METEOROLOGICAL APPLICATIONS
Meteorol. Appl. **25**: 292–301 (2018)
Published online 28 November 2017 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/met.1691

Verification and bias correction of ECMWF forecasts for Irish weather stations to evaluate their potential usefulness in grass growth modelling

Jack McDonnell,[a,b]* Keith Lambkin,[c] Rowan Fealy,[d] Deirdre Hennessy,[b] Laurence Shalloo[b] and Caroline Brophy[a]
[a] *Department of Mathematics and Statistics, Maynooth University, Ireland*
[b] *Grassland Science Research Department, Teagasc, Animal and Grassland Research and Innovation Centre, Co. Cork, Ireland*
[c] *Met Éireann - Agricultural Meteorology Unit, Dublin, Ireland*
[d] *Department of Geography, Maynooth University, Ireland*

ABSTRACT: Typical weather in Ireland provides conditions favourable for sustaining grass growth throughout most of the year. This affords grass based farming a significant economic advantage due to the low input costs associated with grass production. To optimize the productivity of grass based systems, farmers must manage the resource over short time scales. While research has been conducted into developing predictive grass growth models for Ireland to support on-farm decision making, short term weather forecasts have not yet been incorporated into these models. To assess their potential for use in predictive grass growth models, deterministic forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) were verified for lead times up to 10 days using observations from 25 Irish weather stations. Forecasts of air temperature variables were generally precise at all lead times, particularly up to 7 days. Verification of ECMWF soil temperature forecasts is limited, but here they were shown to be accurate at all depths and most precise at greater depths such as 50 cm. Rainfall forecasts performed well up to approximately 5 days. Seven bias correction techniques were assessed to minimize systematic biases in the forecasts. Based on the root mean squared error values, no large improvement was identified for rainfall forecasts on equivalent ECMWF forecasts, but the optimum bias corrections improved air and soil temperature forecasts greatly. Overall, the results demonstrated that forecasts predict observations accurately up to approximately a week in advance and therefore could prove valuable in grass growth prediction at farm level in Ireland.

KEY WORDS forecast verification; bias correction; Ireland; air temperature; rainfall; soil temperature; grass growth; agriculture

Received 12 January 2017; Revised 25 May 2017; Accepted 7 July 2017

1. Introduction

Agriculture is the largest indigenous industry in Ireland; specifically, the agri-food industry accounts for approximately €26 billion of total turnover (~7% GDP) and 8.4% of employment (DAFM, 2015). At the primary agricultural production level, the grass-based beef and dairy sectors account for almost 70% of production (DAFM, 2015). In recognition of the 2015 abolition of EU quotas restricting milk production, the Irish Government established targets to increase the value of primary production by 65% and agri-food exports by 85% by 2025, compared to 2012–2014 levels (DAFM, 2015). To meet these targets and the growing demands of the international community, farming practices in Ireland will need to optimize the use of valuable natural resources such as grass. Shalloo (2009) and Dillon (2011) previously identified a strong positive linear association between grass use (tonnes dry matter *per* hectare (t DM ha^{-1})) and on-farm net profitability (€ ha^{-1}) in Ireland. Due to the mild, maritime climate, with mean annual temperatures from 9 to 11 °C and the typically low intensity, long duration of rainfall throughout the year, optimum conditions for grass growth are achieved during most of the year (Hurtado-Uria *et al.*, 2013a). The production of grass, with low input cost requirements, also minimizes the need for costly alternative feed supplements. Consequently, primary agricultural production has largely developed around grass/pasture based systems, making it a key element of Irish agricultural productivity (Finneran *et al.*, 2010) and profitability.

