810	71,210	Kilkenny	1908	1	1	1940	12	31	6.4	Inch
Castleconnell rectory	524,300	83,000	Limerick	1914	1	1	1939	12	31	0.9	Inch
Castleforbes	534,610	75,130	Longford	1911	1	1	1940	12	31	3.4	Inch
Castletownshend (Glenbarrahane)	513,200	91,000	Cork	1892	1	1	1940	12	31	5.1	Inch
Castletownshend (Seafield)	513,200	91,000	Cork	1914	1	1	1940	12	31	0.0	Inch
Clonakilty (Shannon vale)	513,900	85,300	Cork	1910	1	1	1931	12	31	14.6	Inch
Clongowes wood coll	531,840	64,100	Kildare	1906	1	1	1940	12	31	32.6	IN/MM
Cloondra O.P.W.	534,400	75,430	Longford	1902	1	1	1940	12	31	5.1	Inch
Cloverhill	540,500	72,030	Cavan	1896	1	1	1933	7	31	0.2	Inch
Cork (Park view)	515,412	82,738	Cork	1910	1	1	1937	12	31	0.4	Inch
Cork (The Palace)	515,340	82,850	Cork	1895	1	1	1911	12	31	0.0	Inch

(Continues)

TABLE 2 (Continued)

Station	Lat	Lon	County	Start year	Start month	Start day	End year	End month	End day	% NA	Units
Derrynane Abbey	514,540	100,750	Kerry	1871	1	1	1940	12	31	23.3	Inch
Dromod (Ruskey)	534,930	75,510	Leitrim	1902	1	1	1940	12	31	2.6	Inch
Drumshanbo (Lough allen sluices)	540,340	80,340	Leitrim	1902	1	1	1940	12	31	2.6	Inch
Drumsna (Albert lock)	535,440	80,000	Roscommon	1903	1	1	1940	12	31	2.6	Inch
Dublin (Clontarf)	532,150	61,330	Dublin	1912	1	1	1940	12	31	17.2	Inch
Dublin (Fitzwilliam square)	532,008	61,512	Dublin	1908	1	1	1937	9	30	0.0	Inch
Dublin (Phoenix park)	532,150	62,050	Dublin	1881	1	1	1959	12	31	0.0	IN/MM
Dublin (Ringsend)	532,030	61,250	Dublin	1912	1	1	1940	12	31	13.9	Inch
Dun laoghaire (Harbour yard)	531,740	60,810	Dublin	1905	1	1	1940	12	31	0.2	Inch
Dunfanaghy (St.Patrick's)	551,100	75,800	Donegal	1,880	1	1	1940	12	31	46.7	Inch
Dunmanway	514,320	90,637	Cork	1905	1	1	1937	6	30	3.1	Inch
Enfield (Summerhill)	532,820	64,340	Meath	1896	1	1	1940	12	31	2.2	Inch
Enniscoe	540,410	91,842	Mayo	1874	1	1	1923	12	31	4.1	Inch
Enniscorthy (Ballyhighland)	523,040	64,312	Wexford	1871	1	1	1919	10	31	0.0	Inch
Enniscorthy (Enniscorthy rectory)	523,024	63,335	Wexford	1926	1	1	1929	12	31	0.0	Inch
Enniscorthy (Monart rectory)	523,144	63,706	Wexford	1903	1	1	1907	12	31	2.2	Inch
Enniscorthy (Monksgrange)	523,104	64,658	Wexford	1925	1	1	1931	12	31	0.3	Inch
Enniscorthy (Summerville)	523,310	63,018	Wexford	1908	1	1	1923	12	31	0.0	Inch
Enniscorthy (Woodbrook)	523,255	64,410	Wexford	1902	1	1	1940	12	31	54.2	Inch
Foulkesmill (Longraigue)	521,838	64,559	Wexford	1874	1	1	1940	12	31	11.2	Inch
Glenasmole D.C.W.W.	531,420	62,200	Dublin	1900	1	1	1940	12	31	13.2	Inch
Glengarriff (Ilnacullin)	514,405	93,245	Cork	1914	1	1	1940	12	31	7.4	Inch
Gorey (Courtown house)	523,910	61,345	Wexford	1908	1	1	1941	12	31	0.0	Inch
Gorey (Ram's gate)	524,010	61,800	Wexford	1928	1	1	1940	12	31	2.4	Inch
Gorey (Wells)	523,000	62,000	Wexford	1893	1	1	1923	12	31	6.6	Inch
Greenore	540,150	60,750	Louth	1876	1	1	1940	12	31	6.3	Inch
Greystones (Burnaby lodge)	530,831	60,356	Wicklow	1928	1	1	1940	12	31	0.1	Inch
Greystones (Rathdown HSE.)	530,949	60,540	Wicklow	1918	1	1	1929	7	31	8.7	Inch
Hacketstown rectory	525,200	63,300	Carlow	1918	1	1	1940	12	31	0.0	Inch
Horn head	551,123	75,916	Donegal	1892	1	1	1917	12	31	15.4	Inch
Howth castle	532,310	60,440	Dublin	1913	1	1	1940	12	31	0.0	Inch

(Continues)

TABLE 2 (Continued)

Station	Lat	Lon	County	Start year	Start month	Start day	End year	End month	End day	% NA	Units
Inagh (MT. Callan)	525,030	91,418	Clare	1908	1	1	1940	12	31	6.1	Inch
Kells (Headfort)	534,410	65,050	Meath	1893	10	1	1952	6	30	0.4	Inch
Kenmare (Derreen)	514,610	94,650	Kerry	1912	1	1	1939	12	31	49.8	Inch
Kenmare (Sheen falls)	515,220	93,340	Kerry	1921	1	1	1940	12	31	10.0	Inch
Kilconnel (Rectory)	531,952	82,404	Galway	1875	1	1	1886	12	31	0.1	Inch
Killaloe (Ballina)	524,830	82,630	Clare	1866	1	1	1940	12	31	22.4	Inch
Kinsale (Scilly HSE.)	514,225	83,120	Cork	1922	1	1	1940	12	31	0.0	Inch
Kylemore castle	533,350	95,439	Galway	1889	1	1	1904	12	31	6.3	Inch
Limerick (Mulgrave St.)	523,940	83,710	Limerick	1924	1	1	1940	12	31	0.0	Inch
Lismore castle	520,824	75,600	Waterford	1909	1	1	1940	12	31	41.6	Inch
Lota lodge NO.1	515,422	82,407	Cork	1909	1	1	1937	3	31	3.5	Inch
Lota lodge NO.2	515,422	82,407	Cork	1911	1	1	1925	12	31	0.0	Inch
Malahide (Seamount)	532,630	60,830	Dublin	1908	1	1	1939	3	31	3.2	Inch
Malin head	552,220	72,020	Donegal	1885	1	1	1955	12	31	0.1	IN/MM
Mallaranny	535,420	94,700	Mayo	1919	1	1	1940	12	31	0.4	IN/MM
Mallow (Hazelwood)	521,125	83,900	Cork	1928	1	1	1943	12	31	13.1	Inch
Mallow (Longueville)	520,814	84,426	Cork	1896	1	1	1937	12	31	0.0	Inch
Mallow (Old dramore)	520,537	83,852	Cork	1,890	1	1	1895	12	31	4.7	Inch
Mallow (Summerhill)	520,732	83,817	Cork	1896	1	1	1921	6	30	0.3	Inch
Markree castle	541,030	82,720	Sligo	1874	11	1	1940	12	31	0.0	Inch
Meelick (Victoria lock)	531,000	80,450	Offaly	1902	1	1	1940	12	31	2.6	Inch
Monasterevin (Moore abbey gardens)	530,735	70,333	Kildare	1898	1	1	1918	12	31	0.8	Inch
Mullingar (Ballynegal)	533,440	72,010	Westmeath	1908	1	1	1940	12	31	12.4	Inch
Mullingar (Belvedere HSE)	532,840	72,200	Westmeath	1898	1	1	1940	12	31	9.3	Inch
Nenagh (Ballygibbon)	525,422	80,653	Tipperary	1901	1	1	1912	12	31	0.0	Inch
Nenagh (Castle lough)	525,340	82,320	Tipperary	1875	2	1	1940	12	31	19.3	Inch
Nenagh (Clashnevin)	525,200	80,720	Tipperary	1884	1	1	1899	9	13	0.0	Inch
Nenagh (Nenagh Lodge)	525,403	81,918	Tipperary	1879	1	1	1889	12	31	4.9	Inch
Nenagh (St.Mary's rectory)	525,200	81,156	Tipperary	1919	1	1	1930	10	31	0.0	Inch
Nenagh (Traverston)	524,746	81,013	Tipperary	1907	1	1	1937	12	31	0.1	Inch

(Continues)

TABLE 2 (Continued)

Station	Lat	Lon	County	Start year	Start month	Start day	End year	End month	End day	% NA	Units
Newport (Burrishoole HSE.)	535,352	93,442	Mayo	1899	1	1	1919	12	31	0.1	Inch
Nuig	531,635	90,350	Galway	1864	1	1	1952	6	30	12.6	Inch
Piltown (Kildalton Abbey)	522,120	71,840	Kilkenny	1877	1	1	1940	12	31	45.4	Inch
Portlaw (Mayfield)	521,710	71,900	Waterford	1881	1	1	1940	12	31	3.9	Inch
Portumna O.P.W.	530,530	81,130	Galway	1929	1	1	1940	12	31	16.7	Inch
Queens College Cork	515,334	82,925	Cork	1866	1	1	1907	12	31	40.7	Inch
Roches point	514,735	81,440	Cork	1873	7	1	1940	12	31	0.0	IN/MM
Royal Botanic Gardens (Glasnevin)	532,221	61,618	Dublin	1,880	1	1	1925	12	31	2.2	Inch
Shanagarry (Kinoith)	515,130	80,200	Cork	1910	1	1	1940	12	31	3.0	Inch
Shillelagh (Coollattin)	524,520	63,010	Wicklow	1908	1	1	1940	12	31	3.6	Inch
Skerries (Milverton hall)	533,410	60,830	Dublin	1908	1	1	1940	12	31	9.1	Inch
Stillorgan (Farmleigh)	531,655	61,138	Dublin	1915	1	1	1935	12	31	4.8	Inch
Straffan HSE	531,830	63,730	Kildare	1883	1	1	1940	12	31	0.1	Inch
Strokestown (Castlenode)	534,540	80,500	Roscommon	1908	1	1	1940	12	31	0.0	Inch
Turraun (Bord Na Mona)	531,540	74,450	Offaly	1913	1	1	1940	12	31	10.8	Inch
UCC (Univ.Coll.Cork)	515,359	82,910	Cork	1910	1	1	1940	12	31	3.3	Inch
Valentia (Observatory)	515,617	101,436	Kerry	1875	1	1	1950	12	31	0.0	IN/MM
Waterford (Brook lodge)	521,345	70,205	Waterford	1876	1	1	1931	12	31	20.6	Inch
Waterford (Tycor)	521,510	70,750	Waterford	1888	1	1	1939	12	31	15.5	Inch
Westport (Murrisk Abbey)	534,655	93,823	Mayo	1900	1	1	1904	12	31	0.0	Inch
Westport (St. Helens)	534,726	93,420	Mayo	1903	1	1	1920	12	31	11.5	Inch
Westport HSE	534,800	93,200	Mayo	1911	1	1	1940	12	31	3.1	Inch

Note: Individual folders containing data files and metadata files are provided for each station listed. Station coordinates are given in degree, minutes, seconds (DMS).

project, a further two iterations have been executed across three cohorts of final year Geography students at Maynooth University, producing in excess of 3,500 station years of historical daily rainfall data. A detailed description of the methodology and access to resources is provided by Ryan *et al.* (2018) (https://doi.org/10.1175/BAMS-D-17-0147.1).

The majority of the data was transcribed from the annual rainfall registers. Additional rainfall data were extracted from the long-term meteorological registers listed in Table 1. Rainfall register observations were, for the most part, recorded in inches of rainfall. Data extracted from the meteorological registers were recorded in millimetres from the start of record up to 1914/15 when the unit of measurement changed to inches. Table 2 provides a quick reference to the different units of measurement used at each station. Detailed changes of units of measurement are provided in the individual station metadata files. The original units of measurement have been preserved here. The number of stations varies significantly throughout the years, with only a small number of stations available in the early record (Figure 5). The earliest observations recovered were taken at the National University of Galway (NUIG) in 1864. From 1900, the number of stations increases significantly. In total, 3,616 station years (~1.32 million daily values) were transcribed. Details of all stations transcribed are provided in Table 2.

2.4 | Metadata

In addition to the rainfall observations, metadata for each station were extracted from the individual station records and recorded in separate text files (Figure 6). These files provide information on station name and location, observer, record length, missing data, diameter of gauge, changes in gauge height, gauge pattern, time of observation, unit of measurement and distance to nearest railway station. A section for 'Additional Information' provides details of observations made while transcribing the data, for instance, the presence of a multiday accumulation. A separate section for 'Notes' provides a transcription of any handwritten notes recorded on the original record by the observer. For example, an observer may note an exceptional rainfall event or a leak in the gauge. These notes are transcribed verbatim.

The British Meteorological Office coordinated the meteorological network by supplying instruments and instructions for observations. Nevertheless, before standard equipment and procedures for meteorological observations were introduced from the late-nineteenth century to mid-twentieth century, the design and placement of rain-gauges varied considerably. Observer bias, instrument changes, sampling periods, as well as external factors relating to site exposure (e.g. proximity to buildings) effect the accuracy and consistency of observations (Kunkel *et al.*, 2005; Daly *et al.*, 2007; Green *et al.*, 2008). In general, such changes were recorded on the station record and subsequently transcribed to the metadata files. Detailed metadata presented here provide the comprehensive account of station changes and can be used as an aid in determining the reliability of station records. This information will be particularly valuable in helping to explain the presence of any abrupt shifts identified when further quality assurance and homogenization techniques are applied (Aguilar *et al.*, 2003).

3 | ERROR CHECKING

At each stage of the transcription process, quality assurance measures were employed to preserve the integrity of the data being rescued. Keying guidelines were developed ensuring conformity to World Meteorological Organisation (WMO) standards (WMO, 2016). Monthly totals were examined against the derived sum of the daily entries to identify potentially incorrect data entries. The data were double keyed and the entries from different transcribers compared.

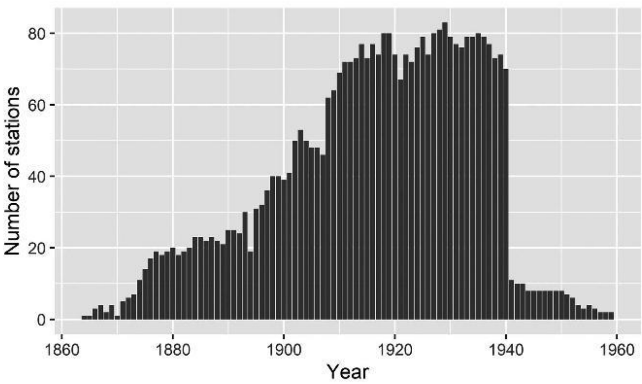


FIGURE 5 Number of stations each year for which daily rainfall data was transcribed. Post-1940 data are available through Met Éireann's website

Sample_Metadata.txt - Notepad

File Edit Format View Help

Station:	**Ardgillan Castle, Balbriggan**
County:	**Dublin**
Met Éireann Reference Code:	**RR/D/14/1**
Lat:	**53°35'09"N**
Lon:	**06°09'42"W**
Observer:	**Records completed by Cap. E. N. Taylor and J. Woodward-Steward.**
Start year:	**1894**
End year:	**1938**
Missing:	**1934 [Monthly totals only]; Oct & Dec 1937 [Monthly totals only].**
Missing data indicator:	**-999**
Time of Observation:	**09:00am**
Unit of measurement:	**inch**
Rain gauge pattern:	**Snowdon**
Diameter:	**5 inches**
Height of top of gauge above ground:	**1 foot**
Height of top of gauge above sea level:	**211 feet [1894-1933]; 203 feet [1934-1938].**
Nearest railway station:	**Balbriggan**
Distance to nearest railway station:	**1.5 miles North-Northeast**
Additional Information:	**1895: Snowfall indicated by *.** **1934: Monthly totals only.** **1937: Monthly totals only - Oct & Dec.** **1938: Record ends in Feb 1938.**
Notes:	**1937: Oct total estimated by nearby stations.** **1937: Dec total received per E.W.M. Murphy.**
Rainfall quality indicator:	**0: Ok** **1: estimated** **2: daily value not available, accumulated value later** **3: Estimated and Cumulative (read)** **4: Trace (<0.05mm, entered as 0.0)** **5: Estimated Trace (same as 4)** **6: Cumulative Trace (same as 4)** **7: Estimated Cumulative Trace (same as 4)** **8: Not Available (no value entered)** **9: Cumulative value including preceding days with indicator 2 (read)**

Windows (CRLF) Ln 1, Col 1 100%

FIGURE 6 Sample metadata file providing information regarding station name and location, observer, record length, missing data, diameter of gauge, changes in gauge height, gauge pattern, time of observation, unit of measurement and distance to nearest railway station. A section for 'Additional Information' provides details of observations made while transcribing the data, for instance, the presence of a multiday accumulation. A separate section for 'Notes' provides a transcription of any handwritten notes recorded on the original record by the observer

Where the entries agreed, the value was provisionally accepted as the raw data value. If the values disagreed, the original record was manually examined to ascertain the true observed value. An examination of errors across all transcriptions revealed a percentage error of <1%. Multiday accumulations were identified and flagged using the original records as a reference. A description of numerical flag values is included in the metadata files. These indicator flags will facilitate the re-distribution of multiday accumulations to the respective days on which no observation was recorded. This will be undertaken prior to the application of further, more sophisticated, quality-assurance techniques. As a final check for transcription errors, the upper and lower 1% of observations (non-zero precipitation) were examined for each individual station record. Values identified as outliers were cross-checked against the original record.

4 | DATASET USE

The data are freely available from the edepositIreland data centre (http://hdl.handle.net/2262/91347). The dataset comprises daily rainfall data for 114 stations at various locations throughout Ireland for varying time periods. Individual station folders contain two files: a data file in ASCII format and a corresponding metadata text file as described in section 2.4. Data files consist of five columns providing the observation date (year, month and day), followed by rain value and indicator value. The indicator value provides information about the nature of the observation and identifies multiday accumulations. A key to the respective indicator values is provided in the metadata files. Rainfall values run continuously from start date to end date of the data recovery period, with missing values denoted using a −999 indicator. Work is currently underway to produce a network of long term, quality assured daily rainfall stations using the datasets whose generation is described in this paper. Preliminary quality assurance checks have been applied to assess the accuracy of the transcription process. A second, comprehensive set of quality assurance techniques will be applied to detect both systematic and non-systematic errors, this will be described in a subsequent paper. Post-1940 daily rainfall records are readily available from Met Éireann's climate database. The newly transcribed data will be added to the database and used to extend these station records back to the late 19th Century. Once joined, the full series will be homogenized and analysed to assess variability and changes in the characteristics of rainfall events over the long-term record.

5 | FUTURE WORK ON DATA RESCUE

The importance of historical climate data is being increasingly acknowledged for its role in supporting effective climate risk management through reanalysis and validation of climate models. As a consequence, climate data rescue has experienced a substantial rejuvenation in recent years, with a number of national and international projects underway, for example Copernicus Climate Change Data Rescue Service (C3S) and the International Data Rescue (I-DARE). The data presented here mark a significant and innovative effort to progress data rescue efforts for Ireland. Nevertheless, a considerable amount of data has not yet been digitized and exist only in hard copy format. Met Éireann, as part of their current operational business plans, has considered opportunities to advance data rescue initiatives within the organization and through collaboration with other agencies to create a comprehensive climate data bank and to facilitate the creation of high-quality data products.

Met Éireann are involved in a number of ongoing projects and initiatives aimed at enhancing climate data availability and accessibility. These include a collaboration with The Central Statistics Office (CSO) to transcribe all parameters and metadata from the eight meteorological registers imaged during the course of this work (see Table 2 for station details). As part of this, transcription of the entire Phoenix Park series has recently been completed. The data series which spans the period 1829 to present is the longest continuous series for the Republic of Ireland and the second longest in Ireland after Armagh Observatory. The extent of the data available for the Phoenix Park makes it one of the most comprehensive series available worldwide.

Engagement of non-experts or 'citizen scientists' on a voluntary basis has become increasingly significant to the rescue and refinement of observational data across multiple scientific disciplines (Bonney *et al.*, 2014). The success of ongoing citizen science applications – for example, OldWeather.org (www.oldweather.org/), the Weather Rescue Project (https://weatherrescue.wordpress.com/) and Data.Rescue@Home (http://www.data-rescue-at-home.org) underscores the potential of crowdsourcing as a data rescue strategy. Further, the data rescue project developed by Ryan *et al.*, (2018) as a collaboration between Met Éireann and Maynooth University will continue as an integral part of the Climate Change module delivered by Dr. Murphy. It is hoped that the project will be developed further through the use of an online platform designed to host Ireland's historical meteorological records. Such an application would facilitate the extension of the data rescue project to other teaching programmes and significantly advance contributions from citizen scientists.

Climate data rescue must be viewed as a continuous, long-term activity (Brönnimann *et al.*, 2018). The priority for Met Éireann over the coming years is to catalogue and image all historical records currently held in the archives for integration into the climate repository, an effort enhanced by the development of resources (including a digital scanner) as part of this work. Continued effort is also being made to recover records held in other libraries and institutes in Ireland and abroad. Met Éireann have recently recruited an archivist to document the vast historical record holdings including meteorological and climatological collections, weather diaries, monthly bulletins, annual reports, weather maps and site inspection reports. Providing access to these records, along with an inventory of what is available is important to promote collaborations and knowledge sharing between interested stakeholders.

ACKNOWLEDGEMENTS

The research presented in this paper was partly funded by the Irish Research Council (IRC) as part of the IRC Employment Based PhD Programme. CM was funded by the Irish Research Council COALESCE scheme (Grant COALESCE/2019/43). The authors are grateful to all the students who contributed to the development of this dataset as part of their GY313 Climate Change module with Dr. Conor Murphy at Maynooth University. We also thank the Maynooth University Research Incentivisation Scheme (RIS) allocation to Dr. Murphy and the Irish Climate Analysis and Research UnitS (ICARUS) for their support in purchasing the Metis digital scanner currently housed in Met Éireann.

OPEN PRACTICES

This article has earned an Open Data badge for making publicly available the digitally shareable data necessary to reproduce the reported results. The data is available at http://hdl.handle.net/2262/91347. Learn more about the Open Practices badges from the Center for OpenScience: https://osf.io/tvyxz/wiki.

ORCID

Ciara Ryan https://orcid.org/0000-0001-6281-5123
Conor Murphy https://orcid.org/0000-0003-4891-2650
Rhonda McGovern https://orcid.org/0000-0002-3931-1944
Mary Curley https://orcid.org/0000-0002-6209-8221

REFERENCES

Aguilar, E., Auer, I., Brunet, M., Peterson, T.C. and Wieringa, J. (2003) Guidelines on climate metadata and homogenization. *World Meteorological Organisation*, 1186, 55.

Allan, R., Endfield, G., Damodaran, V., Adamson, G., Hannaford, M., Carroll, F. *et al.* (2016) Toward integrated historical climate research: The example of Atmospheric Circulation Reconstructions over the Earth. *Wiley Interdisciplinary Reviews: Climate Change*, 7(2), 164–174.

Ashcroft, L., Coll, J.R., Gilabert, A., Domonkos, P., Brunet, M. and Aguilar, E. (2018) A rescued dataset of sub-daily meteorological observations for Europe and the southern Mediterranean region, 1877–2012. *Earth System Science Data*, 10(3), 1613–1635. https://doi.org/10.5194/essd-10-1613-2018

Bonney, R., Shirk, J.L., Phillips, T.B., Wiggins, A., Ballard, H.L., Miller-Rushing, A.J. and Parrish, J.K. (2014) Next steps for citizen science. *Science*, 343, 1436–1437. https://doi.org/10.1126/science.1251554

Brönnimann, S., Brugnara, Y., Allan, R.J., Brunet, M., Compo, G.P. and Crouthamel, R.I. (2018) A roadmap to climate data rescue services. *Geoscience Data Journal*, 5, 28–39. https://doi.org/10.1002/gdj3.56

Butler, C.J., Garcia-Suarez, A. and Palle, E. (2007) Trends and cycles in long Irish meteorological series. *Biology and Environment: Proceedings of the Royal Irish Academy*, 107B, 157–165.

Coll, J.R., van der Schrier, G., Aguilar, E., Rasol, D., Coscarelli, R. and Bishop, A. (2019) *Daily rescued meteorological observations across Europe (1917-1990), PANGAEA.* https://doi.org/10.1594/PANGAEA.896957

Cornes, R.C., Jones, P.D., Briffa, K.R. and Osborn, T.J. (2012) A daily series of mean sea-level pressure for London, 1692–2007. *International Journal of Climatology*, 32, 641–656. https://doi.org/10.1002/joc.2301

Daly, C., Gibson, W.P., Taylor, G.H., Doggett, M.K. and Smith, J.I. (2007) Observer bias in daily precipitation measurements at U.S. Cooperative Network stations. *The Bulletin of the American Meteorological Society*, 88, 899–912.

Dixon, F.E. (1987) Early Irish weather records. In: Shields, L. (Ed.) *The Irish meteorological service: The first fifty years* (pp. 1936–1986). Dublin, Ireland: Met Éireann.

Eden, P. (2009) Traditional weather observing in the UK: An historical overview. *Weather*, 64, 239–245. https://doi.org/10.1002/wea.469

Glasspoole, J. (1952) The British rainfall organisation. *Weather*, 7(3), 85–87. https://doi.org/10.1002/j.1477-8696.1952.tb01449.x

Green, J.S., Klatt, M., Morrissey, M. and Postawko, S. (2008) The comprehensive pacific rainfall database. *Journal of Atmospheric and Oceanic Technology*, 25, 71–82.

Hawkins, E., Burt, S., Thomas, S., Lockwood, M., Brohan, P., Richardson, H. *et al.* (2019) Hourly weather observations from the Scottish Highlands (1883–1904) rescued by volunteer citizen scientists. *Geoscience Data Journal*, 6, 2, 160–173. https://doi.org/10.1002/gdj3.79

Kunkel, K.E., Easterling, D.R., Redmond, K., Hubbard, K., Andsager, K., Kruk, M. *et al.* (2005) Quality control of pre-1948 cooperative observer network data. *Journal of Atmospheric and Oceanic Technology*, 22, 1691–1705.

Mateus, C., Potito, A. and Curley, M. (2020) Reconstruction of a long-term historical daily maximum and minimum air temperature network dataset for Ireland (1831–1968). *Geoscience Data Journal*, 1–14. https://doi.org/10.1002/gdj3.92

Matthews, T., Murphy, C., Wilby, R.L. and Harrigan, S. (2014) Stormiest winter on record for Ireland and UK. *Nature Climate Change*, 4, 738–740. https://doi.org/10.1038/nclimate2336

Murphy, C., Broderick, C., Burt, T.P., Curley, M., Duffy, C., Hall, J. *et al.* (2018) A 305-year continuous monthly rainfall series for the island of Ireland (1711–2016). *Climate of the Past*, 14, 413–440. https://doi.org/10.5194/cp-14-413-2018

Murphy, C.Wilby, R.L., Matthews, T.K.R., Thorne, P., Broderick, C., Fealy, R. *et al.* (2019) Multi-century trends to wetter winters and drier summers in the England and Wales precipitation series explained by observational and sampling bias in early records. *International Journal of Climatology*, 40(1), 610–619. https://doi.org/10.1002/joc.6208

Noone, S., Murphy, C., Coll, J., Matthews, T., Mullan, D., Wilby, R.L. *et al.* (2015) Homogenization and analysis of an expanded long-term monthly rainfall network for the Island of Ireland (1850–2010). *International Journal of Climatology*, 36, 2837–2853. https://doi.org/10.1002/joc.4522

Noone, S., Broderick, C., Duffy, C., Matthews, T., Wilby, R.L. and Murphy, C. (2017) A 250-year drought catalogue for the island of Ireland (1765–2015). *International Journal of Climatology*, 37, 239–254. https://doi.org/10.1002/joc.4999

Ryan, C., Duffy, C., Broderick, C., Thorne, P.W., Curley, M., Walsh, S. *et al.* (2018) Integrating data rescue into the classroom. *Bulletin of the American Meteorological Society*, 99(9), 1757–1764. https://doi.org/10.1175/BAMS-D17-0147.1

Treanor, M., Keane, A. and Lonergan, T. (2011) *Meteorological Observations in Ireland 1855-1976: an archival Finding Aid.* Technical Note 6. Met Éireann.

Walsh, S. (2013) Setting the scene: the climate of Ireland 1900-2012. In Gleeson, E., McGrath, R. and Treanor, M. (Eds) *Ireland's climate: The road ahead.* Dublin, Ireland: Met Éireann, pp. 17–19.

Wilby, R.L., Noone, S., Murphy, C., Matthews, T., Harrigan, S. and Broderick, C. (2016) An evaluation of persistent meteorological drought using a homogeneous Island of Ireland precipitation network. *International Journal of Climatology*, 36, 2854–2865. https://doi.org/10.1002/joc.4523

World Meteorological Organization (2016) *Guidelines on Best Practices for Climate Data Rescue 2016.*

How to cite this article: Ryan C, Murphy C, McGovern R, Curley M, Walsh S; 476 students. Ireland's pre-1940 daily rainfall records. *Geosci Data J.* 2020;00:1–13. https://doi.org/10.1002/gdj3.103

International Journal of Climatology

Long-term trends in extreme precipitation indices in Ireland

Journal:	*International Journal of Climatology*
Manuscript ID	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a

Complete List of Authors:	Ryan, Ciara; National University of Ireland Maynooth, Irish Climate Analysis and Research UnitS (ICARUS), Department of Geography Curley, Mary; Met Eireann, Climate and Observations Division Walsh, Séamus; Met Éireann, Climatology and Observations Murphy, Conor; Maynooth University, Irish Climate Analysis and Research UnitS (ICARUS), Department of Geography

Keywords:	Data rescue, Precipitation extremes, Long-term records, Quality control, Homogenisation
Country Keywords:	Ireland

Long-term trends in extreme precipitation indices in Ireland

Ciara Ryan[1,2], Mary Curley[2], Séamus Walsh[2], Conor Murphy[1]

[1]Irish Climate Analysis and Research Units, Department of Geography, Maynooth University, Maynooth, Co. Kildare

[2]Climate Services, Research and Applications Division, Met Éireann, Dublin, Ireland

Corresponding Author: Ciara Ryan, ciara.ryan@mu.ie

Funding information: This article was partly funded by a grant from the Irish Research Council (IRC) as part of the IRC Employment Based PhD Programme (EBP/2017/441). CM was funded by the Irish Research Council COALESCE scheme (Grant COALESCE/2019/43).

Keywords: Ireland, data rescue, precipitation extremes, long-term trends.

Abstract

Knowledge of long-term changes in climate extremes is vital to better understand climate variability and place present day extreme events in historical context. Analysis of trends in extreme precipitation in Ireland have largely been limited to the second half of the 20th century due to lack of data availability in digital format. Using recently digitised data, this study provides the first assessment of long-term changes in extreme daily precipitation observed at 30 locations across Ireland. Quality control of rescued data is carried out before selected long-term stations are tested for homogeneity using RHtests software. Details of detected breakpoints and the application of adjustments to daily values are discussed. Eleven extreme precipitation indices are calculated on an annual basis and analysed to determine spatial and temporal trends in the frequency, intensity and magnitude of observed precipitation. The persistence of trends for varying record lengths and for two fixed periods (1910-2019 and 1940-2019) of analysis is assessed for all stations and indices. Results show increases in precipitation intensity, especially notable in the east and southeast of the island. Our findings also show that the contribution of heavy and extreme precipitation events to annual totals is increasing, while there was no persistent trends in annual totals or consecutive wet or dry days.

1. Introduction

Understanding long-term variability and trends in precipitation is critical for water resource management and assessing potential flood risk. Long-term series of precipitation have been produced for many regions to detect changes in the spatial and temporal variation of this essential climate variable (e.g. Ashcroft et al., 2018; Brunet et al. 2014; Auer at al. 2007; De Jongh et al., 2006; Moberg et al., 2006). In Ireland, previous research has focused on the development and analysis of long-term monthly precipitation series (e.g. Murphy et al. 2018; Noone et al. 2015), while analysis of daily precipitation has largely been limited to the post 1940 period, for which digitised observational and gridded precipitation products are available. Moreover, no previous analysis has assessed long-term changes in precipitation extremes for Ireland using the suite of indices defined by the CCl/ CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI).

Previous analysis of Irish annual precipitation totals reveals an increase in mean annual precipitation of approximately 5% in 1981–2010, compared with the 30-year period 1961–1990 (Walsh, 2016). In general, larger increases in precipitation amounts are observed in the west of the country. An increase in the frequency of heavy rain days (days with rainfall greater than 10 mm) has been reported at the majority of stations analysed over the period 1961-2010, however, regional variations are evident with occasionally conflicting trends from stations that are geographically relatively close (Walsh and Dwyer, 2012). Sheridan (2001) reported an increase in annual accumulations over the period 1941-1999, particularly in the west of the country, with more notable increases reported since the early 1970s. McElwain and Sweeney (2007) analysed monthly data at Birr and Malin Head from 1890 to 2003 and found significant increases in annual totals at Malin Head with no significant trends detected in the annual precipitation series at Birr. Leahy and Kiely (2011) examined changes in short-duration (1hr-24hr) precipitation extremes and found strong evidence for increasing trends since the late 1970s, particularly in the west of the country, coinciding with changes in the North Atlantic Oscillation.

Knowledge of changes in extreme precipitation events are a key aspect of monitoring climate change as increases or decreases in the frequency and/or magnitude of extreme precipitation events have high environmental and socio-economic impacts. However, the relative sparseness of long-term digitally available records at appropriate spatial and temporal scales hampers our ability to fully understand changes in precipitation extremes (Brunet and Jones, 2011). The spatial variability of precipitation, together with changes in observer practices, gauge location and design, mean that developing reliable, long-term precipitation series is often a challenging task (Wilby et al., 2017). Quality control procedures are crucial to identify both systematic and non-systematic errors that can confound the detection and interpretation of trends. Errors can occur as a result of changes in measuring techniques, observational and recording practices, and transfer of data from paper to digital format (Slater et al., 2020). A well-defined quality control process should be able to flag data errors that could compromise the analysis of natural climate variability, particularly in the study of extreme events (Llabrés-Brustenga et al., 2019). Homogenisation of data to remove spurious non-climatic features from long-term instrumental records is integral to the development of accurate climate data and the assessment of climate variability (Freitas et al., 2013; Vertačnik et al., 2015). While many methods have been developed to evaluate the homogeneity of monthly to annual-resolution climate data, homogenisation of daily data is a highly complex task that is still under development within the research community (Venema et al., 2018).

Recently, Ryan et al. (2020) completed the rescue of 3616 station years of daily precipitation data (~1.32 million daily values) for stations across Ireland. Details of the data rescue process and dataset development are presented in previous publications (Ryan et al., 2018; Ryan et al., 2020) to provide a fully traceable dataset. The availability of this data now makes it possible to assess changes in daily precipitation extremes back to the early 20th Century. While Ryan et al. (2020) implement basic quality assurance tests, in this paper we conduct a more rigorous quality control of rescued data before extracting and homogenising long-term daily series for stations across Ireland. We then derive indices of annual extremes and assess these for evidence of trends.

The remainder of the paper is organized as follows: Section 2 provides a description of the data, the quality control (QC) and homogenisation techniques used to reduce errors in the dataset and methods used to test the direction and magnitude of trends. Section 3 presents the results, while Section 4 provides a discussion of key findings, limitations and directions for future work, before drawing conclusions in section 5.

2. Data and Methods

2.1 Quality control of data

The dataset used in this analysis was constructed from recently rescued and digitised historical (pre-1940) daily precipitation data (Ryan et al. 2020) and post 1940 daily precipitation data extracted from Met Éireann's database. The latter contains daily and monthly precipitation records from 1941 onwards which have previously undergone quality control tests, full details of which are provided by Walsh (2016). These include a combination of visual checks, spatial techniques, cross-validation of daily interpolated values, comparison with daily radar accumulations, where available, and an automated system for the re-distribution of cumulative daily values. The pre-1940 dataset comprises daily precipitation data for 114 stations at various locations throughout Ireland for varying time periods (Figure 1).

Ryan et al. (2020) undertook an initial quality assessment of the pre-1940 digitised data. This involved manually cross-checking all flagged observations against the original record and metadata. At each stage of the transcription process, quality assurance measures were employed to preserve the integrity of the data being rescued. Keying guidelines were developed ensuring conformity to World

Meteorological Organisation (WMO) standards (WMO, 2016). Monthly totals were compared to the derived sum of the daily entries to identify potentially incorrect data entries. The data were double keyed and the entries from different transcribers compared. Where the entries agreed, the value was provisionally accepted as the raw data value. If the values were different, the original record was manually examined to ascertain the true observed value. An examination of errors across all transcriptions revealed a percentage error of <1%. As a final check for transcription errors, the upper and lower 1% of observations (non-zero precipitation) were examined for each individual station record. Values identified as outliers were cross-checked against the original record. Following the application of these initial quality assurance checks the raw data were published to enable users to access the original data (Ryan et al., 2020). Additional quality control checks are undertaken here. These can be grouped into four categories: basic integrity tests, tolerance tests, temporal consistency tests, and spatial consistency tests. Details of the QC procedures deployed are provided in Table 1. Each test produced a list of values flagged at each station which were subject to manual inspection. The detection of false positives (e.g. extreme events flagged as outliers) is a common issue in QC assessments and manual inspection, combined with local/regional knowledge of climatological processes ensured that these important observations were retained.

For the majority of values flagged by the basic integrity tests, the value was adjusted following manual inspection, for example, in cases where zeros were used in place of the missing data indicator or undocumented cumulative totals. Values classified as suspect were accepted if they satisfied each of the following criteria: the value agreed with that recorded on the original record; similar observations were recorded at neighbouring stations; the value was deemed physically reasonable for that station/region/season. Values were rejected and set to missing if they were flagged as suspect and did not satisfy these criteria or if there was evidence reported in the metadata to suggest that the gauge was defective at the time of observation. In total, 59,274 values were flagged and investigated (~4.5% of the data), with slightly less than 1% of the values removed from the quality-controlled version of the dataset used to assess trends. Cumulative daily values which were identified and flagged during the transcription process were re-distributed to the respective days on which no observation was recorded. For cumulative values requiring redistribution, a first guess value for days on which no observation was recorded is estimated by means of inverse distance weighted interpolation; the estimated daily amounts are summed and the ratio of this total to the observed cumulative total is calculated; the final estimated daily values are obtained by adjusting the first guess values according to this ratio (Walsh, 2016).

2.2 Deriving long-term station series

The quality controlled pre-1940 rescued data was added to the database containing post 1940 data. For stations with matching station identifiers and locations, the pre-1940 data was joined to the post-1940 data to produce continuous station series. To address gaps in data series, the method developed at Met Éireann by Walsh (2016) which performed well for producing gridded precipitation datasets, was used to infill and extend station series using the rescued data. First a regression-kriging interpolation model was used to generate monthly values for all stations in the database. Thus, for stations which had gaps in data, the monthly total was estimated using nearby station data. Next, to generate daily values, inverse distance weighted interpolation was used to disaggregate the monthly interpolated values to provide a complete series of daily values, using the method described in section 2.1 for cumulative daily values.. Finally, stations for analysis were selected. Given the lower station density in the pre-1940 period, only sites with neighbouring station data available were selected when developing long-term series. In total 36 stations were identified based on record length, continuity, completeness and proximity of pre-1940 station locations. The mean record length is 118 years. The longest series length is 146 years for Foulkesmill (Longraigue), the shortest is 91 years for Portumna O.P.W. The selected stations (Table 2) were grouped into two categories as follows:

- Group A stations are those with continuous observations at the same site from start year to end year. Missing data were infilled using the output of the regression kriging interpolation model. The average amount of missing data across all group A stations is 4.4%. Valentia and Phoenix Park stations have 0% missing data over the period of record. The maximum amount of missing data for any station in group A is 13.6% for Kenmare.
- Group B stations represent those from the post-1940 dataset for which nearby (<15km) rescued data were available to extend records into the pre-1940 period. B stations were generated using the output of the regression kriging interpolation model.

2.3 Homogenisation of derived station series

Relative homogenisation methods (i.e. testing with respect to a homogenous neighbouring station) are considered more robust than absolute methods, provided station records are sufficiently correlated (Gubler et al., 2017). However, the high spatial variability of daily precipitation make it difficult to find suitable reference series except in the case of parallel measurements (Wang et al., 2010). Moreover, given the sparse spatial density of the station network in the early period (Figure 1), the current homogeneity assessment is restricted to the application of absolute methods. We employ a two-step approach. First, for each station, breakpoints in the monthly series, derived from daily precipitation values, are detected and adjusted using the RHtests software developed by Wang and Feng (2013) at the Climate Research Division, Atmospheric Science and Technology Directorate, Science and Technology Branch, Environment Canada (http://etccdi.pacificclimate.org/software.shtml). Second, the output from the RHtests software is examined for further breaks by testing the homogeneity of the derived annual indices using the standard normal homogeneity test (SNHT) (Alexandersson, 1986) and the Pettitt test (Pettitt, 1979). This approach supports the selection of series for the analysis of trends.

The RHtests_dlyPrcp software package is specifically designed for homogenisation of daily precipitation data which are non-continuous, non-negative, and non-normally distributed. Breakpoints were detected in the monthly log-transformed series using the RHtests software which is based on the penalized maximal *t* test (Wang et al., 2007) and the penalized maximal *F* test (Wang, 2008b). These are embedded in a recursive testing algorithm (Wang, 2008a), with the lag-1 autocorrelation of the time series being empirically accounted for. The software facilitates metadata integration by testing both known and unknown breakpoints; hence significant breakpoints (0.05 level) were identified using a combination of the statistical methods and station reports.

Specifically, the software was first run to detect all breakpoints that could be significant at the nominal level even without metadata support, referred to as Type-1 breakpoints. Next, a second function was run to detect additional breakpoints (Type-0 breakpoints) which are deemed significant only if supported by reliable metadata. For each station series, the available metadata was investigated to determine whether there was evidence at or near the identified breakpoint dates to support the shifts. Only those Type-0 breakpoints that were supported by metadata, along with all Type-1 changepoints were retained.

Before adjusting the daily series, discontinuities in the occurrence frequency of precipitation were investigated. Such discontinuities result from changes in the unit of measurement, measuring precision, and observing practices (Wang et al., 2010). For each station, the series of daily precipitation amounts greater than a given threshold ($>P_{thr}$) was examined by varying P_{thr} over a set of small values that reflect changes in the measuring precision (i.e., 0.2mm, 0.3mm, 0.5mm). This assessment revealed the presence of discontinuities in the frequency of small measured amounts in the pre-1940 record, most likely related to changes in the unit of measurement (i.e. from imperial to metric). Given that the objective of the present study is to examine trends in indices which require daily precipitation >=1mm, adjustments were applied to the series of $>P_{thr}$ daily precipitation data that was found to be free of frequency discontinuity. This P_{thr} varies across station series but does not exceed $P_{thr} = 0.5$mm. It should be noted that all $<P_{thr}$ values in the series are left unadjusted, and thus frequency discontinuities remain, affecting values between 0 and P_{thr} inclusive.

Where breaks were detected, adjustments were applied to daily station series by means of the quantile mapping (QM) algorithm in the RHtests software (Wang et al., 2010). The algorithm was applied as follows: the non-zero daily precipitation series were detrended using the linear trend estimated from a multiphase linear regression fit that accounts for the mean shifts at detected breakpoints in the data series (Wang, 2008b). For each breakpoint to be adjusted, the data in the segment immediately before and after the breakpoint is used to estimate the probability distribution function (PDF). We use all available data in a segment to estimate the PDF, as estimates of the PDF (and hence the QM adjustments) using shorter periods (e.g. < 30 years) can contain large sampling uncertainty (Wang et al., 2010). Each segment is divided into Mq quantile categories, and for each quantile, the differences between the means of the two periods is computed. A natural cubic spline is then fitted to the Mq category-mean differences between the segment to be adjusted and the base segment which is then used to estimate the QM adjustments needed to make the data series homogeneous. The choice of the number of quantile categories (Mq) used to estimate the spline to derive the QM adjustments was set at Mq = 4 ~ 16, depending on the length of the shortest segment in a given data series. We use the most recent segment as the base on the assumption that technologies and observational practices generally improve over time. We recognise that the choice of parameters is subjective and may not be optimal for all series, however, the use of detailed metadata to inform decisions increases confidence in the performance of the algorithm, while sensitivity analysis to the choice of different parameters resulted in minor changes to adjustments made.