Many factors influence grass productivity, some of which are within the farmers' control (e.g. stocking rate, fertilizer application), while others are outside (e.g. meteorological conditions, soil type). Frame (1992) has previously highlighted the influence of weather on grass production and use. In particular, air and soil temperatures as well as rainfall are critical factors determining both growing season length and rate of growth (Brereton, 1995; Thorvaldsson *et al.*, 2004). Perennial ryegrass (*Lolium perenne* L.) is the most widely sown grass species in Ireland (DAFM, 2016), with leaf growth beginning at 5 °C and reaching its peak between 20 and 25 °C (Hopkins, 2000). Temperatures during spring and autumn typically vary between 5 and 10 °C and therefore determine the growing season length (Burke *et al.*, 2004). Soil temperatures between 0 and 10 cm below the surface are particularly important for grass growth in Ireland (Hurtado-Uria *et al.*, 2013a). The year-round rainfall and the low permeability of many Irish soil types means that excessive moisture is usually more problematic for grass growth in Ireland than insufficient moisture availability (Burke *et al.*, 2004). However, adequate rainfall during the growing season is essential for grass growth, with winter rainfall necessary to establish sufficient soil water levels in spring for growth to occur (Frame, 1992).

* Correspondence: J. McDonnell, Department of Mathematics and Statistics, Maynooth University, Maynooth, Co. Kildare, Ireland. E-mail: jack.mcdonnell.2011@mumail.ie

To ensure sufficient supply and the optimized use of grass throughout the growing season, grasslands are required to be managed on short time scales, optimally on a daily to weekly basis. The timing of management decisions, such as removing excess herbage at times of peak growth to maintain grass quality, directly affects farm efficiency and ultimately profitability (O'Donovan, 2000; O'Donovan *et al.*, 2011). While tools currently exist to assist on-farm decision making for grass budgeting/accounting (e.g. grazing planners; PastureBase Ireland – a national grassland database), the use of an operational grass growth model, employing 'local' weather conditions, would assist farmers with management decisions influencing grass growth such as nitrogen fertilizer application, stocking rate and rotation length. Grass growth models using weather observations retrospectively have previously been developed (Hurtado-Uria *et al.*, 2013b). The research outlined here seeks to build on these developments through the evaluation of outputs from a numerical weather prediction (NWP) model for use in forecasting meteorological parameters that influence grass growth, over time scales of relevance for improving on-farm decision making.

Over the last decade, NWP forecasts have become more skilful largely due to significant advances in computational resources, resolution and improved parameterization schemes; but model skill varies depending on the meteorological parameter and the forecast lead time being evaluated. Despite these improvements, they also contain systematic biases (Ebert and McBride, 2000; Sun *et al.*, 2003; Auligne *et al.*, 2007; Roberts, 2008). The biases are most apparent at the surface–atmosphere boundary where errors in fluxes arising from the land surface and atmospheric model interact (Galanis and Anadranistakis, 2002; Mass *et al.*, 2008). Another bias arises due to the model grid representation, an integrated value representative of an area, which is typically evaluated against a proximal weather station.

The presence of systematic biases within NWP outputs can be minimized using appropriate post-processing methods (e.g. Harrison *et al.*, 2000; Boi, 2004; Sweeney *et al.*, 2011; Vannitsem and Hagedorn, 2011). A variety of approaches, including the moving window technique and model output statistics, have been found to improve the accuracy of rainfall and 2 m air temperature forecasts (see Yussouf and Stensrud, 2007; Huang *et al.*, 2012, for example). Many studies attempt to correct forecasts spatially over a domain of interest (Louka *et al.*, 2008; Vrac and Friederichs, 2015, for example), and applying bias correction techniques at individual station locations can yield improvements in forecast accuracy (Taylor and Leslie, 2005).

At long forecast lead times, rainfall is more difficult to forecast, and consequently bias correct, than air temperature. This is largely due to the fact that the processes that give rise to rainfall can occur over small space and time scales (Hamill *et al.*, 2008; Fan and van den Dool, 2011), requiring parameterizations rather than being resolved dynamically. Verification and bias correction of soil temperature forecasts remain limited internationally (but see Albergel *et al.*, 2015), in part due to a lack of suitable databases recording soil temperature observations. However, Met Éireann, the Irish National Meteorological Service, maintains a comprehensive soil temperature observation database for Ireland, with observations taken at six depths for 23 locations. This database is analysed in this paper, providing a detailed case study for soil forecast verification at multiple depths.