Following Santo et al. (2014), further tests of the adjusted series were carried out by testing the homogeneity of the derived PRCPTOT and R5mm (see next section) annual indices using the standard normal absolute homogeneity test (SNHT) and the Pettitt test. The p-value of potential breakpoints were plotted to facilitate a relative assessment of breakpoints across all stations. To determine the significance of potential break points, the null hypothesis (no breakpoint detected) was tested against the alternative at the 0.05 level. As a final assessment, a visual inspection was carried out by plotting the original and the adjusted annual precipitation totals for each series found to be inhomogenous.

2.4 Indices and Trend detection

The derived quality controlled, homogenised long-term daily series were assessed for evidence of trends in the characteristics of extreme precipitation. For this purpose, eleven indices defined by the CCl/ CLIVAR/JCOMM Expert Team on Climate Change Detection and Indices (ETCCDI, http://cccma.seos.uvic.ca/ETCCDI/) were calculated for each series and investigated for evidence of trends. The selected indices allow examination of changes in intensity, frequency, and duration of extreme precipitation events (Alexander et al., 2006). Precipitation indices are summarised in Table 3. For deriving indices, we define a wet day as a day with ≥1 mm precipitation, while the common 30-year base period used to define thresholds for percentile-based indices is 1961-1990.

Indices were derived for two set periods: 1910-2019 and 1940-2019. The 1910-2019 period was selected as the longest common period for which all stations have available records, the shorter 1940-2019 timespan represents the period of available digital records prior to our data rescue efforts. Evidence for monotonic trends was assessed using the Modified Mann-Kendall (MK) test (Hamed and Rao, 1998), a non-parametric rank-based method. The standardized MK test statistic (MKZs) follows the standard normal distribution with a mean of zero and variance of one. A positive (negative) value of MKZs indicates an increasing (decreasing) trend. Statistical significance was evaluated with probability of Type I error set at the 5 % significance level. A two tailed MK test was applied, hence the null hypothesis of no trend (increasing or decreasing) is rejected when |MKZs| > 1.96.

The MK test requires data to be independent (i.e. free from serial correlation) as positive serial correlation increases the likelihood of Type 1 errors or incorrect rejection of a true null hypothesis. Therefore, application of the Modified MK test includes a test to detect positive lag-1 serial correlation at the 5% level using the autocorrelation function (ACF). The existence of trend influences the correct estimate of serial correlation; therefore, original time-series are detrended to form a 'trend-removed' residual series before the ACF is applied (Hamed and Rao, 1998).

The magnitude of trend (β) is estimated using the approach proposed by Theil (1950) and Sen (1968), hereinafter referred to as TSA. To facilitate a relative comparison among different sites, the approach presented by Yue and Hashino (2007) is utilised where the magnitude of trend TSA_{rel} (%) for each time-series is expressed as a percentage change over the period of record of n years relative to the mean (μ) for the period, given by:

$$\text{TSArel}\ (\%) = \left(\frac{\beta\ \times n}{\mu}\right) \times\ 100$$

For brevity, Tables 6 and 7 present the direction of change and proportion of statistically significant (5 % level) trends along with the median magnitude of trend (TSA_{rel} %) and the median number of days per year for the 11 extreme precipitation indices for both fixed periods. Full details of trend statistics calculated for each station and indicator are provided as Supplementary Information.

The persistence or dependency of trends on period of record was examined for each station by systematically reducing the start year of analysis from the whole record to a minimum of 30 years following the approach of Murphy et al. (2013) and Noone et al. (2015). For each indicator, the MKZs statistic was derived first for the full record e.g. 1890-2019, then 1891-2019, and so on until 1990-2019. MKZs values were plotted for each iteration of start years to examine the temporal evolution of trend throughout the period of record. Unlike the fixed period analysis above, the persistence of trends was evaluated for the full record available at each station.

3. Results

3.1 Homogenisation

Following application of the RHtest software, twenty stations were found to be homogenous and did not require adjustment. In total, 22 breaks were detected in 16 of the 36 stations. Multiple breaks were found in five records: three in the Newport series; two each in the Malin Head, Belmullet, Glengarriff, and Drumsna series; and a single break point in each of the remaining eleven series. Approximately 80% of detected breakpoints are supported by information contained in the metadata regarding instrument and site changes. Details extracted from metadata for each identified break are presented in Table 4.

To examine the impact of the adjustments on the station series, the original and adjusted annual precipitation totals were plotted for the 16 inhomogeneous series (Figure 2). The magnitude of the adjustments applied to the original data varies across station series and datum as per the QM approach, with larger (smaller) values undergoing a greater (lesser) adjustment. Table 5 provides details of the mean, minimum and maximum annual percentage difference between the original and the adjusted series.

The Pettitt test and the SNHT were applied to the annual series of PRCPTOT and R5mm to detect residual breakpoints following the application of the RHtests software. Across the 36 station series tested, the Pettitt test detected nine significant breakpoints in the PRCPTOT series. Five significant breakpoints were detected in the series of R5mm. The SNHT test detected eleven significant breakpoints in the PRCPTOT series, five of which coincide with breakpoints detected by the Pettitt test. Five significant breakpoints were detected in the R5mm series, four of which were also detected by the Pettitt test. Of the total 30 breakpoints detected in PRCTOT and R5mm by both tests, the timing for 21 are consistent across the network if stations. These occur in the 1970's and are consistent in timing with a shift to more positive NAO index resulting in stronger mid-latitude westerlies and increased precipitation over Ireland (Kiely, 1999). As would be expected in the presence of widespread natural climate variability, these breakpoints were not detected by Coll et al., 2018 using relative homogeneity methods. Given their consistency in timing across the network and their concurrent timing with a well know shift in the NAO we attribute these break points to natural climate variability.

Six stations were identified as having potential remaining discontinuities and were subsequently excluded from further analysis. These stations are Abbeyfeale (st3310), Carndolla (st2227), Westport (st1433), Creeslough (st944), Drumsna (st1529) and Kenmare (st603). The Drumsna gauge is known to have underestimated precipitation prior to 1942 when a new gauge was installed. A comparison of precipitation totals from both gauges over a 13-month period revealed that the new gauge reported 154% of the old gauge. The Drumsna series was corrected for a breakpoint detected in 1942, however, remaining discontinuities are evident in the adjusted series. The other five stations are located in the west and the northwest of the country where station density is lowest in the pre-1940 period. The Creeslough series was corrected for a breakpoint detected in 1928, however, a visual assessment of the annual series following adjustment suggests the presence of a remaining shift in the early 1960's that was not detected by RHtests. No breakpoints were detected by RHtests in the series for Abbeyfeale, Carndolla, Kenmare and Westport. However, a visual inspection of the original annual time series, the annual PRCPTOT and the R5mm time series for individual stations as well as a comparison of the series across all stations revealed large trends and potential breakpoints that were not detected by RHtests in testing the monthly series.

Following the removal of these six stations, the Pettitt test detected four significant breakpoints in the annual PRCPTOT and one significant breakpoint in the annual series of R5mm when applied to the remaining 30 station series. Figure 3 displays the p-values of potential break points for the Pettitt and SNHT tests. For the Pettitt test, all five breakpoints were detected in the period 1976-1979 and are consistent with the timing of other non-significant break points across the network. The SNHT detected six significant breakpoints in the annual PRCPTOT series and one significant breakpoint in the annual R5mm series. Two of these were detected in the period 1977-78. In the absence of evidence for widespread changes in measurement practices, the timing of these breakpoints suggest that a genuine climatic process was identified as inhomogeneities in the majority of series. Specifically, a shift in the North Atlantic Oscillation (NAO) to a predominantly positive phase since the mid-1970s, (Kiely, 1999). The remaining breakpoints were detected in 2007 at Strokestown (Carrowclogher) and Inagh (Mt. Callan) and in the years 1987 and 1910 at Ballyhaise and Glenasmole, respectively. Only the break point at Glenasmole in 1910 occurs at a time when no other stations show non-significant breakpoints. It was noted during the QC process that the occurrence of heavy precipitation during December 2010 at Glenasmole resulted in large monthly and annual totals. The values were investigated and accepted based on consistency with observations recorded at neighbouring stations. Such large values can introduce spurious breakpoints in the time series. Moreover, this breakpoint is detected in both the PRCPTOT and R5mm series by the SNHT but is not detected by the Pettitt test in either series. Previous studies (e.g. Ducre-Robitaille et al., 2003; Toreti et al., 2011) have reported increases in SNHT false break detection at the beginning and the end of the series. Given the presence of this breakpoint at the beginning of the series we retain it for examining trends.

3.2 Fixed period trend analysis

Trends were assessed for the network of 30 homogenised daily precipitation series for two fixed periods: 1910-2019 (Table 6) and 1940-2019 (Table 7). For both periods, increasing trends in extreme precipitation indices are more prevalent than decreasing, with the exception of CDD and CWD in the period 1910-2019 and CDD in the period 1940-2019.

The percentage of stations reporting increasing trends in Rx1day and Rx5day is consistent over both periods of analysis. However, there is a marked increase in the number of stations reporting statistically significant increases in Rx5day in the period 1910-2019. Of the stations reporting decreasing trends none are found to be statistically significant. There is a shift in the direction of trends at both Valentia and Glengarriff located in the southwest from a negative trend in the period 1940-2019 to a positive trend over the longer period, with a significant trend reported for Glengarriff. Large variations in trend magnitude exist across the network. Largest trends found for Rx5day with a median increase across all stations of 6.06 percent over the period 1910-2019, increasing to 8.35 percent for 1940-2019. For the latter, the 95 percent confidence interval in trend magnitude across the network spans -6.81 to 23.77 percent. Largest trends are evident in the midlands for the period 1910-2019, extending to the southeast for the 1940-2019 period of analysis.

The percentage contribution from wet days (R95pTOT) and very wet days (R95pTOT) to total precipitation is also dominated by increasing trends. In the period 1910-2019, 83.3 percent of stations show increasing trends, 13.3 percent significant, with 76.7 percent of stations showing increasing trends in R99pTOT of which 16.7 percent are significant. For both periods, R95pTOT has the largest median increase in trend magnitude (TSArel ~14 percent) across all stations. Again, there is large variability in trend magnitude across stations. For the period 1910-2019, for instance, R95pTOT shows a median increase of 5.07 percent across the network, with the 95 percent confidence intervals ranging from -5.07 to 61.87 percent. Largest trends in both periods are found in the midlands and the east and southeast of the island. Significant trends reported in R95pTOT and R99pTOT at Foulkesmill (Longraigue), Portlaw (Mayfield II) and Cashel (Ballinamona) over the shorter period of analysis (1940-2019) are not observed in the longer record (1910-2019), while the longer record shows significant trends at Inagh (Mt. Callan) and Mullingar. The record for Markree station, situated in the northwest, shows the opposite. For R95pTOT, positive trends are reported in both periods of record with significant trends observed in the 1940-2019 period. For R99pTOT, negative trends are observed in both periods of record with significant trends observed in the 1940-2019 period.

CDD reveals little change, with only one significant trend reported in either period. Overall, there is a relatively even split between increasing and decreasing trends in both periods. Figures 4 and 5 indicate that decreasing trends in CDD tend to be more prevalent in the south and southwest, while increasing trends tend to be more prevalent in the midland region. In the period 1910-2019, 43.3 percent of stations report increases (10 percent significant) and 56.7 percent of stations report decreases (13.3 percent significant) in CWD. However, there is shift in the direction of trends reported in the period 1940-2019, with 70 percent of stations reporting increasing trends in CWD, 20 percent of which are significant. The shift in direction of trend is most prevalent in the southwest of the country. Median trend in number of days per year across all stations reveals no change in either fixed period for both CDD and CWD.

Increasing trends in indices measuring precipitation threshold frequency are reported in both periods of analysis for all three indicators i.e., R5mm, R10mm, R20mm, however, the magnitude of trends are small (see Tables). The number of stations reporting increasing trends in R5mm decreases slightly in the period 1910-2019. Increasing trends in R5mm are observed at stations in the north and northwest, namely Malin Head, Belmullet and Newport which are significant over the longer period of record, whereas stations in the southwest show significant trends in the shorter period which are not observed in the period 1910-2019. The number of stations showing increasing trends in R10mm remains stable over both periods, however, a higher number of stations report significant increases over the period 1940-2019. There is an increase in the number of stations reporting increasing trends in R20mm for the period 1910-2019 compared to the period 1940-2019. The largest trends in R20mm in both periods of analysis are observed in the south and southeast of the country. No stations report statistically significant decreases in the frequency of precipitation events ≥5 mm, ≥10 mm or ≥20mm in either of the fixed periods assessed.

Both the number of stations reporting increasing trends and the median magnitude of change in intensity (SDII) is relatively consistent over the two periods of analysis, however, there is a slight increase in the number of stations reporting statistically significant increases over the period 1910-2019. These significant increasing trends in SDII are most notable in the east and southeast of the country. The magnitude of trends in intensity are also very consistent across both study periods. For the period 1910-2019 the median magnitude of change in SDII across the network is 3.84 percent, with the 95 percent confidence interval ranging from -2.02 to 9.75 percent.

The number of stations reporting statistically significant increases in total annual wet-day (≥ 1 mm) precipitation (PRCPTOT) more than doubles in the period 1940-2019 compared to the period 1910-2019. There is also a shift in the direction of trend in some stations in the southeast over the shorter period of analysis. No significant decreases are reported in either period analysed. A slight decrease is noted in the median magnitude of change reported in the period 1910-2019 compared to the period 1940-2019.

3.3 Persistence of trends over full period of record

While the previous section highlights the predominance of increasing trends in daily precipitation indices, the dependence of results on the two different periods of record was also noted. In this section we assess the robustness of trends for different start dates over the full period of record for all 30 stations and indices. Plotting the MKZs for all stations/indices (see Figure 6) also allows inspection of trends in individual stations relative to the pattern of change across the network, potentially identifying stations with trends that depart from the pattern of the wider network.

CDD shows a lack of statistically significant trends for all stations across the network for varying start years. Similarly, the majority of stations show no significant trends for varying start years for the CWD index. However, for some stations there is evidence that tests commencing before the 1900s show significant increasing trends, together with tests commencing between 1950 and 1970. Significant decreasing trends in CWD are evident for other stations for tests commencing in the early decades of the 1900s. However, these are not persistent for varying start dates. Two stations show trends in CWD that depart from the general pattern of trends across the network. For tests commencing prior to 1950, Strokestown (Carrowclogher) shows large and significant increasing trends (e.g. MKZs >4 for tests commencing in 1910). Kilskyre (Robinstown) shows large and significant decreasing trends for tests commencing after 1960. These results suggest remaining inhomogeneities at these stations affecting trends in CWD.

PRCPTOT shows a predominance of increasing trends across the full period of record for the majority of stations. Notably, there is greater variation in trend direction and magnitude for tests commencing prior to the 1900s. Persistently significant increasing trends are evident for stations/tests commencing prior to the 1970s. For tests commencing after the 1970s, the vast majority of stations show non-significant trends. The Markree series shows a notable deviation for the general pattern of trends in PRCPTOT for tests commencing before 1900, with raises suspicion about the quality of the early part of this series. Similarly, the Tycor series shows significant decreasing trends for tests commencing prior to 1905.

SDII shows the largest variation in trend magnitude and direction across the network of stations, with different stations showing persistent significant increasing and decreasing trends for tests commencing throughout the record. R95pTOT and R99pTOT are dominated by increasing trends for different start years, with persistent significant increasing trends evident for some stations for tests commencing since the 1930s. Only a small number of stations show weak, non-significant decreasing trends in these indices. Both R5mm and R10mm show a predominance of increasing trends for most stations for different test periods. For tests commencing post 1970 there is a tendency for non-significant trends. For stations with records extending prior to 1900 there is greater variation across stations in trend tests commencing in the late 1800s. For R10mm Markree and Newport station series show significant increasing and decreasing trends, respectively, for tests commencing prior to 1890. While the number of stations available for comparison decreases at this time, the trends appear larger than those at other stations. Trends in Rx1day and Rx5day precipitation show coherence across the network for varying start dates. On the whole, trends in Rx5day tend to be persistently positive for the majority of stations, irrespective of start date.

4. Discussion

This study derived quality controlled long-term daily precipitation series for 36 stations across Ireland by leveraging data rescue efforts to extend existing digital records available from 1940. The average series length is 118 years, with some stations dating back to 1870s. Following quality control and homogenisation, trends in precipitation extremes from annual ETCCDI indices were examined for 30 stations. Results indicate the predominance of increasing trends across the network. However, the direction, magnitude and significance of trends is dependent on the period of analysis, emphasising the importance of the long records developed here. The analysis of trend persistence for varying start years indicates the influential role of low frequency variability on trend results. Of particular influence has been the shift to more prevalent positive NAO conditions around the late 1970s and a greater predominance of westerly airflow, the impact of which is most evident for the CWD, PRCPTOT, R5mm and R10mm indices. The impact of this change on regional circulation has been widely recognised (e.g Hurrel, 1995; Kiely, 1999; Murphy et al., 2013).

Most previous assessments of precipitation in Ireland have been limited to the post-1940 record and have not undertaken assessment of ETCCDI indices (e.g Walsh and Dwyer, 2012; Sheridan, 2001) making direct comparison difficult. Assessments undertaken for longer records have focused on seasonal and annual totals (e.g. Butler et al., 1998; McElwain and Sweeney, 2007; Noone et al, 2015). These studies have drawn attention to the increase in annual totals, particularly for stations on the western seaboard. Our results show that of the indices examined here, PRCPTOT is most sensitive to the influence of low frequency variability and period assessed. We find the greatest magnitude of trends for indices measuring the contribution of heavy precipitation events to annual totals (i.e.

R95pTOT and R99pTOT), particularly in the south and south east of the island. The largest number of significant increasing trends is reported for intensity (SDII), also in the east and south of the island.

A crucial step in assessing trends is the quality assurance of available data to ensure derived results are not an artefact of measurement practice or data processing. It is widely accepted that the reliability and accuracy of homogenisation procedures generally decrease with the increasing temporal resolution of homogenisation (Venema et al., 2012). The task is further complicated for early records as station density typically decreases with time (Hollis et al., 2019). Therefore, homogenisation of daily data is considered to be a much more challenging problem than homogenisation at monthly or annual scales (Coll et al., 2018). Relatively few homogenisation algorithms include methods for the homogenisation of daily climate series and those that do e.g. MASH (Szentimrey, 1999, 2014) and ACMANT (Domonkos, 2015) tackle the homogenisation of daily precipitation data based on a procedure in accordance with the multiplicative (or cumulative) model that is assumed for monthly data. Moreover, the theoretical minimum number of well-correlated reference stations needed for the application of relative homogenisation methods as recommended by the World Meteorological Organisation is three, with at least four reference stations required to obtain good results in more complex situations (WMO, 2020). Obviously, this is challenging for records that extend into the 1800s, given decreasing network density.

Our study represents a relatively rare attempt to detect and correct inhomogeneities in daily precipitation data for long-term records spanning more than a century. We employed the RHtests homogenisation software as it provides one of the few methodologies for the detection and adjustments of breaks in daily precipitation data without the use of a reference series. Reeves et al. (2007) found that two-phase regression methods, as implemented in the RHtests software, had a comparable level of performance to methods such as the standard normal homogeneity test (SNHT), with the performance of the algorithm dependant on the parameters defined by the user. In total, 22 breaks were identified by RHtests software in 16 of the 36 stations, with approximately 80% of detected breakpoints supported by metadata relating to instrument and site changes and were adjusted using quantile mapping. However, subsequent assessment of residual breakpoints in PRCPTOT and R5mm using the Pettitt and SHNT tests resulted in six stations being excluded from the analysis of trends. For four of these stations (Abbeyfeale, Carndolla, Kenmare and Westport) no significant breakpoints were detected by RHtests, despite the fact that potential breakpoints documented in metadata were manually entered and tested.

For Drumnsa station, two significant residual breakpoints were detected; the first in 1917 had no associated metadata but was also detected in a previous study by Noone et al. (2015) using HOMER

(Homogenisation Software in R) software. The second breakpoint detected in 1942 corresponds to the installation of a new gauge at the time. While this break was detected by RHtests and corrected, discontinuities evident in the adjusted series suggest that the adjustment applied was too modest. The final station to be excluded was Creeslough. This series was corrected for a breakpoint detected in 1928, however, a visual assessment of the annual series following adjustment suggests the presence of a remaining shift in the early 1960's that was not detected by RHtests. Given these uncertainties surrounding break detection and correction we suggest that other studies also examine output from RHtests algorithms for residual breaks, as was also done by Santos et al. (2014) in their assessment of Portuguese seasonal extremes.

While the nature of the network analysed required the use of absolute homogeneity tests, our analysis built in assessments of residual breakpoints and trend tests that leverage information from across the network, thereby building confidence in our results. For instance, the plotting of p-values for residual breaks across series (Figure 3) allowed comparison of the timing of potential break points across the network, helping to isolate breaks unlikely to be associated with climate variability. Furthermore, the assessment of trend persistence following Noone et al. (2015) and Murphy et al. (2016) aided the visual detection of stations for which the behaviour of trends through the record deviated significantly from other stations. In this regard, cautionary flags were raised for trend in some stations/indices where remaining inhomogeneities may exist, particularly in early records. These include Strokestown, Markree and Newport, especially for trends in CWD, PRCPTOT and R5mm.

There is much scope for further work to build on the dataset and results presented. Future work will assess changes in seasonal and monthly extremes across the derived network. Noone et al. (2015) presented a long-term monthly precipitation network for the island of Ireland. Future updates of that network should include monthly totals for the additional stations developed here. While this work has improved understanding of the changing nature of extreme precipitation on the Island of Ireland, it would not have been possible without concerted efforts at data rescue. While significant progress has been made in making pre-1940 daily precipitation available for scientific scrutiny through the PhD research of the lead author (Ryan et al 2019; Ryan et al. 2020), much data remains to be rescued and digitised for Ireland. Ongoing data rescue initiatives at Met Éireann could be directed at filling the spatial gaps in this network and improving the density of data from early stations to both assist with understanding the spatial distribution of changes and to open the possibilities for deploying relative homogenisation methods. There is also the potential to develop gridded monthly products back to at least the start of the 20th Century (Hollis et al., 2019). Lastly, given the uncertainties associated with homogenisation of daily data, future work should also evaluate the findings presented here using other homogenisation software. Insights from initiatives such as the European Research Area for

Climate Services (ERA4CS) INDECIS project that are assessing approaches to homogenisation of daily data will be critical in this regard.

Finally, with increasing concerns about the impacts of climate change, understanding changes in the characteristics of precipitation is increasingly important given that climate extremes often exhibit different behaviours than changes in average conditions. The UN Intergovernmental Panel on Climate Change (IPCC) Fifth Annual Assessment Report (AR5) states that over most of the mid-latitude land masses, extreme precipitation events are very likely to be more intense and more frequent in a warmer world (IPCC, 2014). Trenberth et al., (2003) estimate a ~7% increase in water vapour content for every 1°C rise in temperature. Our findings are largely consistent with expected changes in daily precipitation extremes in a warming world. The overall tendency of increasing trends observed here suggest that heavy or extreme precipitation events are contributing significantly more to annual precipitation totals in Ireland, while precipitation is becoming more intense, particularly in the east and southeast of the country. However, potential changes in North Atlantic atmospheric circulation patterns will also have a direct impact on future precipitation, with changes in the frequency and intensity of Atlantic depressions and their associated weather fronts and convective activity important factors for extreme precipitation (Matthews et al., 2015). Therefore, it is important to further investigate the driving mechanisms behind the trends reported.

5. Conclusion

This study examines trends in annual precipitation for Ireland using the ETCCDI indices for a newly developed long-term daily precipitation network. Stations within the network have an average record length of 118 years, with some stations dating back to the 1870s. Our results show the predominance of increasing trends across the island, the growing contribution of heavy precipitation events to annual totals and increases in precipitation intensity, especially in the east and southeast of the island. While results are consistent with expectations in a warming world, the influence of low frequency variability is substantial. Understanding variability and change in long-term climate is dependent on data rescue activities and approaches to homogenisation of daily data. While approaches to undertake the latter are still in their infancy, this work shows the benefits of combining existing techniques to increase confidence in the derived series.

ORCID
Ciara Ryan https://orcid.org/0000-0001-6281-5123
Conor Murphy https://orcid.org/0000-0003-4891-2650
Mary Curley https://orcid.org/0000-0002-6209-8221

References

Aguilar, E. and Prohom, M. (2011) RClimDex-extraQC (EXTRAQC Quality Control Software). User Manual, Centre for Climate Change, University Rovira i Virgili, Tarragona, Spain. http://www.c3.urv.cat/data/manual/Manual_rclim dex_extraQC.r.pdf (accessed 14th November 2020).

Alexander, L.V., Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Klein Tank, A.M.G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., RupaKumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D.B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., Vazquez-Aguirre, J.L. (2006) Global observed changes in daily climate extremes of temperature and precipitation. J Geophys Res 111: D05109. https://doi.org/10.1029/2005jd006290

Alexandersson, H. (1986) A homogeneity test applied to precipitation data. J Climatol 6:661–675. doi:10.1002/joc.3370060607

Ashcroft, L., Coll, J.R., Gilabert, A., Domonkos, P., Brunet, M. and Aguilar, E. (2018) A rescued dataset of sub-daily meteorological observations for Europe and the southern Mediterranean region, 1877–2012. Earth System Science Data, 10(3), 1613–1635.

Auer, I., Böhm, R., Jurković , A., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Brunetti, M., Nanni, T., Maugeri, M., Briffa, K., Jones, P., Efthymiadis, D., Mestre, O., Moisselin, JM., Begert, M., Brazdil, R., Bochníček, O., Cegnar, T., Gajić-Čapka, M., Zaninović, K., Majstorović, Ž., Szalai, S., Szentimery, T., Mercalli, L. (2005) A new instrumental precipitation dataset for the Greater Alpine Region for the period 1800–2002. Int. J. Climatol. 25(2): 139–166.

Brunet, M., Gilabert, A., Jones, P., and Efthymiadis, D. (2014) A historical surface climate dataset from station observations in Mediterranean North Africa and Middle East areas, Geosci. Data J., 1, 121–128, https://doi.org/10.1002/gdj3.12

Brunet, M., and Jones, P. (2011) Data rescue initiatives: Bringing historical climate data into the 21st century. Climate Res., 47, 29–40, https://doi.org/10 .3354/cr00960.

Butler, C.J., Coughlin, A.D.S. and Fee, D.T. (1998) Precipitation at Armagh Observatory 1838–1997. Biol. Environ. 98B(2): 123–140.

Coll, J., Curley, C., Walsh, S. and Sweeney, J., (2018) HOMERUN: Relative Homogenisation of the Irish Precipitation Network. Environmental Protection Agency, Johnstown Castle Estate: Wexford, Ireland Report No. 242. EPA RESEARCH PROGRAMME 2014–2020.

De Jongh, I.L.M., Verhoest, N.E.C. and De Troch, F.P. (2006) Analysis of 105-year time series of precipitation observed at Uccle, Belgium. Int. J. Climatol. 26: 2023–2039.

Domonkos, P. (2015) Homogenization of precipitation time series with ACMANT. Theoretical and Applied Climatology, 122, 303–314. http://doi.org/10.1007/s00704-014-1298-5.

Ducre-Robitaille, J.F., Vincent, L.A., Boulet, G. (2003) Comparison of techniques for detection of discontinuities in temperature series. International Journal of Climatology 23: 1087–1101.

Eischeid, J.K., Baker, C.B., Karl, T.R. and Diaz, H.F. (1995) The quality control of long-term climatological data using objective data analysis. Journal of Applied Meteorology 34: 2787–2795.

Espírito Santo, F., Ramos, A.M., de Lima, M.I.P. and Trigo, R.M. (2014). Seasonal changes in daily precipitation extremes in mainland Portugal from 1941 to 2007. *Reg Environ Change* 14**,** 1765–1788.

Freitas, L., Gonzalez Pereira, M., Caramelo, L., Mendes, M. and Nunes, L.F., (2013) Homogeneity of monthly air temperature in Portugal with HOMER and MASH. *Quarterly Journal of the Hungarian Meteorological Service,* 117: 69–90.

Gubler, S., Hunziker, S., Begert, M., Croci-Maspoli, M., Konzelmann, T., Brönnimann, Schwierz, C., Oria, C. and Rosas, G., (2017) The influence of station density on climate data homogenization. *International Journal of Climatology,* 37: 4670–4683.

Hamed, K.H. and Rao, A.R., (1998) A modified Mann-Kendall trend test for autocorrelated data. Journal of Hydrology, 204, 182–196.

Hollis, D., McCarthy, M.P., Kendon, M., Legg, T. and Simpson, I. (2019) HadUK-Grid – a new UK dataset of gridded climate observations. *Geoscience Data Journal.* 6: 151– 159. https://doi.org/10.1002/gdj3.78

Hurrell, J. W. (1995) Decadal Trends in the North Atlantic Oscillation: Regional Temperatures and Precipitation, *Science*, 269(5224), 676–679, doi:10.1126/science.269.5224.676.

IPCC (2014) *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Kiely, G. (1999) Climate change in Ireland from precipitation and streamflow observations, *Advances in Water Resources*, 23(2), 141–151, doi:10.1016/S0309-1708(99)00018-4.

Leahy, P. G. and Kiely, G. (2011) Short Duration Rainfall Extremes in Ireland: Influence of Climatic Variability, *Water Resour Manage*, 25(3), 987–1003, doi:10.1007/s11269-010- 9737-2.

Llabrés-Brustenga, A., Rius, A., Rodríguez-Solà, R. *et al.* (2019) Quality control process of the daily rainfall series available in Catalonia from 1855 to the present. *Theor Appl Climatol* 137**,** 2715–2729. https://doi.org/10.1007/s00704-019-02772-5

Matthews, T., Murphy, C., Wilby, R.L., and Harrigan, S. (2015) A cyclone climatology of the British-Irish Isles 1871–2012. *Int. J. Climatol.*, doi: 10.1002/joc.4425.

McElwain, L. and Sweeney, J. (2007) Key Meteorological Indicators of Climate Change in Ireland. Environmental Protection Agency, Johnstown Castle Estate: Wexford, Ireland.

Moberg, A., Jones, P.D., Lister, D., Walther, A., Brunet, M., Jacobeit, J., Alexander, L.V., Della-Marta, P.M., Luterbacher, J., Yiou, P., Chen, D., Klein Tank, A.M.G., Saladié, O., Sigró, J., Aguilar, E., Alexandersson, H., Almarza, C., Auer, I., Barriendos, M., Begert, M., Bergström, H., Böhm, R., Butler, C.J., Caesar, J., Drebs, A., Founda, D., Gerstengarbe, F-W., Micela, G., Maugeri, M., Österle, H., Pandzic, K., Petrakis, M., Srnec, L., Tolasz, R., Tuomenvirta, H., Werner, P.C., Linderholm, H., Philipp, A., Wanner, H. and Xoplaki, E. (2006) Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901–2000. *J. Geophys. Res.* 111: D22106.

Murphy, C., Broderick, C., Burt, T.P., Curley, M., Duffy, C., Hall, J. *et al.* (2018) A 305-year continuous monthly rainfall series for the island of Ireland (1711–2016). *Climate of the Past*, **14**, 413– 440.

Murphy, C., Harrigan, S., Hall, J. and Wilby, R. L. (2013) Climate-driven trends in mean and high flows from a network of reference stations in Ireland, *Hydrological Sciences Journal*, 58(4), 755–772, doi:10.1080/02626667.2013.782407.

Noone, S., Murphy, C., Coll, J., Matthews, T., Mullan, D., Wilby, R.L. and Walsh, S. (2015) Homogenization and analysis of an expanded long-term monthly rainfall network for the Island of Ireland (1850–2010). International Journal of Climatology, 36, 2837–2853. https://doi.org/10.1002/joc.4522

Pettitt, A.N. (1979) A non-parametric approach to the change-point problem. *J. Roy. Stat. Soc. C-App.*, 28, 126-135, 1979.

Reeves, J., J. Chen, X. L. Wang, R. Lund, and Q. Lu (2007), A review and comparison of changepoint detection techniques for climate data, *J. Appl. Meteorol. Climatol.*, **46**, 900–915.

Ryan, C., Duffy, C., Broderick, C., Thorne, P.W., Curley, M., Walsh, S., Treanor, M., Daly, C. *and Murphy, C.*. (2018) Integrating data rescue into the classroom. *Bulletin of the American Meteorological Society*, 99(9), 1757– 1764. https://doi.org/10.1175/BAMS-D17-0147.1

Ryan, C., Murphy, C., McGovern, R., Curley, M., Walsh, S. and 476 students (2020) Ireland's pre-1940 daily rainfall records. *Geoscience Data Journal*, 1–13. https://doi.org/10.1002/gdj3.1

Sen, P. K. (1968) Estimates of the regression coefficient based on Kendall's tau, *J. Am. Stat. Assoc*., 63, 1379-1389.

Sheridan, T. (2001) Analysis of trends at some Irish rainfall stations. Dublin, Ireland: Met Éireann Technical Report 59.

Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., Kelder, T., Kowal, K., Lees, T., Matthews, T., Murphy, C., and Wilby, R. L. (2020) Nonstationary weather and water extremes: a review of methods for their detection, attribution, and management, *Hydrol. Earth Syst. Sci. Discuss*. [preprint], https://doi.org/10.5194/hess-2020-576.

Szentimrey, T. (1999) Multiple analysis of series for homogenization (MASH). In: Szalai, S., Szentimrey, T., Szinell, C. (eds.) Proceedings of the Second Seminar for Homogenization of Surface Climatological Data. Seminar for Homogenization of Surface Climatological Data, 2nd session (9–13 November 1998), Budapest, Hungary, WCDMP-No. 41, pp. 27–46. World Meteorological Organization (WMO).

Szentimrey, T. (2014) Manual of homogenization software MASHv3.03. In. Hungarian Meteorological Service, Budapest, Hungary.

Theil, H. (1950) A Rank-invariant Method of Linear and Polynomial Regression Analysis, I., Nederlands Akad. Wetensch. Proc., 53, 386-392.

Toreti, A., Kuglitsch, F.G., Xoplaki, E., Della-Marta, P.M., Aguilar, E., Prohom, M. and Luterbacher, J. (2011), A note on the use of the standard normal homogeneity test to detect inhomogeneities in climatic time series, Int. J. Climatol., 31, 630–632, doi:10.1002/joc.2088.

Trenberth, K. E. (2011) Changes in precipitation with climate change, *Clim Res*, 47(1-2), 123– 138. https://doi.org/10.3354/cr00953.

Venema, V., Trewin, B., Wang, X., Szentimrey, T., Lakatos, M., Aguilar, E., Auer, I., Guijarro, J. A., Menne, M., Oria, C., Louamba, W.S.R.L., Rasul, G. (2018) Guidance on the homogenization of climate station data. WMO, Geneva. https://doi.org/10.31223/osf.io/8qzrf

Venema, V.K.C., Mestre, O., Aguilar, E., Auer, I., Guijarro, J.A., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P., Viarre, J., Müller-Westermeier, G., Lakatos, M., Williams, C.N., Menne, M.J., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andresen, L., Acquaotta, F., Fratianni, S., Cheval, S., Klancar, M., Brunetti, M., Gruber, C., Prohom Duran, M., Likso, T., Esteban, P., Brandsma, T. (2012) Benchmarking homogenization algorithms for monthly data. *Clim. Past* 8(1), 89–115.

Vertačnik, G., Dolinar, M., Bertalanič, R., Klančar, M., Dvoršek, D. and Nadbath, M., 2015. Ensemble homogenization of Slovenian monthly air temperature series. *International Journal of Climatology*, 35: 4015–4026.

Walsh, S. (2016) Long-term rainfall averages for Ireland. Met Éireann Climatological Note 15. Dublin.

Walsh, S. and Dwyer, N. (2012) Rainfall. In: The Status of Ireland's Climate, 2012. Dwyer N (ed). CCRP Report No. 26. EPA, Wexford, Ireland.

Wang, X. L., H. Chen, Y. Wu, Y. Feng, and Q. Pu, (2010) New techniques for detection and adjustment of shifts in daily precipitation data series. *J. Appl. Meteor. Climatol*., 49 (No. 12), 2416-2436. DOI: 10.1175/2010JAMC2376.1

Wang, X. L., (2008a) Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal t or F test. *J. Appl. Meteor. Climatol*., 47, 2423-2444.

Wang, X. L., (2008b) Penalized maximal F-test for detecting undocumented mean-shifts without trend-change. *J. Atmos. Oceanic Tech*., 25 (No. 3), 368-384. doi:10.1175/2007/JTECHA982.1.

Wang, X. L., Q. H. Wen, and Y. Wu, (2007) Penalized maximal t test for detecting undocumented mean change in climate data series. *J. Appl. Meteor. Climatol*., 46 (No. 6), 916-931. DOI:10.1175/JAM2504.1

Wang, X. L. and Y. Feng (2013) RHtests_dlyPrcp User Manual. Climate Research Division, Atmospheric Science and Technology Directorate, Science and Technology Branch, Environment Canada. 17 pp. [Available online at http://etccdi.pacificclimate.org/software.shtml]

Wilby, R. L., Clifford, N. J., De Luca, P., Harrigan, S., Hillier, J. K., Hodgkins, R., Johnson, M. F., Matthews, T. K., Murphy, C., Noone, S. J., Parry, S., Prudhomme, C., Rice, S., Slater, L., Smith, K. and Wood, P. (2017) The 'dirty dozen'of freshwater science: detecting then reconciling hydrological data biases and errors, Wiley Interdisciplinary Reviews (WIREs) Water, 4 (3), e1209.

World Meteorological Organization (2020) *Guidelines on climate metadata and homogenization* (WMO/TD-No. 1186, WCDMP No. 53). Geneva.

World Meteorological Organization (2016) *Guidelines on Best Practices for Climate Data Rescue 2016* (WMO-No. 1182). Geneva.

Yue, S. and Hashino, M. (2007) Long term trends of annual and monthly precipitation in Japan. *Journal of the American Water Resources Association*, Vol 39 (3), 587-596.

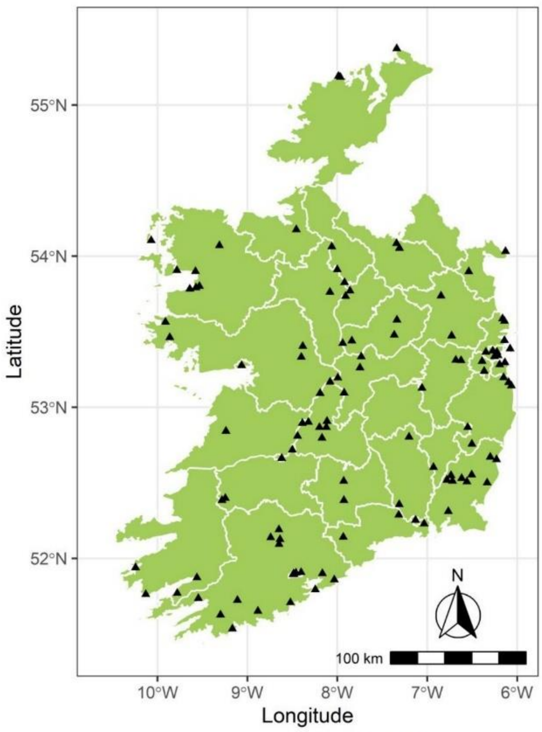


Figure 1: Location of stations for which pre-1940 daily precipitation data were rescued and digitised. Individual folders containing a data file and metadata file are available for each station (Ryan et al. 2020).

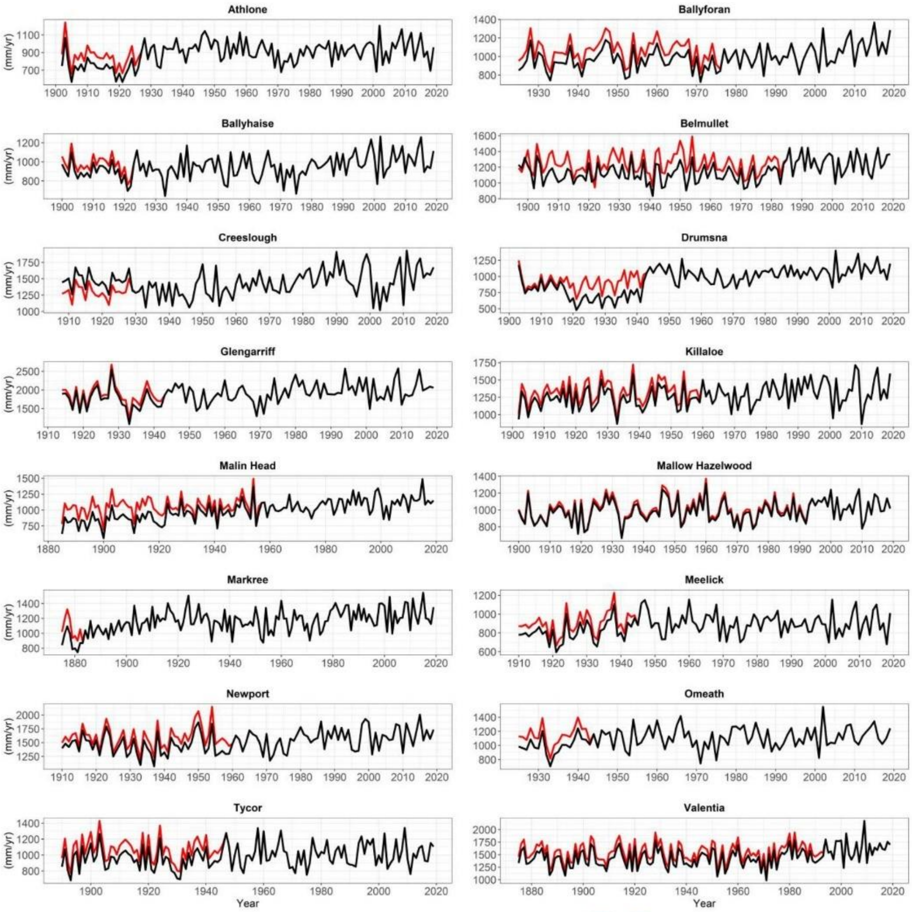


Figure 2: Original (black) and adjusted (red) total annual precipitation for the 16 station series with breakpoints detected and adjusted using RHtests software. Units are millimetres per year.

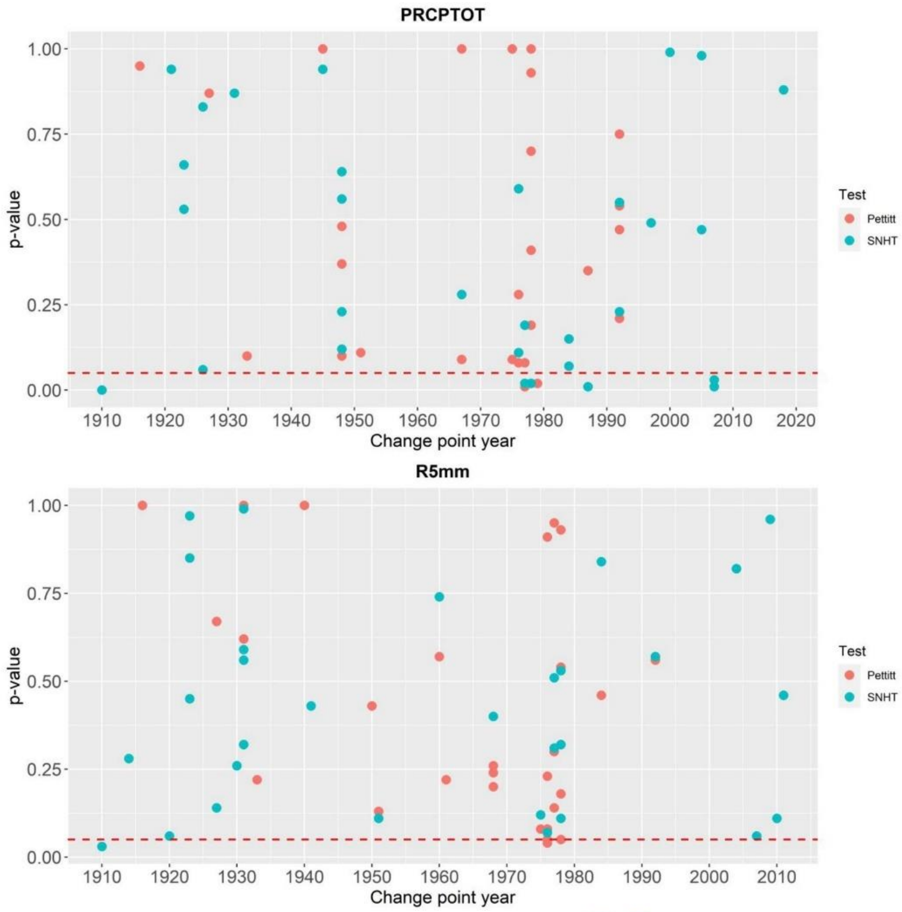


Figure 3: P-values of potential breakpoints detected in the annual series of PRCPTOT (top) and R5mm (bottom) indices using the Pettitt test and the standard normal homogeneity test (SNHT). Red dashed line indicates significance at the 0.05 level.