The purpose of this case study was to identify the accuracy of European Centre for Medium-Range Weather Forecasts (ECMWF) weather forecasts in Ireland and to improve forecast accuracy where possible using bias correction for potential future inclusion in predictive grass growth models. This objective was achieved using data from a distributed network of 25 weather stations in Ireland over a period of 7 years (2007–2013). The quality of ECMWF forecasts of rainfall, soil temperature and maximum, minimum and mean 2 m air temperature for lead times from 1 to 10 days was assessed. Various bias correction techniques were compared and the resulting forecasts were verified. Air and soil temperature, and rainfall are highly influential in grass growth, but investigating and bias correcting their forecasts is an essential prior step to their inclusion in grass growth models.

2. Data and methods

2.1. Data collation

Weather observations were collated for each of the 25 Met Éireann synoptic weather stations in Ireland for the period from 2007 to 2013 (Figure 1). Hourly observations of air temperature were obtained, while soil temperature (°C) at six depths (5, 10, 20, 30, 50 and 100 cm) was available at 6 h intervals (0300, 0900, 1500 and 2100) at 5, 10, 20 cm and at once *per* day (0900 UTC) for soil depths at 30, 50 and 100 cm. Rainfall data (mm) were available daily. Corresponding forecasts for the 7 years were obtained from the ECWMF operational forecasting model, for model grids matching the locations of the surface weather stations, for forecast lead times from day 1 to day 10. For example, the day 1 forecast for 10 January 2007 was run at 0000 on 10 January 2007, while the day 10 forecast was run at 0000 on 1 January 2007. Day 1 to day 10 are hereafter referred to as 'forecast periods'. As eight of the weather stations became operational after 1 January 2007, observations were not available at all locations for the entire 7 year period examined (Table S1). There was also a small number of missing observations at some stations due to servicing, calibrations or instrument outages. Persistence and mean climatological forecasts were also obtained, with the exception that mean climatological observations were unavailable for soil temperatures.

2.1.1. Observations

The daily mean 2 m air temperature was the mean of the 24 h 2 m air temperature values recorded on each day starting at 0000 Coordinated Universal Time (UTC). If fewer than 13 hourly temperatures were available for a day, the daily value was excluded from analysis. The daily maximum and minimum 2 m air temperatures were the highest and lowest values recorded at the station in the 24 h of the day beginning at 0000. Soil temperature observations were available at 23 of the 25 stations (measurements of soil temperature were unavailable from Finner and Macehead). At each station, soil temperature measurements were taken at 5, 10, 20, 30, 50 and 100 cm. The daily mean soil temperature at 5, 10 and 20 cm was derived from the mean of the observed temperature values for each depth at 0300, 0900, 1500 and 2100. Observations at 30, 50 and 100 cm were measured once daily at 0900. The total daily rainfall was the total precipitation (mm) recorded in the 24 h beginning at 0000.

2.1.2. Forecasts

Daily forecast values for model grids corresponding to each weather station were obtained from the ECMWF atmospheric model high resolution 10 day 0000 forecast. This forecast had a horizontal grid resolution of approximately 25 km over Ireland until 26 January 2010, when it was increased to approximately