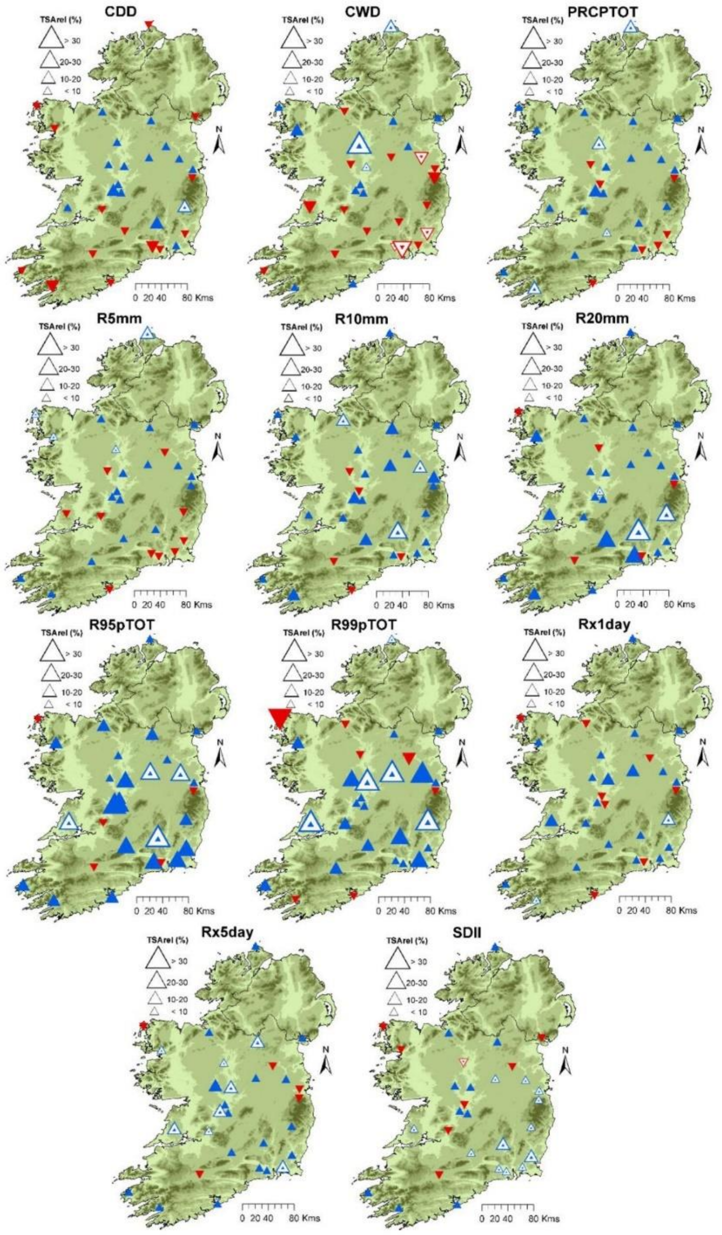


Figure 4: Magnitude and direction of trends for 1910-2019 fixed period for extreme precipitation indices. Blue triangles represent increasing trends and red decreasing trends, with magnitude proportional to size. Magnitude derived using the relative Theil-Sen Approach TSArel (%). Significant trends (5 % level) shown by white triangles and derived from the Modified Mann Kendall (MKZs).

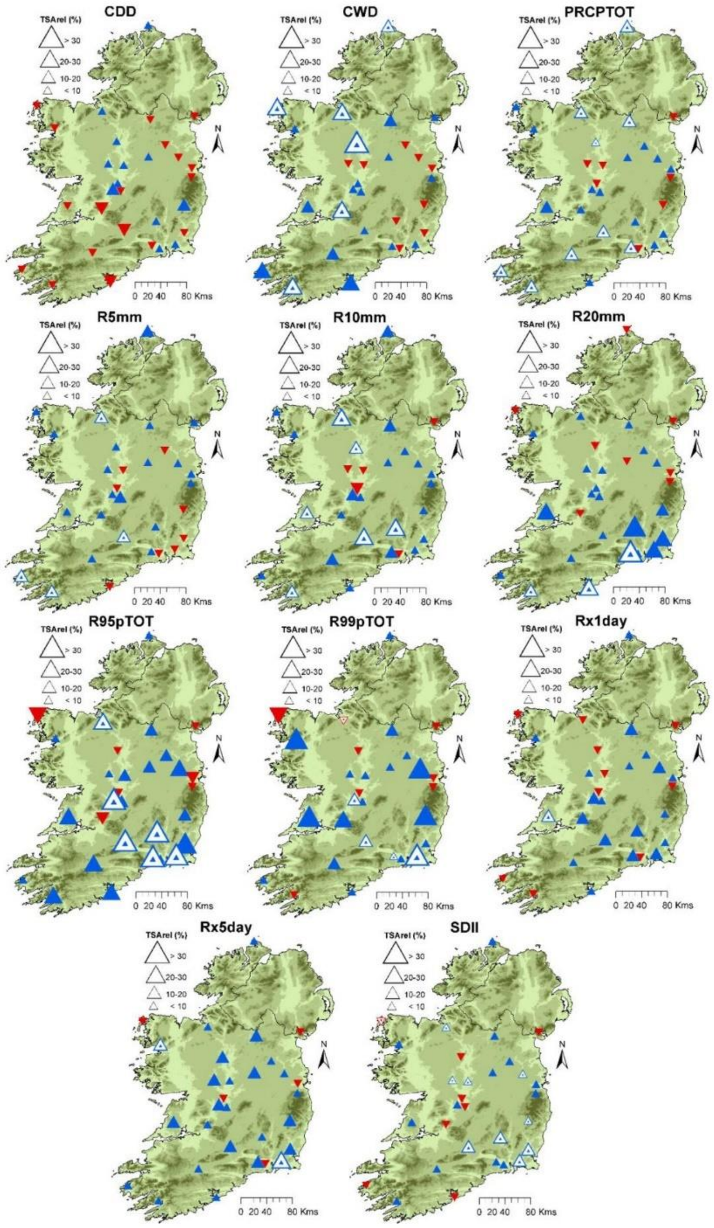


Figure 5: Magnitude and direction of trends for 1940-2019 fixed period for extreme precipitation indices. Blue triangles represent increasing trends and red decreasing trends, with magnitude proportional to size. Magnitude derived using the relative Theil-Sen Approach TSArel (%). Significant trends (5 % level) shown by white triangles and derived from the Modified Mann Kendall (MKZs).

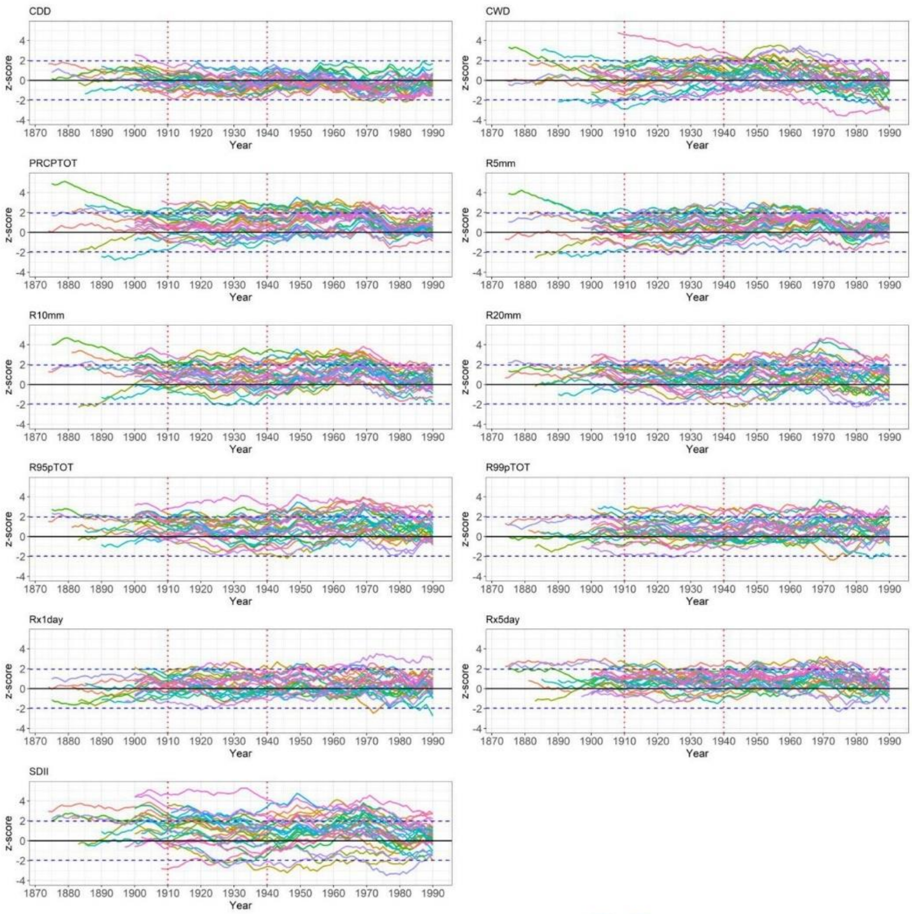


Figure 6: Persistence plots for full available time-series of extreme precipitation indices. Coloured lines represent MKZs statistics for varying start years for individual stations. Dashed blue lines are the threshold for statistically significant trends at the 5 % level with MKZs above (below) these indicating significant increasing (decreasing) trends since the corresponding start date. The vertical red lines mark the start year of the 1910-2019 and 1940-2019 fixed periods.

Table 1: Summary of Quality Control tests applied to the pre-1940 data. * Tests applied using RClimdex_extraQC software (Aguilar & Prohom, 2011).

QC test	Description
Monthly cross-check	Flag values when sum of daily rr data does not equal monthly rr total
Non-numeric values	Check file structure
Physically impossible values	Flag values where rr <0
Duplicate date control*	Flag any sequence of dates that appear more than once in time series. RClimdex_extraQC software (Aguilar & Prohom, 2011).
Unrealistic values*	Flag out of range values based on fixed threshold of 243.5mm i.e., highest daily total recorded in Ireland 18th Sep 1993 at Cloone, Co.Kerry. RClimdex_extraQC software (Aguilar & Prohom, 2011).
Anomalous sequence of zero rr	Flag if zeros persist for ≥ 1-month duration
Rounding problems evaluation*	Visual assessment to identify potential biases in the rounding of decimal values. RClimdex_extraQC software (Aguilar & Prohom, 2011).
Accumulated values	Flag months with only one value recorded which is in excess of 2 times the mean daily rr intensity for that month
Climatological outliers*	Flag outliers based on IQR exceedance: percentile-based approach and therefore suitable for asymmetric distribution. RClimdex_extraQC software (Aguilar & Prohom, 2011).
Temporal consistency	Flag outliers based on limits determined from a multiple of the interquartile range (IQR) calculated for each station/month. Outliers are identified using the sample distribution of each calendar month separately for each station. An outlier is flagged when $X_i - {}_q50 > f(IR)$ where X_i is the monthly mean of the year i, ${}_q50$ is the median, and f is the multiplication factor. A value of f = 4 was used as the multiplicative factor as in Eischeid et al. (1995).
Spatial consistency	Check consistency with neighbouring stations: Number of dry days: flag months with 2 days more/less than max/min of nearest neighbour.

	Number of wet days (rr >= 0.2mm): flag months with 2 days more/less than max/min of nearest neighbour. Number of days with >= 5mm: flag months with 2 days more/less than max/min of nearest neighbour. Number of days with >=10mm: flag months with 2 days more or less than min/max nearest neighbour. Total Monthly Fall: (normalised as a fraction of the long-term average) flag months with ± 25% of the max/min normalised rr of nearest neighbours i.e., > 1.25 max of NN or < 0.75 min of NN.

Table 2: Details of selected long-term stations including station ID, station name, location, height (m) above msl, start and end year of series, series category.

st_id	station_name	lat	lon	height	start	end	group
108	Foulkesmill (Longraigue)	52.31056	-6.76639	71	1874	2019	A
175	Phoenix Park	53.36361	-6.34972	48	1881	2019	A
417	Inagh (Mt.Callan)	52.84167	-9.23833	122	1908	2019	A
1075	Roches Point	51.79306	-8.24444	40	1883	2019	A
1275	Markree	54.175	-8.45556	34	1875	2019	A
1519	Meelick (Victoria Lock)	53.16667	-8.08056	39	1910	2019	A
1529	Drumsna (Albert Lock)	53.91111	-8.00000	45	1903	2019	A
1575	Malin Head	55.37194	-7.33917	20	1885	2019	A
2375	Belmullet	54.22750	-10.00694	9	1897	2019	A
1819	Portumna O.P.W.	53.09167	-8.19167	35	1929	2019	A
1929	Athlone O.P.W.	53.42222	-7.94167	37	1902	2019	A
2012	Cashel (Ballinamona)	52.51056	-7.92861	80	1911	2019	A
2275	Valentia Observatory	51.93833	-10.24083	24	1875	2019	A
1812	Waterford (Tycor)	52.25278	-7.13056	49	1890	2019	A
201	Glengarriff (Ilnacullin)	51.73472	-9.54583	7	1914	2019	A
603	Kenmare (Dereen)	51.76944	-9.78056	24	1912	2019	A

1923	Glenasmole D.C.W.W.	53.23889	-6.36667	158	1900	2019	A
8212	Portlaw (Mayfield II)	52.29083	-7.30083	8	1900	2019	A
675	Ballyhaise	54.05139	-7.30972	78	1900	2019	B
706	Mallow (Hazelwood)	52.19028	-8.65000	94	1900	2019	B
875	Mullingar	53.53722	-7.36222	101	1910	2019	B
944	Creeslough (Carrownamaddy)	55.13333	-7.95000	88	1908	2019	B
1338	Omeath	54.08667	-6.25639	12	1925	2019	B
1375	Dunsany	53.51583	-6.66000	83	1900	2019	B
1475	Gurteen	53.05306	-8.00861	75	1900	2019	B
2115	Hacketstown (Voc.Sch.)	52.86111	-6.55278	189	1918	2019	B
2528	Ballyforan (Bord na Mona)	53.43972	-8.30333	47	1925	2019	B
4015	Enniscorthy (Brownswood)	52.46306	-6.56083	18	1900	2019	B
4513	Kilkenny (Lavistown House) II	52.63611	-7.19722	58	1900	2019	B
5131	Kilskyre (Robinstown)	53.69278	-6.96333	87	1900	2019	B
6019	Killaloe Docks	52.81	-8.44889	40	1902	2019	B
6329	Strokestown (Carrowclogher)	53.753	-8.10800	52	1908	2019	B
3310	Abbeyfeale (Caherlane)	52.35194	-9.28361	155	1925	2019	B
1175	Newport	53.92222	-9.57222	22	1910	2019	B
1433	Westport (Carrabawn)	53.7925	-9.52722	56	1909	2019	B
2227	Carndolla	53.40306	-9.01556	24	1900	2019	B

Table 3: List of extreme precipitation indices used in this study. All indices are calculated annually.

ID	Indicator name	Indicator definitions	Units
RX1day	Max 1-day precipitation amount	Annual maximum 1-day precipitation	mm
RX5day	Max 5-day precipitation amount	Annual maximum consecutive 5-day precipitation	mm
SDII	Simple daily intensity index	Ratio of annual total precipitation to number of wet days (≥1 mm)	mm/day
R5mm	Number of days when precipitation ≥5 mm	Annual count when precipitation ≥5 mm	days
R10mm	Number of heavy precipitation days	Annual count when precipitation ≥10 mm	days

R20mm	Number of very heavy precipitation days	Annual count when precipitation ≥ 20 mm	days
CDD	Consecutive dry days	Maximum number of consecutive days when precipitation <1 mm	days
CWD	Consecutive wet days	Maximum number of consecutive days when precipitation ≥1 mm	days
R95pTOT	Very wet days	Annual sum of precipitation on days when precipitation exceeds the 95th percentile of daily precipitation in the base period.	mm
R99pTOT	Extremely wet days	Annual sum of precipitation on days when precipitation exceeds the 99th percentile of daily precipitation in the base period.	mm
PRCPTOT	Annual total wet-day precipitation	Annual total precipitation from days ≥ 1 mm	mm

Table 4: List of breakpoints detected by RHtests software and supporting information extracted from metadata files.

ID	Station	Breakpoint	Reason
1929	Athlone	1926	Reports of gauge leaking resulting in low readings. Recommendation for new gauge to be installed, however, no documentation to say that this was carried out. The height of the gauge above ground was changed from 2ft to 1ft in Jan 1928.
2528	Ballyforan	1976	Ballyforan station commenced observations in 1976. Record extended to 1925 using nearby station data.
675	Ballyhaise	1923	Ballyhaise station commenced observations in 1923. Record extended to 1900 using nearby station data.
2375	Belmullet	1956; 1984	New station established in September 1956. No documented reason for break detected in 1984. Significant (prob >0.99) changepoint detected in 1983 by Leahy and Kiely (2011) using the Mann-Pettitt-Whitney test on annual precipitation total over the period 1957-2008.
944	Creeslough	1928	Creeslough record was extended to 1908 using nearby station data. Dunfanaghy is among the neighbouring stations available in the pre-1940 period. Dunfanaghy was established as a rainfall station in 1880, the first location being St. Helen's (55°11'N; 7°58'W). In July 1926, the station was moved to St. Patrick's, a site which had been set up approx. a quarter of a mile from St. Helen's. The record at St. Patrick's was considered a continuation of the observations taken at St. Patrick's.

1529	Drumsna	1917; 1942	No documented reason for break detected in 1917 but this break was also detected by HOMER in previous work by Noone et al. (2015). New gauge installed in 1942. Comparison of the old site over a 13-month period showed the new gauge recording 154% of the old site.
201	Glengarriff	1931; 1942	In 1929 the height of the gauge changed from 15ft to 25ft above msl. The height of the gauge above ground also changed from 1 ft to 2ft 6in at some point during the period 1929-1936. In December 1942 the site was moved to a new site. The site was inspected in 1950 and the report states that the new site satisfies conventions.
6019	Killaloe	1959	Killaloe station was established in 1983. Record extended to 1900 using nearby station data. Killaloe OPW and Killaloe (Ballina) are among the neighbouring stations available prior to 1983. Killaloe OPW commenced in 1957 and ceased observations in 1984. Killaloe (Ballina) commenced in 1866 and ceased observations in 1959.
1575	Malin Head	1923; 1957	From 1885 readings were made at the telegraphic reporting station (Lloyds tower at a height of 230 ft). In Dec 1921 the station was moved to the coast guard station at a height of approximately 22ft above msl. New station established May 1955. Following the installation of the new site, comparison readings were taken which suggested discrepancies between readings taken at the old site and the newly established station.
706	Mallow	1993	New rain gauge installed and moved to new site due to concerns about trees that had grown high around the site.
1275	Markree	1883	The original square gauge (1 sq. yard) was positioned on top of the library. Over the period 1875-1881 a comparison 5" gauge positioned 6" above ground gave a correction multiple of 1.2045. In 1870 the diameter of the gauge is recorded as 24 inches. In 1883 the diameter of the gauge is recorded as 5 inches. No record of the gauge diameter is noted in the intervening years. In July 1884 the height of the top of the gauge above ground changed from 16.5 ft to 1 ft. In July 1884 the height of the top of the gauge above sea level changed from 148 ft to 130ft.
1519	Meelick	1944	No documented reason for break detected in 1944. Station inspection report states that no change of site took place prior to 1952. However, a report in 1942 documents concerns about over-exposure. It states that the ground around the gauge has been hallowed out saucer-like due to over-exposure. It states that the rain recorder is now accurate and recommends that a standard gauge be issued.

1175	Newport	1928; 1952; 1959	Newport TUCSON AWS was established in 2005. Prior to this, readings were taken at Newport (Furnace) which commenced observations in 1959. The record was extended to 1910 using available nearby station data.
1338	Omeath	1943	Omeath station commenced observations in March 1943. Record extended to 1925 using nearby station data.
1812	Tycor	1946	In March 1946 a rain-recorder and 5" Snowdon gauge (E.S.B. type) were installed on a new site further East (50 yds) and 10ft. lower than the existing gauge - which was over-exposed on the concrete roof of the service reservoir. Comparative readings were taken for a period of 2.5 years and the results showed that the ratio of the readings of the new gauge to those of the old gauge was 115:110. The new gauge was forthwith adopted as the official gauge.
2275	Valentia	1993	April 1993: trees removed 150m south of enclosure.

Table 5: Mean/min/max annual % difference between the original and the adjusted series for 16 inhomogeneous station series.

station_id	station_name	mean (%)	min (%)	max (%)
201	Glengarriff	7.4	4.8	11.5
675	Ballyhaise	8.4	0.1	9.6
706	Mallow_HW	2.7	0.2	4.4
944*	Creeslough	-11.3	-12.5	1.3
1175	Newport	10.2	5.9	17.8
1275	Markree	18.4	1.9	21.2
1338	Omeath	14	3.5	16.4
1519	Meelick	11.2	3	12.5
1529*	Drumsna	24.4	4.5	38.9
1575	Malin Head	17.2	8	26.6
1812	Tycor	12.8	2.1	13.9
1929	Athlone	16.8	15.4	18.4
2275	Valentia	8.8	5.5	9.3
2375	Belmullet	11.7	-9.9	25.6
2528	Ballyforan	11.6	2.3	13.1
6019	Killaloe	8.4	3.7	9.1

Table 6: Direction of change and proportion of statistically significant (5 % level) trends for 1910-2019 fixed period extreme precipitation indices. Direction and significance tested using Modified

MannKendall (MKZs) and magnitude tested with the relative Theil-Sen Approach (TSArel). Magnitude of change is based on the median of the test statistics with confidence intervals given by the lower/upper bounds. Number of days is the medium numbers of days per year across all stations with confidence intervals given by the lower/upper bounds.

Indicator	Stat.	Positive (sig.) %	Negative (sig.) %	Stat.	Magnitude (CI) %	Stat.	No. of days/yr (CI)
CDD	MKZs	53.3 (3.3)	46.7 (0.0)	TSA_{rel}	0(-16.82,17.41)	TSA	0(-0.03,0.03)
CWD	MKZs	43.3 (10)	56.7 (13.3)	TSA_{rel}	0(-19.01,15.08)	TSA	0(-0.02,0.02)
PRCPTOT	MKZs	76.7 (13.3)	23.3 (0.0)	TSA_{rel}	2.79(-6.25,11.39)	TSA	NA
R95pTOT	MKZs	83.3 (13.3)	16.7 (0.0)	TSA_{rel}	13.72(-9.3,37.29)	TSA	NA
R99pTOT	MKZs	76.7 (16.7)	23.3 (0.0)	TSA_{rel}	5.07(-10.2,61.87)	TSA	NA
Rx1day	MKZs	73.3 (6.7)	26.7 (0.0)	TSA_{rel}	3.64(-11.27,19.41)	TSA	NA
Rx5day	MKZs	83.3 (26.7)	16.7 (0.0)	TSA_{rel}	6.06(-6.96,19.23)	TSA	NA
SDII	MKZs	73.3 (40)	26.7 (3.3)	TSA_{rel}	3.84(-2.02,9.75)	TSA	NA
R5mm	MKZs	66.7 (13.3)	33.3 (0.0)	TSA_{rel}	0(-8.47,11.64)	TSA	0(-0.05,0.08)
R10mm	MKZs	83.3 (10)	16.7 (0.0)	TSA_{rel}	6.22(-6.46,20.54)	TSA	0.02(-0.02,0.06)
R20mm	MKZs	83.3 (10)	16.7 (0.0)	TSA_{rel}	0(0,38.25)	TSA	0(0,0.02)

Table 7: Direction of change and proportion of statistically significant (5 % level) trends for 1940-2019 fixed period extreme precipitation indices. Direction and significance tested using Modified MannKendall (MKZs) and magnitude tested with the relative Theil-Sen Approach (TSArel). Magnitude of change is based on the median of the test statistics with confidence intervals given by the lower/upper bounds. Number of days is the medium numbers of days per year across all stations with confidence intervals given by the lower/upper bounds.

Indicator	Stat.	Positive (sig.) %	Negative (sig.) %	Stat.	Magnitude (CI) %	Stat.	No. of days/yr (CI)
CDD	MKZs	40 (0.0)	60 (0.0)	TSA_{rel}	0 (-21.13,18.48)	TSA	0(-0.05,0.04)
CWD	MKZs	70 (20)	30 (0.0)	TSA_{rel}	0 (-11.05,26.25)	TSA	0(-0.02,0.04)
PRCPTOT	MKZs	76.7 (30)	23.3 (0.0)	TSA_{rel}	6.4 (-3.94,16.53)	TSA	NA
R95pTOT	MKZs	76.7 (20)	23.3 (0.0)	TSA_{rel}	13.61 (-17.05,44.69)	TSA	NA
R99pTOT	MKZs	73.3 (13.3)	26.7 (0.0)	TSA_{rel}	0 (-30.89,66.84)	TSA	NA
Rx1day	MKZs	66.7 (3.3)	33.3 (0.0)	TSA_{rel}	6.02 (-11.71,24.22)	TSA	NA
Rx5day	MKZs	83.3 (6.7)	16.7 (0.0)	TSA_{rel}	8.35 (-6.81,23.77)	TSA	NA
SDII	MKZs	73.3 (30)	26.7 (3.3)	TSA_{rel}	3.78 (-3.8,10.88)	TSA	NA
R5mm	MKZs	73.3 (13.3)	26.7 (0.0)	TSA_{rel}	5.34 (-6.64,15.82)	TSA	0.05(-0.06,0.14)
R10mm	MKZs	83.3 (20)	16.7 (0.0)	TSA_{rel}	6.09 (-9.76,22.79)	TSA	0.02(-0.03,0.09)
R20mm	MKZs	73.3 (10)	26.7 (0.0)	TSA_{rel}	0 (0,41.04)	TSA	0(0,0.03)

Appendix III: Chapter 2 Supplementary Information

This appendix provides details of project resources included as supplementary information to the paper presented in Chapter 2. Images of sample meteorological records and transcription templates are included in Chapter 2. A copy of the student guidelines is provided below. Full project resources including transcription templates, MATLAB code and student-aids can be accessed at the following link to the article published in the Bulletin of the American Meteorological Society: https://journals.ametsoc.org/view/journals/bams/99/9/bams-d-17-0147.1.xml?tab_body=supplementary-materials

Continuous Assessment Guidelines

Data Rescue and Digitisation: Historical climatology and the power of the crowd

Overview

This assignment is a research led teaching experiment to reflect on the importance of historical climatology and to explore the power of the crowd in rescuing and transcribing long-term daily rainfall data. The assignment is designed to provide insight into the processes involved in data rescue, digitisation and quality assurance procedures that are critical to understanding past variability and change. Your work will feed into ongoing research and make a major contribution to our scientific understanding. Together we will transcribe over 1300 pre-1940 annual rainfall sheets from 45 stations across Ireland. Such an undertaking would be impossible by one individual, but we are better together! By taking small chunks of the data across all the people in the class we can achieve it. At the end of this process you will have made a telling contribution to knowledge, helping to create one of the longest running daily rainfall networks in Europe! During the last class of the module I aim to show you just how much of a contribution you, and us as a class have made.

Learning Outcomes

This assignment is closely tied to the learning outcomes of the module GY313 Climate Change. In addition, on completion you will have:

- A firsthand awareness of the do's and don'ts of working with historical climate data.
- A critical appreciation of how important data quality is to understanding past and future climate change.

- First-hand experience of the powerful contribution that citizen science can make to the study of climatology, geography and other disciplines.

What you will do

Full details on the stages and steps involved are provided below. Essentially, we want to create a digital copy of the rainfall sheets by inputting the handwritten information from the sheets into individual Excel files. To achieve this, you will be required to transcribe, and quality assure 18 annual rainfall sheets each and return these to me. You will then be required to write a short reflection on the importance of historical climatology. The continuous assessment (CA) accounts for one third of the final marks for this module. The CA marks will be divided as follows:

Data rescue and digitisation – 60 %

You will get 3.3% for each annual sheet that is transcribed correctly and returned to me (via Moodle – see below). That equates to 1% of the entire module mark per sheet! So, if you follow the instructions and take some care and pride in your work you could be guaranteed at least 60% in the CA (20% of module mark).

Critical reflection – 40%

Once the transcription process is complete you are required to write a critical reflection (max 1,200 words) entitled '*Why is the study of historical climatology important in the context of climate change?'*. It is up to you to take it where you will, but I will provide some readings for you on Moodle.

The deadline for both components is **midnight Friday 10th November.**

Data Rescue and Digitisation – the details (Please read carefully)

The data rescue and digitisation process is structured around three phases:

Phase 1 – Receiving the images and template

Phase 2 – Transcribing and quality assuring the data

Phase 3 – Returning the data

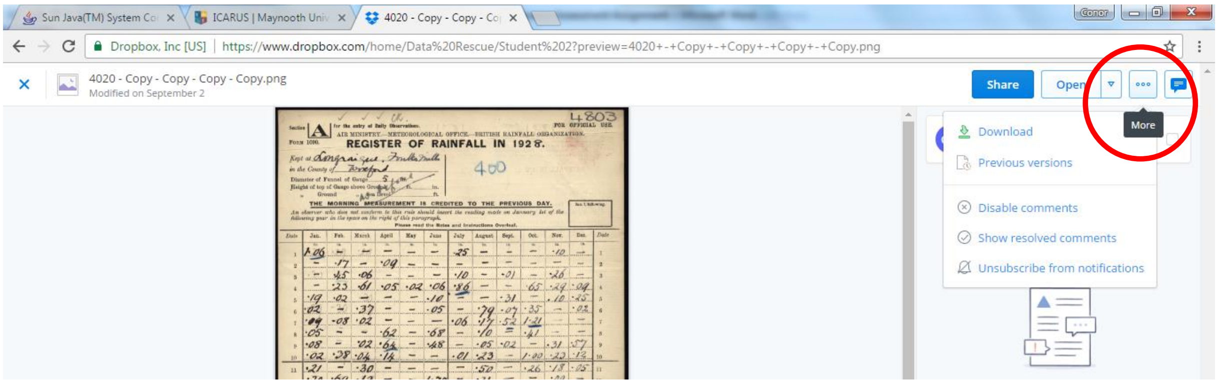
The key steps for each phase and potential pitfalls (that I can foresee) are outlined below. It is VITAL that you pay close attention to these and follow them as closely as possible.

Phase 1 – Receiving the images

You will find your images here:

https://www.dropbox.com/GY313

Press Ctrl and click on the link above. You will be brought to a Dropbox folder where you will see a list of subfolders labelled by student name. Please find your folder. In this folder you will find the 18 sheets that you will be required to transcribe. You will also find the Excel template that you will use to transcribe each sheet. Each pdf file is an image of an original rainfall sheet for a year between 1866 and 1939. Click on an image and it will open in a viewer and you can transcribe from there or you can download by clicking the more ' … ' tab in the top right corner and then clicking download (see screenshot below). You can do this for each individual sheet as you need it.



<u>KEY POINT – each image is named a specific way and it is important that you keep track of it. The naming convention is image number.pdf</u>

You now have your images and template and are ready for phase two.

Phase 2 – Digitising and quality assuring the data

The next step is to input the data and all other information recorded on the sheet into the Excel template provided in your assigned folder. Each rainfall sheet will be transcribed to a new Excel file so by the end you will have 18 Excel files to submit – each saved using the format: image number_student number.

<u>KEY POINT – once finished your transcription save the excel file as image number underscore student number</u> e.g. 1336_12345678.xls

<u>DO NOT save the EXCEL FILE as something different. We will not be able to trace it back to the student and you will not receive any marks for the sheet. If you are unsure please ask!</u>

You will need to view the video tutorial (available on Moodle) before you begin transcribing the data. In the video I use a sample rainfall sheet to demonstrate the transcription process. I describe each section of the Excel template to show you what information from the sheet goes in which part of the template form. You will have access to the video for the duration of the project so that you may watch the tutorial as many times as necessary.

TIPS for keying in data:

1. Download your 18 images and Excel template from Dropbox and save them to a GY313 work folder. Transcribe one image at a time and save the Excel file to a subfolder folder using the file name convention image number_student number.xls for each file. By the end you should have 18 Excel files in the subfolder. The folder containing the 18 images will be compressed (zip folder) and uploaded to Moodle.
2. Create a list (a simple word document that you can save in your new GY313 work folder) of the image numbers that you have been assigned. Before you submit the files you can do a quick check to make sure that the file names match the image numbers on your list.
3. Use a split screen so you can view the Excel template and the image on the screen at the same time.
4. Zoom in on the image for writing that is difficult to read.
5. It is not advisable to print out the image as it will make it more difficult to keep track of the image number. Also, you will not be able to zoom in to enhance the image.
6. N.B. If you are having difficulty reading any of the handwriting on the sheet or have any other questions, please post your query to the Moodle discussion forum that has been set up for the project (see supports below).
7. Any information transcribed should be typed <u>exactly</u> as it appears on the sheet – this includes observer names, station names and any handwritten notes.
8. Try to fill in the data working across the rows rather than down each column. We found last year that greater amounts of errors arose when students entered data by column.

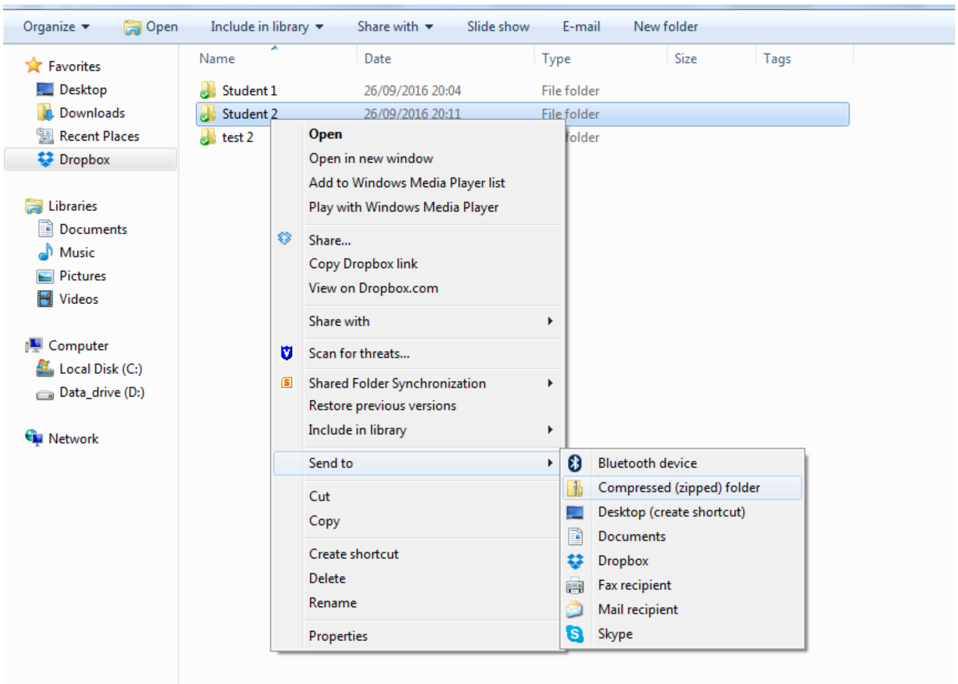
Phase 3 – Return of data

Once you have keyed in the data you need to save the Excel file as follows. **DO NOT DEVIATE** from this naming convention or you will not receive the marks for that sheet as I will not be able to trace it.

image number_student number.xlsx

For example, if you transcribe image number 1336 and your student number is 12345678 you should save the Excel file as 1336_12345678.xlsx

When you have finished transcribing your sheets you will submit the Excel files to Moodle. As mentioned above, I strongly advise that you create a folder where you can save the Excel files. To submit all 18 files: create a zip file of the folder containing the files (right click on the folder and click send to 'compressed zipped folder') and then upload this to Moodle. The screenshot below shows you how to create a zip file of your folder. I will set up a submission tab in Moodle closer to the submission date.



Student Supports

This is an innovative and interesting assignment, but I am sure you will need to call on some help as you come across oddities in the sheets and throughout the process. If you need help don't be afraid to ask! We will use to Moodle discussion forum to address any questions you have with regard to transcribing the data. For example, some of the sheets may have information that is difficult to read. In this instance, you can post a question to the discussion forum explaining the problem and stating the image number so that I may refer to the original sheet. I will run a check-in clinic in class to go through any common issues that arise and to answer any questions. I will also have weekly office hours on Tuesdays from 3:00-5:30pm when you can come and speak to me. There is plenty of support and I encourage you to take advantage of the help!

Appendix IV: Chapter 4 Supplementary Information

Table S1: PRCPTOT trend statistics for individual stations for 1910-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	-0.61	0.54	±1.96	(-)	-0.281	-1.163	0.526
175	0.41	0.68	±1.96	(+)	0.097	-0.48	0.723
201	3.02	0	±1.96	(+)*	1.858	0.105	3.565
417	0.49	0.63	±1.96	(+)	0.354	-0.953	1.711
675	1.65	0.1	±1.96	(+)	0.66	-0.078	1.445
706	1.15	0.25	±1.96	(+)	0.35	-0.485	1.16
875	0.95	0.34	±1.96	(+)	0.313	-0.393	0.97
1075	-1.34	0.18	±1.96	(-)	-0.471	-1.419	0.436
1175	1.09	0.28	±1.96	(+)	0.847	-0.373	1.998
1275	1.48	0.14	±1.96	(+)	0.702	-0.127	1.651
1338	0.01	0.99	±1.96	(+)	0.01	-1.112	1.097
1375	1.01	0.31	±1.96	(+)	0.062	-0.524	0.68
1475	1.39	0.16	±1.96	(+)	0.443	-0.269	1.132
1519	-0.04	0.97	±1.96	(-)	-0.009	-0.7	0.667
1575	3	0	±1.96	(+)*	0.967	0.2	1.673
1812	-1.56	0.12	±1.96	(-)	-0.788	-1.717	0.087
1819	1.83	0.07	±1.96	(+)	0.947	-0.054	1.935
1923	-0.46	0.65	±1.96	(-)	-0.212	-1.308	0.933
1929	1.39	0.16	±1.96	(+)	0.518	-0.175	1.226
2012	2.47	0.01	±1.96	(+)*	0.708	-0.138	1.538
2115	0.19	0.85	±1.96	(+)	0.106	-0.903	0.997
2275	1.51	0.13	±1.96	(+)	0.889	-0.304	2.085
2375	0.75	0.45	±1.96	(+)	0.357	-0.523	1.173
2528	-0.46	0.64	±1.96	(-)	-0.269	-1.326	0.786
4015	-1.31	0.19	±1.96	(-)	-0.419	-1.325	0.482
4513	0.78	0.44	±1.96	(+)	0.21	-0.454	0.881
5131	0.54	0.59	±1.96	(+)	0.081	-0.653	0.755
6019	0.29	0.77	±1.96	(+)	0.147	-0.917	1.25
6329	2.97	0	±1.96	(+)*	0.954	0.295	1.667
8212	0.53	0.6	±1.96	(+)	0.256	-0.68	1.236

Table S2: R5mm trend statistics for individual stations for 1910-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	-1.54	0.12	±1.96	(-)	-0.043	-0.105	0.013
175	0.11	0.91	±1.96	(+)	0	-0.043	0.05

201	1.69	0.09	±1.96	(+)	0.085	0	0.179
417	-0.72	0.47	±1.96	(-)	-0.03	-0.118	0.063
675	0.57	0.57	±1.96	(+)	0.015	-0.05	0.087
706	1.04	0.3	±1.96	(+)	0.02	-0.042	0.089
875	0.5	0.62	±1.96	(+)	0	-0.044	0.071
1075	-1.64	0.1	±1.96	(-)	-0.049	-0.118	0.015
1175	2.17	0.03	±1.96	(+)*	0.076	0	0.161
1275	1.02	0.31	±1.96	(+)	0.044	-0.032	0.122
1338	0.19	0.85	±1.96	(+)	0	-0.074	0.091
1375	0.22	0.83	±1.96	(+)	0	-0.047	0.057
1475	1.18	0.24	±1.96	(+)	0.038	-0.022	0.101
1519	0.15	0.88	±1.96	(+)	0	-0.059	0.063
1575	2.86	0	±1.96	(+)*	0.071	0	0.133
1812	-1.74	0.08	±1.96	(-)	-0.054	-0.117	0
1819	1.29	0.2	±1.96	(+)	0.049	-0.026	0.125
1923	0.2	0.85	±1.96	(+)	0	-0.067	0.078
1929	1.25	0.21	±1.96	(+)	0.037	-0.025	0.101
2012	0.64	0.52	±1.96	(+)	0.022	-0.039	0.095
2115	-1.25	0.21	±1.96	(-)	-0.033	-0.105	0.026
2275	1.08	0.28	±1.96	(+)	0.054	-0.034	0.143
2375	2.84	0	±1.96	(+)*	0.054	-0.014	0.127
2528	-0.78	0.43	±1.96	(-)	-0.03	-0.114	0.04
4015	-1.78	0.08	±1.96	(-)	-0.041	-0.103	0.016
4513	0.86	0.39	±1.96	(+)	0.02	-0.027	0.077
5131	-0.48	0.63	±1.96	(-)	-0.01	-0.065	0.041
6019	-0.54	0.59	±1.96	(-)	-0.014	-0.111	0.061
6329	2.06	0.04	±1.96	(+)*	0.05	0	0.118
8212	-0.28	0.78	±1.96	(-)	0	-0.065	0.052

Table S3: R10mm trend statistics for individual stations for 1910-2019 fixed period.

st_id	**MMK Z-value**	**P-value**	**Z critical value (α = 0.5)**	**MMK trend**	**Sen's slope**	**CI lower**	**CI upper**
108	0.13	0.9	±1.96	(+)	0	-0.032	0.041
175	1.93	0.05	±1.96	(+)	0.019	0	0.044
201	1.78	0.07	±1.96	(+)	0.071	0	0.148
417	1.18	0.24	±1.96	(+)	0.035	-0.018	0.091
675	1.43	0.15	±1.96	(+)	0.033	0	0.07
706	-0.27	0.79	±1.96	(-)	0	-0.039	0.031
875	1.5	0.13	±1.96	(+)	0.024	0	0.056
1075	-0.5	0.61	±1.96	(-)	0	-0.047	0.026
1175	0.6	0.55	±1.96	(+)	0.015	-0.039	0.073

1275	2.03	0.04	±1.96	(+)*	0.042	0	0.083
1338	0.4	0.69	±1.96	(+)	0	-0.036	0.053
1375	2.85	0	±1.96	(+)*	0.03	0	0.056
1475	1.07	0.28	±1.96	(+)	0.015	-0.011	0.044
1519	-0.73	0.47	±1.96	(-)	-0.01	-0.046	0.018
1575	1.07	0.29	±1.96	(+)	0.02	-0.012	0.061
1812	-0.37	0.71	±1.96	(-)	0	-0.043	0.029
1819	1.57	0.12	±1.96	(+)	0.042	0	0.092
1923	0.4	0.69	±1.96	(+)	0	-0.036	0.051
1929	0.64	0.52	±1.96	(+)	0	-0.02	0.043
2012	1.44	0.15	±1.96	(+)	0.033	0	0.071
2115	0.86	0.39	±1.96	(+)	0.019	-0.02	0.069
2275	1.44	0.15	±1.96	(+)	0.029	-0.02	0.087
2375	1.36	0.17	±1.96	(+)	0.029	0	0.071
2528	-0.3	0.76	±1.96	(-)	0	-0.053	0.036
4015	0.75	0.46	±1.96	(+)	0	-0.026	0.051
4513	2.83	0	±1.96	(+)*	0.047	0.014	0.077
5131	0.62	0.53	±1.96	(+)	0	-0.019	0.043
6019	0.28	0.78	±1.96	(+)	0	-0.043	0.06
6329	1.03	0.31	±1.96	(+)	0.015	-0.013	0.056
8212	0.71	0.48	±1.96	(+)	0.012	-0.024	0.057

Table S4: R20mm trend statistics for individual stations for 1910-2019 fixed period.

st_id	**MMK Z-value**	**P-value**	**Z critical value (α = 0.5)**	**MMK trend**	**Sen's slope**	**CI lower**	**CI upper**
108	0.53	0.6	±1.96	(+)	0	-0.01	0.023
175	1.8	0.07	±1.96	(+)	0	0	0.014
201	1.76	0.08	±1.96	(+)	0.029	0	0.067
417	1.52	0.13	±1.96	(+)	0.023	0	0.05
675	1.12	0.26	±1.96	(+)	0	0	0.014
706	-0.16	0.87	±1.96	(-)	0	-0.013	0
875	1.63	0.1	±1.96	(+)	0	0	0.022
1075	1.7	0.09	±1.96	(+)	0	0	0.026
1175	1.68	0.09	±1.96	(+)	0.011	0	0.034
1275	0.24	0.81	±1.96	(+)	0	0	0.01
1338	1.89	0.06	±1.96	(+)	0	0	0.026
1375	1.54	0.12	±1.96	(+)	0	0	0.016
1475	0.1	0.92	±1.96	(+)	0	0	0
1519	2.1	0.04	±1.96	(+)*	0	0	0.018
1575	0.53	0.6	±1.96	(+)	0	0	0.013
1812	-0.25	0.8	±1.96	(-)	0	-0.021	0.015
1819	1.17	0.24	±1.96	(+)	0	0	0.022