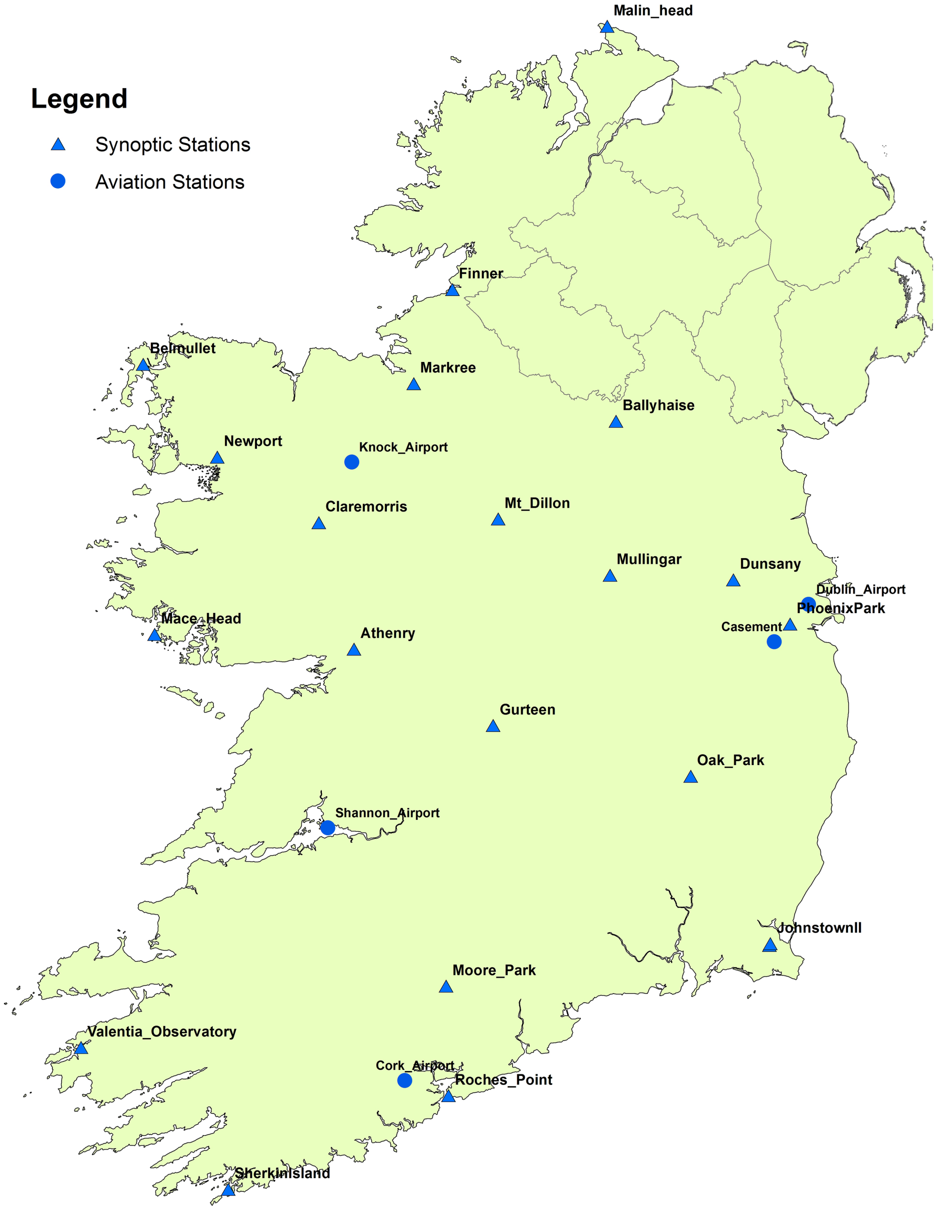

Figure 1. Map of Ireland showing the locations of the 25 Met Éireann synoptic stations. Observations are taken manually at the manual stations, which are located at airports. [Colour figure can be viewed at wileyonlinelibrary.com].

16 km. The data extracted were mapped onto a latitude/longitude grid ($0.125° \times 0.125°$) for ease of comparison of different model versions operating on different grid resolutions over the years. At some coastal stations, the grid box containing the station was describing a sea area in the model before the update and a land area afterwards. It was found that this could lead to biases in the model. In order to overcome issues associated with the improvement of land–sea boundaries in the updated model, the forecast value used was that from the nearest land grid box output by the model. The ECMWF model predicts the minimum and maximum 2 m air temperatures in every consecutive 6 h interval from +0 h (0 h after the model run) to +240 h. The daily minimum and maximum temperatures were the lowest and highest, respectively, of the forecasts for the day in question. Each ECMWF model run gave 2 m air temperature and soil temperature forecasts at 3 h intervals from +0 h to +144 h and at 6 h intervals from +144 h to +240 h. The model forecast mean temperature was computed as the mean of the 2 m air temperature forecast values for the day. The day 1 to day 10 ECMWF forecasts for the four soil temperature forecast parameters (STL1 (0–7 cm), STL2 (7–28 cm), STL3 (28–100 cm) and STL4 (100–289 cm)) were obtained similarly. For example, the day 2 forecast for 2 January was the mean of the +24, 27, 30, 33, 36, 39, 42 and 45 h forecasts from 1 January whereas the day 7 forecast for 7 January was the mean of the +150, 156, 162 and 168 h forecasts from 1 January. Day 1 to day 10 rainfall forecasts were output directly by the ECMWF forecast model.