1923	-0.64	0.52	±1.96	(-)	0	-0.026	0.014
1929	1.42	0.16	±1.96	(+)	0	0	0.023
2012	1.7	0.09	±1.96	(+)	0.011	0	0.026
2115	3	0	±1.96	(+)*	0.013	0	0.031
2275	0.63	0.53	±1.96	(+)	0	-0.016	0.035
2375	-0.71	0.48	±1.96	(-)	0	-0.02	0
2528	1	0.32	±1.96	(+)	0	0	0.026
4015	1.01	0.31	±1.96	(+)	0	0	0.032
4513	2.16	0.03	±1.96	(+)*	0.013	0	0.026
5131	0.36	0.72	±1.96	(+)	0	0	0.013
6019	0.12	0.91	±1.96	(+)	0	-0.019	0.023
6329	-0.16	0.87	±1.96	(-)	0	0	0
8212	1.9	0.06	±1.96	(+)	0.019	0	0.042

Table S5: R95pTOT trend statistics for individual stations for 1910-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	1.59	0.11	±1.96	(+)	0.444	-0.083	0.99
175	0.53	0.59	±1.96	(+)	0.063	-0.346	0.462
201	1.7	0.09	±1.96	(+)	0.748	-0.238	1.74
417	2.28	0.02	±1.96	(+)*	0.755	0.097	1.427
675	0.99	0.32	±1.96	(+)	0.332	-0.109	0.827
706	-0.3	0.76	±1.96	(-)	-0.07	-0.489	0.379
875	2.03	0.04	±1.96	(+)*	0.444	0.013	0.871
1075	1.7	0.09	±1.96	(+)	0.291	-0.204	0.763
1175	1.44	0.15	±1.96	(+)	0.367	-0.221	0.973
1275	1.25	0.21	±1.96	(+)	0.285	-0.168	0.681
1338	0.37	0.71	±1.96	(+)	0.128	-0.507	0.761
1375	2.11	0.04	±1.96	(+)*	0.363	-0.017	0.735
1475	0.77	0.44	±1.96	(+)	0.189	-0.226	0.627
1519	1.3	0.2	±1.96	(+)	0.209	-0.155	0.571
1575	0.59	0.56	±1.96	(+)	0.122	-0.289	0.57
1812	-0.15	0.88	±1.96	(-)	-0.038	-0.592	0.549
1819	2.17	0.03	±1.96	(+)*	0.603	0.067	1.175
1923	-0.26	0.79	±1.96	(-)	-0.081	-0.836	0.662
1929	1.53	0.13	±1.96	(+)	0.394	-0.041	0.829
2012	1.71	0.09	±1.96	(+)	0.499	0.026	0.951
2115	1.24	0.21	±1.96	(+)	0.311	-0.174	0.818
2275	1.78	0.07	±1.96	(+)	0.366	-0.277	1.038
2375	-0.36	0.72	±1.96	(-)	-0.082	-0.62	0.436
2528	0.75	0.45	±1.96	(+)	0.212	-0.338	0.735
4015	1.26	0.21	±1.96	(+)	0.414	-0.144	1.025

4513	2.41	0.02	±1.96	(+)*	0.623	0.217	1.074
5131	0.01	0.99	±1.96	(+)	0.002	-0.473	0.491
6019	-0.19	0.85	±1.96	(-)	-0.07	-0.653	0.579
6329	0.2	0.84	±1.96	(+)	0.052	-0.433	0.493
8212	1.7	0.09	±1.96	(+)	0.448	-0.067	0.992

Table S6: R99pTOT trend statistics for individual stations for 1910-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	1.69	0.09	±1.96	(+)	0.138	0	0.577
175	0.91	0.36	±1.96	(+)	0.01	-0.038	0.32
201	-0.1	0.92	±1.96	(-)	0	-0.3	0.138
417	5.46	0	±1.96	(+)*	0.406	0	0.846
675	0.31	0.76	±1.96	(+)	0.027	-0.205	0.303
706	0.88	0.38	±1.96	(+)	0.058	-0.049	0.383
875	2.78	0.01	±1.96	(+)*	0.302	0	0.56
1075	-0.02	0.99	±1.96	(-)	0	-0.074	0.057
1175	1.27	0.21	±1.96	(+)	0.135	-0.055	0.559
1275	-0.83	0.41	±1.96	(-)	-0.035	-0.341	0.072
1338	0.02	0.98	±1.96	(+)	0	-0.265	0.206
1375	1.72	0.09	±1.96	(+)	0.152	0	0.414
1475	0.05	0.96	±1.96	(+)	0	-0.233	0.272
1519	1.22	0.22	±1.96	(+)	0.012	-0.054	0.26
1575	2.23	0.03	±1.96	(+)*	0.029	-0.04	0.35
1812	0.65	0.51	±1.96	(+)	0	-0.095	0.182
1819	1.1	0.27	±1.96	(+)	0.038	-0.024	0.385
1923	-0.09	0.93	±1.96	(-)	0	-0.389	0.408
1929	2.79	0.01	±1.96	(+)*	0.246	0	0.487
2012	1.53	0.13	±1.96	(+)	0.104	0	0.467
2115	2.84	0	±1.96	(+)*	0.32	0	0.673
2275	1.94	0.05	±1.96	(+)	0.131	0	0.668
2375	-1.51	0.13	±1.96	(-)	-0.2	-0.513	0.004
2528	1.19	0.24	±1.96	(+)	0.132	-0.088	0.484
4015	0.94	0.34	±1.96	(+)	0	0	0.247
4513	1.27	0.2	±1.96	(+)	0.15	0	0.507
5131	-0.95	0.34	±1.96	(-)	-0.108	-0.396	0.14
6019	0.9	0.37	±1.96	(+)	0.075	-0.106	0.387
6329	-0.02	0.99	±1.96	(-)	0	-0.227	0.231
8212	0.03	0.98	±1.96	(+)	0	-0.069	0.11

Table S7: Rx1day trend statistics for individual stations for 1910-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	0.62	0.54	±1.96	(+)	0.02	-0.045	0.08
175	0.47	0.64	±1.96	(+)	0.013	-0.033	0.066
201	2.94	0	±1.96	(+)*	0.027	-0.068	0.125
417	1.69	0.09	±1.96	(+)	0.059	-0.001	0.125
675	1.11	0.27	±1.96	(+)	0.014	-0.029	0.06
706	0.25	0.8	±1.96	(+)	0.006	-0.042	0.058
875	1.75	0.08	±1.96	(+)	0.035	-0.016	0.085
1075	-0.25	0.8	±1.96	(-)	-0.009	-0.066	0.048
1175	1.36	0.17	±1.96	(+)	0.032	-0.02	0.09
1275	-0.66	0.51	±1.96	(-)	-0.012	-0.057	0.035
1338	0.05	0.96	±1.96	(+)	0.002	-0.078	0.086
1375	1.49	0.14	±1.96	(+)	0.025	-0.021	0.074
1475	-0.49	0.62	±1.96	(-)	-0.015	-0.056	0.033
1519	-0.04	0.97	±1.96	(-)	0	-0.04	0.038
1575	1.02	0.31	±1.96	(+)	0.023	-0.029	0.076
1812	-1.67	0.09	±1.96	(-)	-0.03	-0.085	0.025
1819	1.07	0.28	±1.96	(+)	0.031	-0.021	0.084
1923	-0.52	0.6	±1.96	(-)	-0.026	-0.131	0.076
1929	1.3	0.19	±1.96	(+)	0.033	-0.018	0.087
2012	0.94	0.35	±1.96	(+)	0.03	-0.028	0.086
2115	2.58	0.01	±1.96	(+)*	0.054	0.003	0.105
2275	0.41	0.68	±1.96	(+)	0.015	-0.048	0.075
2375	-1.27	0.2	±1.96	(-)	-0.031	-0.08	0.016
2528	0.75	0.45	±1.96	(+)	0.023	-0.039	0.083
4015	0.4	0.69	±1.96	(+)	0.009	-0.038	0.059
4513	0.38	0.7	±1.96	(+)	0.013	-0.038	0.062
5131	-0.33	0.74	±1.96	(-)	-0.009	-0.057	0.038
6019	0.38	0.71	±1.96	(+)	0.01	-0.041	0.057
6329	0.21	0.83	±1.96	(+)	0.005	-0.04	0.051
8212	0.35	0.72	±1.96	(+)	0.009	-0.048	0.065

Table S8: Rx5day trend statistics for individual stations for 1910-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	2	0.05	±1.96	(+)*	0.091	0.002	0.183
175	-0.03	0.98	±1.96	(-)	0	-0.086	0.098
201	1.33	0.18	±1.96	(+)	0.101	-0.053	0.255
417	2.68	0.01	±1.96	(+)*	0.128	0.03	0.231
675	2.25	0.02	±1.96	(+)*	0.088	0.014	0.163
706	-1.21	0.23	±1.96	(-)	-0.027	-0.113	0.052

875	1.4	0.16	±1.96	(+)	0.046	-0.029	0.123
1075	0.79	0.43	±1.96	(+)	0.031	-0.047	0.109
1175	2.25	0.02	±1.96	(+)*	0.08	-0.015	0.181
1275	0.49	0.62	±1.96	(+)	0.013	-0.053	0.091
1338	0.94	0.35	±1.96	(+)	0.051	-0.065	0.178
1375	0.81	0.42	±1.96	(+)	0.025	-0.052	0.106
1475	0.83	0.41	±1.96	(+)	0.029	-0.036	0.094
1519	0.44	0.66	±1.96	(+)	0.015	-0.054	0.079
1575	1.08	0.28	±1.96	(+)	0.03	-0.035	0.102
1812	0.71	0.48	±1.96	(+)	0.03	-0.051	0.123
1819	2.46	0.01	±1.96	(+)*	0.082	0	0.181
1923	-0.16	0.87	±1.96	(-)	-0.013	-0.181	0.156
1929	2.82	0	±1.96	(+)*	0.089	0.016	0.156
2012	0.47	0.64	±1.96	(+)	0.021	-0.067	0.102
2115	1.58	0.11	±1.96	(+)	0.071	-0.042	0.18
2275	1.65	0.1	±1.96	(+)	0.073	-0.015	0.158
2375	-0.85	0.39	±1.96	(-)	-0.033	-0.107	0.044
2528	1.91	0.06	±1.96	(+)	0.093	-0.001	0.179
4015	0.36	0.72	±1.96	(+)	0.019	-0.076	0.116
4513	1.09	0.28	±1.96	(+)	0.036	-0.031	0.104
5131	-0.61	0.54	±1.96	(-)	-0.024	-0.087	0.049
6019	2.91	0	±1.96	(+)*	0.067	-0.033	0.163
6329	2.11	0.04	±1.96	(+)*	0.045	-0.042	0.126
8212	0.89	0.37	±1.96	(+)	0.047	-0.058	0.156

Table S9: CDD trend statistics for individual stations for 1910-2019 fixed period.

st_id	**MMK Z-value**	**P-value**	**Z critical value (α = 0.5)**	**MMK trend**	**Sen's slope**	**CI lower**	**CI upper**
108	0.66	0.51	±1.96	(+)	0	-0.02	0.037
175	0.17	0.87	±1.96	(+)	0	-0.028	0.033
201	-1.45	0.15	±1.96	(-)	-0.021	-0.05	0
417	0.1	0.92	±1.96	(+)	0	-0.024	0.025
675	1.01	0.31	±1.96	(+)	0.013	-0.014	0.049
706	-0.95	0.34	±1.96	(-)	0	-0.041	0.016
875	0.06	0.95	±1.96	(+)	0	-0.024	0.024
1075	-0.35	0.73	±1.96	(-)	0	-0.045	0.029
1175	-0.88	0.38	±1.96	(-)	-0.01	-0.038	0
1275	0.51	0.61	±1.96	(+)	0	-0.017	0.032
1338	-0.12	0.9	±1.96	(-)	0	-0.048	0.039
1375	0.32	0.75	±1.96	(+)	0	-0.026	0.034
1475	0.27	0.79	±1.96	(+)	0	-0.022	0.031
1519	0.84	0.4	±1.96	(+)	0	-0.017	0.037
1575	-1.05	0.29	±1.96	(-)	0	-0.039	0

1812	-0.15	0.88	±1.96	(-)	0	-0.037	0.035
1819	1.23	0.22	±1.96	(+)	0.022	0	0.055
1923	-0.08	0.93	±1.96	(-)	0	-0.029	0.027
1929	0.02	0.99	±1.96	(+)	0	-0.029	0.029
2012	-0.41	0.68	±1.96	(-)	0	-0.042	0.024
2115	2.72	0.01	±1.96	(+)*	0.035	0	0.07
2275	-0.34	0.74	±1.96	(-)	0	-0.03	0.018
2375	-0.86	0.39	±1.96	(-)	0	-0.036	0.015
2528	0.45	0.66	±1.96	(+)	0	-0.022	0.038
4015	-0.54	0.59	±1.96	(-)	0	-0.044	0.025
4513	1.71	0.09	±1.96	(+)	0.024	0	0.056
5131	0.83	0.4	±1.96	(+)	0	-0.017	0.039
6019	-1.08	0.28	±1.96	(-)	-0.015	-0.043	0
6329	0.67	0.5	±1.96	(+)	0	-0.012	0.034
8212	-1.37	0.17	±1.96	(-)	-0.031	-0.069	0

Table S10: CWD trend statistics for individual stations for 1910-2019 fixed period.

st_id	**MMK Z-value**	**P-value**	**Z critical value (α = 0.5)**	**MMK trend**	**Sen's slope**	**CI lower**	**CI upper**
108	-1.2	0.23	±1.96	(-)	0	-0.025	0
175	-0.24	0.81	±1.96	(-)	0	0	0
201	0.07	0.94	±1.96	(+)	0	-0.042	0.043
417	-1.64	0.1	±1.96	(-)	-0.028	-0.065	0
675	1.02	0.31	±1.96	(+)	0	0	0.031
706	-0.2	0.84	±1.96	(-)	0	-0.022	0.017
875	-1.09	0.27	±1.96	(-)	0	-0.022	0
1075	0.05	0.96	±1.96	(+)	0	-0.015	0.019
1175	1.74	0.08	±1.96	(+)	0.024	0	0.063
1275	-0.35	0.72	±1.96	(-)	0	-0.033	0.025
1338	0.1	0.92	±1.96	(+)	0	-0.019	0.022
1375	-2.29	0.02	±1.96	(-)*	-0.013	-0.031	0
1475	0.83	0.41	±1.96	(+)	0	0	0.032
1519	1.12	0.26	±1.96	(+)	0	0	0.03
1575	2.03	0.04	±1.96	(+)*	0.019	0	0.044
1812	-2.87	0	±1.96	(-)*	-0.025	-0.043	0
1819	0.61	0.54	±1.96	(+)	0	-0.014	0.032
1923	-1.8	0.07	±1.96	(-)	-0.015	-0.036	0
1929	3	0	±1.96	(+)*	0	0	0.03
2012	-0.66	0.51	±1.96	(-)	0	-0.022	0
2115	-0.05	0.96	±1.96	(-)	0	-0.015	0.015
2275	-0.52	0.6	±1.96	(-)	0	-0.05	0.024
2375	0.27	0.79	±1.96	(+)	0	-0.026	0.034
2528	-0.27	0.79	±1.96	(-)	0	-0.038	0.027
4015	-2.42	0.02	±1.96	(-)*	-0.014	-0.032	0

4513	-1.09	0.27	±1.96	(-)	0	-0.023	0
5131	0.62	0.54	±1.96	(+)	0	0	0.029
6019	-0.24	0.81	±1.96	(-)	0	-0.026	0.021
6329	3.96	0	±1.96	(+)*	0.056	0.029	0.08
8212	-4.25	0	±1.96	(-)*	0	-0.02	0

Table S11: SDII trend statistics for individual stations for 1910-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	2.95	0	±1.96	(+)*	0.006	0.002	0.01
175	2.31	0.02	±1.96	(+)*	0.003	0	0.005
201	0.6	0.55	±1.96	(+)	0.001	-0.004	0.007
417	3.86	0	±1.96	(+)*	0.005	0.001	0.009
675	0.67	0.5	±1.96	(+)	0.002	-0.002	0.004
706	-0.27	0.79	±1.96	(-)	-0.001	-0.004	0.003
875	3.54	0	±1.96	(+)*	0.005	0.002	0.008
1075	1.41	0.16	±1.96	(+)	0.003	-0.001	0.007
1175	-0.04	0.97	±1.96	(-)	0.00E+00	-0.004	0.004
1275	1.55	0.12	±1.96	(+)	0.002	-0.001	0.005
1338	-0.47	0.64	±1.96	(-)	-0.001	-0.006	0.003
1375	3.79	0	±1.96	(+)*	0.004	0.001	0.006
1475	0.79	0.43	±1.96	(+)	0.002	-0.001	0.004
1519	-0.08	0.94	±1.96	(-)	0	-0.004	0.003
1575	0.42	0.67	±1.96	(+)	0.001	-0.002	0.004
1812	2.65	0.01	±1.96	(+)*	0.004	0	0.008
1819	1.91	0.06	±1.96	(+)	0.004	0	0.009
1923	2.45	0.01	±1.96	(+)*	0.005	0	0.011
1929	1.04	0.3	±1.96	(+)	0.001	-0.002	0.005
2012	2.11	0.03	±1.96	(+)*	0.004	0.001	0.008
2115	4.9	0	±1.96	(+)*	0.005	0.001	0.009
2275	1.08	0.28	±1.96	(+)	0.003	-0.002	0.007
2375	-0.04	0.97	±1.96	(-)	0.00E+00	-0.003	0.003
2528	0.52	0.61	±1.96	(+)	0.001	-0.002	0.004
4015	3.82	0	±1.96	(+)*	0.007	0.003	0.011
4513	3.94	0	±1.96	(+)*	0.007	0.004	0.01
5131	-0.4	0.69	±1.96	(-)	-0.001	-0.004	0.002
6019	-0.28	0.78	±1.96	(-)	0	-0.004	0.003
6329	-1.98	0.05	±1.96	(-)*	-0.004	-0.007	-0.001
8212	2.25	0.02	±1.96	(+)*	0.005	0.001	0.01

Table S12: PRCPTOT trend statistics for individual stations for 1940-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	0.76	0.44	±1.96	(+)	0.52	-0.914	1.943
175	0.06	0.95	±1.96	(+)	0.03	-0.939	1.103
201	5.97	0	±1.96	(+)*	3.092	0.712	5.609
417	1.94	0.05	±1.96	(+)	2.17	-0.03	4.27
675	2.18	0.03	±1.96	(+)*	1.453	0.119	2.779
703	2.18	0.03	±1.96	(+)*	1.453	0.119	2.779
875	1.42	0.16	±1.96	(+)	0.777	-0.521	1.949
1075	0.57	0.57	±1.96	(+)	0.341	-1.058	1.7
1175	1.33	0.19	±1.96	(+)	1.361	-0.607	3.188
1275	2.7	0.01	±1.96	(+)*	2.1	0.516	3.673
1338	-0.37	0.71	±1.96	(-)	-0.276	-1.732	1.295
1375	1.47	0.14	±1.96	(+)	0.628	-0.333	1.575
1475	1.46	0.15	±1.96	(+)	0.99	-0.222	2.205
1519	-0.81	0.42	±1.96	(-)	-0.487	-1.72	0.631
1575	2.5	0.01	±1.96	(+)*	1.447	0.22	2.596
1812	-0.28	0.78	±1.96	(-)	-0.24	-1.711	1.3
1819	1.47	0.14	±1.96	(+)	0.961	-0.271	2.297
1923	-0.15	0.88	±1.96	(-)	-0.094	-1.725	1.521
1929	-0.32	0.75	±1.96	(-)	-0.271	-1.476	0.929
2012	2.78	0.01	±1.96	(+)*	1.82	0.584	3.043
2115	-0.35	0.72	±1.96	(-)	-0.273	-1.659	1.125
2275	3.08	0	±1.96	(+)*	2.141	0.365	3.925
2375	0.57	0.57	±1.96	(+)	0.439	-1.024	1.798
2528	-0.62	0.54	±1.96	(-)	-0.415	-1.889	0.982
4015	0.11	0.91	±1.96	(+)	0.108	-1.488	1.567
4513	1.4	0.16	±1.96	(+)	0.878	-0.34	1.903
5131	0.87	0.38	±1.96	(+)	0.287	-1.043	1.535
6019	1.01	0.31	±1.96	(+)	0.885	-0.911	2.744
6329	2.13	0.03	±1.96	(+)*	1.082	-0.007	2.354
8212	1.97	0.05	±1.96	(+)*	1.344	0.002	2.878

Table S13: R5mm trend statistics for individual stations for 1940-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	-0.6	0.55	±1.96	(-)	-0.028	-0.133	0.063
175	0.15	0.88	±1.96	(+)	0	-0.067	0.077
201	2.04	0.04	±1.96	(+)*	0.158	0	0.294
417	1.54	0.12	±1.96	(+)	0.091	-0.056	0.232
675	0.93	0.35	±1.96	(+)	0.067	-0.062	0.188
703	0.93	0.35	±1.96	(+)	0.067	-0.062	0.188
875	0.47	0.64	±1.96	(+)	0.022	-0.07	0.118

1075	-0.45	0.66	±1.96	(-)	-0.018	-0.115	0.078
1175	1.75	0.08	±1.96	(+)	0.132	0	0.267
1275	2.12	0.03	±1.96	(+)*	0.143	0	0.273
1338	0.41	0.68	±1.96	(+)	0.024	-0.086	0.138
1375	1.17	0.24	±1.96	(+)	0.047	-0.037	0.133
1475	1.75	0.08	±1.96	(+)	0.091	0	0.182
1519	-1.09	0.28	±1.96	(-)	-0.056	-0.158	0.045
1575	1.92	0.05	±1.96	(+)	0.107	0	0.211
1812	-0.53	0.59	±1.96	(-)	-0.026	-0.13	0.077
1819	1.11	0.26	±1.96	(+)	0.05	-0.043	0.146
1923	0.91	0.36	±1.96	(+)	0.051	-0.063	0.158
1929	-0.3	0.77	±1.96	(-)	-0.013	-0.125	0.094
2012	2.37	0.02	±1.96	(+)*	0.115	0.015	0.2
2115	-1.14	0.25	±1.96	(-)	-0.046	-0.16	0.044
2275	2.68	0.01	±1.96	(+)*	0.188	0.063	0.318
2375	1.23	0.22	±1.96	(+)	0.053	-0.067	0.171
2528	0.09	0.93	±1.96	(+)	0	-0.097	0.096
4015	-0.83	0.41	±1.96	(-)	-0.043	-0.151	0.063
4513	1.59	0.11	±1.96	(+)	0.066	0	0.139
5131	-0.03	0.98	±1.96	(-)	0	-0.097	0.098
6019	1.25	0.21	±1.96	(+)	0.061	-0.071	0.174
6329	0.91	0.36	±1.96	(+)	0.048	-0.05	0.162
8212	0.74	0.46	±1.96	(+)	0.032	-0.056	0.128