Monthly mean climatological values for rainfall and maximum, minimum and mean temperature between 1981 and 2010 were obtained at each station (Walsh, 2012). The monthly mean climatological observation at each station was forecast for every day in the month to generate mean climatology forecasts. Persistence forecasts were generated by forecasting the observation of the day preceding the forecast generation (Joliffe and Stephenson, 2011). For example, day 10 persistence forecasts were the observed weather conditions of 10 days ago.

2.2. Accuracy assessment of direct model output forecasts

A range of verification statistics were used to assess forecast accuracy and to identify biases for each weather variable. The statistics included mean systematic bias:

$$\mathrm{MSB} = \frac{\sum_{i=1}^{n} \left(f_i - o_i\right)}{n}$$

root mean squared error:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(f_i - o_i\right)^2}{n}}$$

and mean absolute error:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} \left|f_i - o_i\right|}{n}$$

where o_i is the i^{th} observation, f_i is the i^{th} forecast and n is the total number of observations used in the calculation (Joliffe and Stephenson, 2011). Standard deviation was used to assess the variability of the forecast and observed values. The forecasts were analysed at each individual forecast period from day 1 to day 10, or using all forecast periods combined if the statistic of interest did not vary substantially across forecast periods, to give a general description of forecast quality. The trends were examined by individual year, season, month and station, as well as across all 7 years of data and 25 stations. When calculating yearly, monthly or seasonal statistics, limits were imposed on the number of missing values permitted: a minimum requirement of 183 values present in a year, 61 in a season and 22 in a month was set (yearly missing values at each station are shown in Table S1). To assess the skill of a forecast, its accuracy was compared with that of a mean climatology or a persistence forecast.

2.3. Bias correction

Seven bias correction (BC) techniques were tested: yearly, seasonal and monthly BC ((1), (2) and (3), respectively), mean and variance (MAV) BC (4), regression model BC (5), regression model BC by station (6) and the composite (COM) post-processing method (7). Each BC method used a cross-validation approach; values from the period to be bias corrected were excluded from the data used to inform the bias corrections (Joliffe and Stephenson, 2011). It was deemed reasonable to use data from before and after the target year (the year in which the forecasts were to be bias corrected) (Joliffe and Stephenson, 2011). The details for each bias correction method are as follows:

1. year BC: one year of forecasts and observations at one station were selected as the target year data. The MSB for all data at the station excluding the target year was calculated. This MSB was subtracted from ECMWF forecasts in the target year to obtain the yearly bias corrected forecasts (Joliffe and Stephenson, 2011). This was repeated for each year at each station. Data from all forecast periods were used to do these bias corrections provided the MSB values did not vary much across forecast periods;
2. season BC: individually for each station, the MSB in the season of interest for all data outside the target year was subtracted from target year ECMWF forecasts for the particular season. The seasons were defined as spring (March, April, May), summer (June, July, August), autumn (September, October, November) and winter (December, January, February);
3. month BC: individually for each station, the MSB in the month of interest for all years outside the target year was subtracted from target year ECMWF forecasts for the particular month;
4. mean and variance (MAV) BC: separately for each forecast period, forecasts were bias corrected by scaling them to have the same mean and variance as the observations from the month being bias corrected (obs) using the cross-validation approach (Sweeney *et al.*, 2011). The bias correction involved two steps:

$$\mathrm{fBC1}_i = f_i \left(\sigma_{\mathrm{obs}}/\sigma_f\right)$$

$$\mathrm{fBC2}_i = \mathrm{fBC1}_i + \mu_{\mathrm{obs}} - \mu_{\mathrm{fBC1}}$$

where f_i is the i^{th} forecast of interest, σ_{obs} and σ_{f} are the standard deviations of the observations and forecasts of interest, respectively, $\mathrm{fBC1}_i$ is the i^{th} forecast scaled for standard deviation, μ_{obs} and μ_{fBC1} are the means of the observations and forecasts (adjusted for standard deviation), respectively, and $\mathrm{fBC2}_i$ is the i^{th} MAV bias corrected forecast;