Table S14: R10mm trend statistics for individual stations for 1940-2019 fixed period.

st_id	**MMK Z-value**	**P-value**	**Z critical value (α = 0.5)**	**MMK trend**	**Sen's slope**	**CI lower**	**CI upper**
108	0.65	0.52	±1.96	(+)	0.02	-0.041	0.091
175	0.48	0.63	±1.96	(+)	0	-0.026	0.045
201	2.38	0.02	±1.96	(+)*	0.142	0.036	0.237
417	2.38	0.02	±1.96	(+)*	0.103	0	0.2
675	1.67	0.09	±1.96	(+)	0.045	0	0.111
703	1.67	0.09	±1.96	(+)	0.045	0	0.111
875	0.73	0.47	±1.96	(+)	0.019	-0.029	0.071
1075	0.42	0.68	±1.96	(+)	0	-0.048	0.063
1175	1.11	0.27	±1.96	(+)	0.055	-0.033	0.143
1275	9.2	0	±1.96	(+)*	0.115	0.049	0.182
1338	-0.15	0.88	±1.96	(-)	0	-0.067	0.056
1375	0.86	0.39	±1.96	(+)	0.016	-0.021	0.065
1475	0.06	0.95	±1.96	(+)	0	-0.055	0.049
1519	-1.04	0.3	±1.96	(-)	-0.028	-0.087	0.025
1575	1.9	0.06	±1.96	(+)	0.054	0	0.111
1812	-0.36	0.72	±1.96	(-)	0	-0.068	0.048

1819	1.26	0.21	±1.96	(+)	0.05	0	0.118
1923	0.03	0.97	±1.96	(+)	0	-0.067	0.068
1929	-0.35	0.73	±1.96	(-)	0	-0.063	0.038
2012	5.43	0	±1.96	(+)*	0.081	0.019	0.14
2115	0.55	0.58	±1.96	(+)	0.007	-0.041	0.081
2275	1.12	0.26	±1.96	(+)	0.044	-0.029	0.13
2375	0	1	±1.96	(+)	0	-0.065	0.064
2528	-0.61	0.54	±1.96	(-)	0	-0.074	0.033
4015	0.77	0.44	±1.96	(+)	0.022	-0.038	0.091
4513	2.8	0.01	±1.96	(+)*	0.054	0	0.103
5131	0.95	0.34	±1.96	(+)	0.028	-0.031	0.083
6019	0.09	0.93	±1.96	(+)	0	-0.083	0.1
6329	2.26	0.02	±1.96	(+)*	0.06	0	0.122
8212	1.31	0.19	±1.96	(+)	0.043	-0.02	0.116

Table S15: R20mm trend statistics for individual stations for 1940-2019 fixed period.

st_id	**MMK Z-value**	**P-value**	**Z critical value (α = 0.5)**	**MMK trend**	**Sen's slope**	**CI lower**	**CI upper**
108	1.51	0.13	±1.96	(+)	0.02	0	0.054
175	-0.14	0.89	±1.96	(-)	0	0	0
201	2.67	0.01	±1.96	(+)*	0.056	0	0.119
417	1.82	0.07	±1.96	(+)	0.048	0	0.087
675	0.48	0.63	±1.96	(+)	0	0	0.019
703	0.48	0.63	±1.96	(+)	0	0	0.019
875	-0.08	0.93	±1.96	(-)	0	-0.02	0.016
1075	1.97	0.05	±1.96	(+)*	0.017	0	0.046
1175	0.41	0.68	±1.96	(+)	0	-0.021	0.038
1275	0.64	0.52	±1.96	(+)	0	0	0.027
1338	-0.42	0.68	±1.96	(-)	0	-0.029	0.016
1375	1.39	0.16	±1.96	(+)	0	0	0.032
1475	0.21	0.83	±1.96	(+)	0	0	0.018
1519	0.31	0.75	±1.96	(+)	0	0	0.017
1575	-0.34	0.74	±1.96	(-)	0	-0.02	0
1812	0.04	0.97	±1.96	(+)	0	-0.03	0.033
1819	1.15	0.25	±1.96	(+)	0	0	0.032
1923	-1.08	0.28	±1.96	(-)	0	-0.048	0
1929	0.48	0.63	±1.96	(+)	0	0	0.024
2012	1.28	0.2	±1.96	(+)	0	0	0.042
2115	1.69	0.09	±1.96	(+)	0.014	0	0.041
2275	0.2	0.84	±1.96	(+)	0	-0.038	0.043
2375	-1.7	0.09	±1.96	(-)	0	-0.043	0
2528	0.06	0.95	±1.96	(+)	0	-0.02	0.023
4015	1.08	0.28	±1.96	(+)	0.019	0	0.061
4513	1.88	0.06	±1.96	(+)	0.02	0	0.044

5131	0.57	0.57	±1.96	(+)	0	0	0.029
6019	-0.93	0.35	±1.96	(-)	0	-0.056	0.02
6329	-0.34	0.73	±1.96	(-)	0	-0.024	0
8212	2.33	0.02	±1.96	(+)*	0.042	0	0.077

Table S16: R95pTOT trend statistics for individual stations for 1940-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	2.17	0.03	±1.96	(+)*	1.037	0.117	1.953
175	-1.11	0.27	±1.96	(-)	-0.307	-1.006	0.364
201	1.9	0.06	±1.96	(+)	1.38	-0.17	2.92
417	1.93	0.05	±1.96	(+)	1.037	-0.024	2.107
675	1.01	0.31	±1.96	(+)	0.519	-0.29	1.287
703	1.01	0.31	±1.96	(+)	0.519	-0.29	1.287
875	0.93	0.35	±1.96	(+)	0.354	-0.386	1.09
1075	1.22	0.22	±1.96	(+)	0.521	-0.294	1.362
1175	0.26	0.79	±1.96	(+)	0.105	-0.825	1.172
1275	2.62	0.01	±1.96	(+)*	0.871	0.202	1.592
1338	-0.65	0.51	±1.96	(-)	-0.276	-1.11	0.569
1375	1.48	0.14	±1.96	(+)	0.534	-0.209	1.176
1475	0.46	0.64	±1.96	(+)	0.266	-0.556	1.007
1519	-0.23	0.82	±1.96	(-)	-0.094	-0.738	0.546
1575	0.52	0.6	±1.96	(+)	0.16	-0.563	0.915
1812	0.42	0.68	±1.96	(+)	0.089	-0.897	1.055
1819	2.47	0.01	±1.96	(+)*	0.763	0.065	1.472
1923	-0.74	0.46	±1.96	(-)	-0.307	-1.467	0.828
1929	0.74	0.46	±1.96	(+)	0.272	-0.518	0.945
2012	3.51	0	±1.96	(+)*	0.936	0.153	1.695
2115	0.86	0.39	±1.96	(+)	0.344	-0.387	1.164
2275	0.58	0.56	±1.96	(+)	0.23	-0.71	1.213
2375	-1.57	0.12	±1.96	(-)	-0.601	-1.427	0.264
2528	0.36	0.72	±1.96	(+)	0.192	-0.62	0.932
4015	1.23	0.22	±1.96	(+)	0.799	-0.14	1.834
4513	3.04	0	±1.96	(+)*	1.181	0.433	1.788
5131	0.91	0.36	±1.96	(+)	0.406	-0.448	1.213
6019	-0.91	0.36	±1.96	(-)	-0.435	-1.5	0.612
6329	-0.13	0.9	±1.96	(-)	-0.087	-0.903	0.802
8212	2.22	0.03	±1.96	(+)*	1.045	0.113	1.953

Table S17: R99pTOT trend statistics for individual stations for 1940-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	2.38	0.02	±1.96	(+)*	0.594	0	1.229
175	-0.16	0.87	±1.96	(-)	0	-0.37	0.348
201	-0.31	0.76	±1.96	(-)	0	-0.79	0.185
417	4.07	0	±1.96	(+)*	0.863	0.12	1.525
675	0.51	0.61	±1.96	(+)	0.102	-0.268	0.589
703	0.51	0.61	±1.96	(+)	0.102	-0.268	0.589
875	0.79	0.43	±1.96	(+)	0.064	-0.129	0.591
1075	0.32	0.75	±1.96	(+)	0	-0.059	0.28
1175	1.87	0.06	±1.96	(+)	0.401	-0.08	1.02
1275	-0.58	0.56	±1.96	(-)	-0.027	-0.477	0.185
1338	-0.72	0.47	±1.96	(-)	-0.015	-0.688	0.129
1375	2.67	0.01	±1.96	(+)*	0.276	0	0.687
1475	0.55	0.58	±1.96	(+)	0.068	-0.369	0.543
1519	-0.36	0.72	±1.96	(-)	0	-0.357	0.229
1575	0.25	0.81	±1.96	(+)	0	-0.397	0.435
1812	0.2	0.84	±1.96	(+)	0	-0.233	0.381
1819	1.2	0.23	±1.96	(+)	0.067	0	0.592
1923	-0.27	0.79	±1.96	(-)	0	-0.786	0.606
1929	0.07	0.95	±1.96	(+)	0	-0.318	0.38
2012	1.12	0.26	±1.96	(+)	0.126	-0.096	0.691
2115	1.57	0.12	±1.96	(+)	0.365	-0.003	0.85
2275	0.38	0.7	±1.96	(+)	0	-0.363	0.769
2375	-1.08	0.28	±1.96	(-)	-0.2	-0.715	0.057
2528	0.59	0.56	±1.96	(+)	0.078	-0.246	0.531
4015	1.18	0.24	±1.96	(+)	0	0	0.577
4513	2.45	0.01	±1.96	(+)*	0.561	0.022	1.044
5131	0.02	0.98	±1.96	(+)	0	-0.457	0.49
6019	1.28	0.2	±1.96	(+)	0.227	-0.117	0.826
6329	-0.49	0.63	±1.96	(-)	-0.034	-0.549	0.311
8212	0.84	0.4	±1.96	(+)	0	0	0.602

Table S18: Rx1day trend statistics for individual stations for 1940-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	1.31	0.19	±1.96	(+)	0.074	-0.031	0.174
175	0.34	0.73	±1.96	(+)	0.009	-0.086	0.1
201	-0.57	0.57	±1.96	(-)	-0.029	-0.18	0.123
417	2.11	0.03	±1.96	(+)*	0.116	0.01	0.212
675	1.14	0.25	±1.96	(+)	0.04	-0.028	0.12
703	1.14	0.25	±1.96	(+)	0.04	-0.028	0.12
875	0.5	0.62	±1.96	(+)	0.019	-0.056	0.097

1075	0.99	0.32	±1.96	(+)	0.036	-0.053	0.136
1175	0.88	0.38	±1.96	(+)	0.041	-0.045	0.143
1275	-0.53	0.6	±1.96	(-)	-0.024	-0.1	0.057
1338	-0.55	0.58	±1.96	(-)	-0.016	-0.12	0.09
1375	1.22	0.22	±1.96	(+)	0.05	-0.027	0.124
1475	0.19	0.85	±1.96	(+)	0.006	-0.07	0.081
1519	-0.15	0.88	±1.96	(-)	-0.004	-0.06	0.053
1575	0.01	0.99	±1.96	(+)	0	-0.093	0.09
1812	-0.64	0.52	±1.96	(-)	-0.026	-0.121	0.06
1819	0.94	0.35	±1.96	(+)	0.04	-0.029	0.111
1923	-0.22	0.82	±1.96	(-)	-0.016	-0.17	0.155
1929	-1.12	0.26	±1.96	(-)	-0.021	-0.1	0.065
2012	1.28	0.2	±1.96	(+)	0.06	-0.032	0.144
2115	0.81	0.42	±1.96	(+)	0.025	-0.055	0.107
2275	-0.03	0.98	±1.96	(-)	-0.001	-0.095	0.088
2375	-1.13	0.26	±1.96	(-)	-0.044	-0.129	0.032
2528	0.87	0.38	±1.96	(+)	0.034	-0.043	0.112
4015	0.97	0.33	±1.96	(+)	0.044	-0.041	0.122
4513	1.32	0.19	±1.96	(+)	0.06	-0.017	0.133
5131	0.65	0.52	±1.96	(+)	0.028	-0.05	0.112
6019	1.6	0.11	±1.96	(+)	0.061	-0.013	0.133
6329	-0.93	0.35	±1.96	(-)	-0.036	-0.116	0.048
8212	1.54	0.12	±1.96	(+)	0.068	-0.027	0.168

Table S19: Rx5day trend statistics for individual stations for 1940-2019 fixed period.

st_id	**MMK Z-value**	**P-value**	**Z critical value (α = 0.5)**	**MMK trend**	**Sen's slope**	**CI lower**	**CI upper**
108	3.1	0	±1.96	(+)*	0.209	0.05	0.367
175	-0.95	0.34	±1.96	(-)	-0.069	-0.211	0.104
201	0.51	0.61	±1.96	(+)	0.079	-0.171	0.317
417	1.51	0.13	±1.96	(+)	0.131	-0.038	0.306
675	1.4	0.16	±1.96	(+)	0.087	-0.034	0.221
703	1.4	0.16	±1.96	(+)	0.087	-0.034	0.221
875	1.77	0.08	±1.96	(+)	0.081	-0.041	0.202
1075	0.49	0.63	±1.96	(+)	0.014	-0.122	0.146
1175	2.38	0.02	±1.96	(+)*	0.166	-0.006	0.329
1275	0.91	0.36	±1.96	(+)	0.056	-0.058	0.17
1338	-0.31	0.76	±1.96	(-)	-0.025	-0.194	0.139
1375	0.56	0.57	±1.96	(+)	0.03	-0.105	0.159
1475	1.03	0.3	±1.96	(+)	0.058	-0.057	0.168
1519	-1.11	0.27	±1.96	(-)	-0.052	-0.156	0.056
1575	0.45	0.65	±1.96	(+)	0.019	-0.085	0.15
1812	-0.2	0.84	±1.96	(-)	-0.02	-0.156	0.142

1819	1.93	0.05	±1.96	(+)	0.1	-0.002	0.232
1923	0.55	0.58	±1.96	(+)	0.07	-0.21	0.3
1929	0.9	0.37	±1.96	(+)	0.067	-0.053	0.18
2012	1.94	0.05	±1.96	(+)	0.104	-0.023	0.229
2115	1.45	0.15	±1.96	(+)	0.119	-0.06	0.291
2275	0.47	0.63	±1.96	(+)	0.033	-0.11	0.178
2375	-0.26	0.8	±1.96	(-)	-0.018	-0.146	0.128
2528	1.87	0.06	±1.96	(+)	0.122	-0.008	0.221
4015	1.73	0.08	±1.96	(+)	0.127	-0.052	0.289
4513	1.33	0.19	±1.96	(+)	0.075	-0.037	0.184
5131	0.75	0.45	±1.96	(+)	0.044	-0.073	0.178
6019	0.73	0.47	±1.96	(+)	0.061	-0.119	0.214
6329	1.33	0.18	±1.96	(+)	0.1	-0.046	0.215
8212	1.53	0.13	±1.96	(+)	0.141	-0.037	0.309

Table S20: CDD trend statistics for individual stations for 1940-2019 fixed period.

st_id	**MMK Z-value**	**P-value**	**Z critical value (α = 0.5)**	**MMK trend**	**Sen's slope**	**CI lower**	**CI upper**
108	0.76	0.45	±1.96	(+)	0	-0.037	0.063
175	-0.32	0.75	±1.96	(-)	0	-0.063	0.04
201	-0.96	0.34	±1.96	(-)	-0.018	-0.067	0.018
417	-1	0.32	±1.96	(-)	-0.018	-0.067	0
675	-0.1	0.92	±1.96	(-)	0	-0.053	0.048
703	-0.1	0.92	±1.96	(-)	0	-0.053	0.048
875	0.95	0.34	±1.96	(+)	0	-0.03	0.053
1075	-0.9	0.37	±1.96	(-)	-0.027	-0.089	0.032
1175	-0.71	0.47	±1.96	(-)	0	-0.063	0.022
1275	0.67	0.5	±1.96	(+)	0	-0.031	0.061
1338	-0.35	0.73	±1.96	(-)	0	-0.067	0.045
1375	-0.74	0.46	±1.96	(-)	0	-0.068	0.033
1475	-0.52	0.6	±1.96	(-)	0	-0.05	0.037
1519	0.98	0.33	±1.96	(+)	0.02	-0.022	0.068
1575	0.09	0.93	±1.96	(+)	0	-0.034	0.038
1812	0.18	0.85	±1.96	(+)	0	-0.054	0.068
1819	1.2	0.23	±1.96	(+)	0.026	0	0.065
1923	-0.13	0.9	±1.96	(-)	0	-0.045	0.04
1929	0.88	0.38	±1.96	(+)	0.018	-0.027	0.063
2012	-1.16	0.25	±1.96	(-)	-0.029	-0.083	0.018
2115	1.43	0.15	±1.96	(+)	0.031	0	0.08
2275	-0.37	0.71	±1.96	(-)	0	-0.045	0.029
2375	-0.21	0.83	±1.96	(-)	0	-0.042	0.034
2528	0.71	0.48	±1.96	(+)	0	-0.024	0.055
4015	-0.19	0.85	±1.96	(-)	0	-0.057	0.047

4513	0.06	0.95	±1.96	(+)	0	-0.053	0.053
5131	-0.08	0.94	±1.96	(-)	0	-0.048	0.043
6019	-1.34	0.18	±1.96	(-)	-0.036	-0.091	0
6329	0.51	0.61	±1.96	(+)	0	-0.029	0.053
8212	-0.65	0.52	±1.96	(-)	-0.019	-0.08	0.036

Table S21: CWD trend statistics for individual stations for 1940-2019 fixed period.

st_id	**MMK Z-value**	**P-value**	**Z critical value (α = 0.5)**	**MMK trend**	**Sen's slope**	**CI lower**	**CI upper**
108	0.59	0.56	±1.96	(+)	0	0	0.033
175	-0.62	0.54	±1.96	(-)	0	-0.021	0
201	2.17	0.03	±1.96	(+)*	0.059	0	0.125
417	1.02	0.31	±1.96	(+)	0.027	-0.024	0.083
675	1.87	0.06	±1.96	(+)	0.028	0	0.054
703	1.87	0.06	±1.96	(+)	0.028	0	0.054
875	1.22	0.22	±1.96	(+)	0	0	0.037
1075	1.33	0.18	±1.96	(+)	0.03	0	0.067
1175	0.73	0.46	±1.96	(+)	0.021	-0.039	0.083
1275	3.51	0	±1.96	(+)*	0.055	0	0.103
1338	0.16	0.87	±1.96	(+)	0	-0.021	0.029
1375	-1.17	0.24	±1.96	(-)	0	-0.036	0
1475	0.67	0.5	±1.96	(+)	0	-0.023	0.049
1519	0.51	0.61	±1.96	(+)	0	-0.022	0.04
1575	2.32	0.02	±1.96	(+)*	0.036	0	0.079
1812	-0.28	0.78	±1.96	(-)	0	-0.029	0.02
1819	0.67	0.5	±1.96	(+)	0	-0.017	0.038
1923	0.77	0.44	±1.96	(+)	0	0	0.03
1929	-1.01	0.31	±1.96	(-)	0	-0.048	0.018
2012	0.76	0.45	±1.96	(+)	0	-0.016	0.042
2115	-0.44	0.66	±1.96	(-)	0	-0.033	0.018
2275	1.95	0.05	±1.96	(+)	0.047	0	0.097
2375	2.37	0.02	±1.96	(+)*	0.056	0	0.108
2528	-0.58	0.56	±1.96	(-)	0	-0.064	0.031
4015	-0.75	0.45	±1.96	(-)	0	-0.038	0
4513	-1.13	0.26	±1.96	(-)	0	-0.04	0
5131	-0.35	0.73	±1.96	(-)	0	-0.042	0.026
6019	4.31	0	±1.96	(+)*	0.043	0	0.082
6329	2.28	0.02	±1.96	(+)*	0.062	0	0.111
8212	0.84	0.4	±1.96	(+)	0	-0.018	0.034

Table S22: SDII trend statistics for individual stations for 1940-2019 fixed period.

st_id	MMK Z-value	P-value	Z critical value (α = 0.5)	MMK trend	Sen's slope	CI lower	CI upper
108	2.62	0.01	±1.96	(+)*	0.01	0.002	0.017
175	0.52	0.6	±1.96	(+)	0.001	-0.003	0.006
201	1.37	0.17	±1.96	(+)	0.005	-0.002	0.012
417	1.88	0.06	±1.96	(+)	0.006	0	0.012
675	0.51	0.61	±1.96	(+)	0.002	-0.004	0.007
703	0.51	0.61	±1.96	(+)	0.002	-0.004	0.007
875	0.63	0.53	±1.96	(+)	0.001	-0.003	0.006
1075	-0.18	0.86	±1.96	(-)	-0.001	-0.007	0.005
1175	0.48	0.63	±1.96	(+)	0.001	-0.005	0.007
1275	2.05	0.04	±1.96	(+)*	0.005	0	0.01
1338	-1.48	0.14	±1.96	(-)	-0.005	-0.011	0.002
1375	2	0.05	±1.96	(+)*	0.003	-0.001	0.008
1475	-0.07	0.94	±1.96	(-)	0	-0.005	0.005
1519	-1.01	0.31	±1.96	(-)	-0.002	-0.007	0.003
1575	0.62	0.54	±1.96	(+)	0.001	-0.003	0.006
1812	1.38	0.17	±1.96	(+)	0.004	-0.003	0.011
1819	1.23	0.22	±1.96	(+)	0.004	-0.001	0.009
1923	1.45	0.15	±1.96	(+)	0.005	-0.003	0.012
1929	2.2	0.03	±1.96	(+)*	0.003	-0.001	0.008
2012	3.37	0	±1.96	(+)*	0.009	0.004	0.014
2115	7.67	0	±1.96	(+)*	0.005	0	0.011
2275	-0.23	0.82	±1.96	(-)	-0.001	-0.008	0.006
2375	-2.41	0.02	±1.96	(-)*	-0.005	-0.009	0
2528	2	0.05	±1.96	(+)*	0.004	0.00E+00	0.008
4015	2.36	0.02	±1.96	(+)*	0.008	0.002	0.015
4513	4.64	0	±1.96	(+)*	0.01	0.005	0.015
5131	0.54	0.59	±1.96	(+)	0.003	-0.003	0.009
6019	-0.76	0.45	±1.96	(-)	-0.002	-0.008	0.004
6329	-0.82	0.41	±1.96	(-)	-0.003	-0.007	0.002
8212	1.72	0.08	±1.96	(+)	0.007	-0.001	0.014