5. model BC: using a linear regression approach (Acharya *et al.*, 2013), separately for each forecast, forecasts were used as a predictor of observations:

$$E\left[y\right] = \beta_0 + \beta_1 f + \delta_j + \alpha_k$$

where y is the observation, f is the ECMWF forecast, δ_j is a categorical month-specific term, $j = 1, ..., 12$, α_k is a categorical station-specific term, $k = 1, ..., 25$. The parameters of the model were estimated using the data outside the target year. The bias corrected forecasts for the target year were obtained by predicting from the estimated model;

6. model BC by station: the model BC approach described in (5) was performed on each station separately with just the forecast and monthly terms as predictors, and
7. composite (COM) post-processing: any subset (of size n) of the six bias corrected forecasts and the ECMWF forecast can be combined to make a composite (COM) forecast (Sweeney and Lynch, 2011). For each day, the cross-validation approach was employed separately on each station using data from the same month as that being bias corrected to obtain an MAE value (MAE_j) for each forecast (f_j) in the subset being used. The following equations were used to calculate the COM forecast for a particular day:

$$\mathrm{NUM} = \prod_{j=1}^{n} \mathrm{MAE}_j$$

$$\mathrm{COM} = \sum_{j=1}^{n} \frac{\mathrm{NUM}^* f_j}{\mathrm{MAE}_j \sum_{j=1}^{n} \left(\mathrm{NUM}/\mathrm{MAE}_j\right)}$$

Any negative rainfall forecasts given by the bias correction methods were set to zero. All bias corrected forecasts were re-assessed for accuracy using the methods described in Section 2.2.

A moving window approach using cross-validation, in which only data from the 31 days centred on the day to be bias corrected were used in the bias correction, was applied to methods (3), (4), (6) and (7) to assess whether there was any improvement in forecast accuracy (Stensrud and Yussouf, 2005). When using this approach, there was generally no overall improvement on the original bias correction methods. Since it was much more computationally expensive and therefore less useful for operational purposes, it is not presented.

3. Results and discussion

3.1. Accuracy assessment of ECMWF forecasts

The accuracy of the ECMWF minimum, mean and maximum air temperature forecasts each decreased linearly as the forecast period increased (Table S2); for example, the minimum temperature RMSE for data from all stations rose from 1.51 to 2.25 to 3.44 °C for forecast periods 1, 5 and 10, respectively. Comparing yearly RMSE values at each individual station and for each air temperature variable identified that persistence forecasts outperformed day 7 ECMWF forecasts over 75% of the time. This indicates that air temperature forecasts of forecast periods longer than a week were not any more useful than low skill forecasts. A number of stations showed considerable within year variability in maximum temperature RMSE (Figure 2).

The highest yearly RMSE value for minimum temperature was recorded in 2010: 1.74 °C at forecast period 1 across all stations in comparison to a mean of 1.47 °C for the other 6 years of data, and this trend was evident in all forecast periods. In 2010, 20.1% of the daily minimum temperature observations were less than 0 °C compared to 7.0% on average across the other 6 years, and these extremes were poorly predicted by the ECMWF model, explaining the poor forecasting performance in 2010. Poor forecasts of freezing temperatures are unlikely to hamper the prediction of grass growth provided the observations remain below 5 °C since this is the lower threshold for grass growth (Hopkins, 2000). Minimum, maximum and mean 2 m air temperature MSB trends were generally constant across forecast periods at each station indicating that the air temperature variables should respond positively to bias correction (minimum temperature, Figure S1). Figure S1 shows pronounced yearly differences in MSB at Belmullet and Mace Head. This is probably due to the horizontal improvements in the model in 2010 better defining the land−sea boundary for these coastal stations. Otherwise, there were no noticeable differences in systematic biases before and after the ECMWF model update (Figures 3, S1 and S2). Minimum temperature displayed a larger range of MSB values across

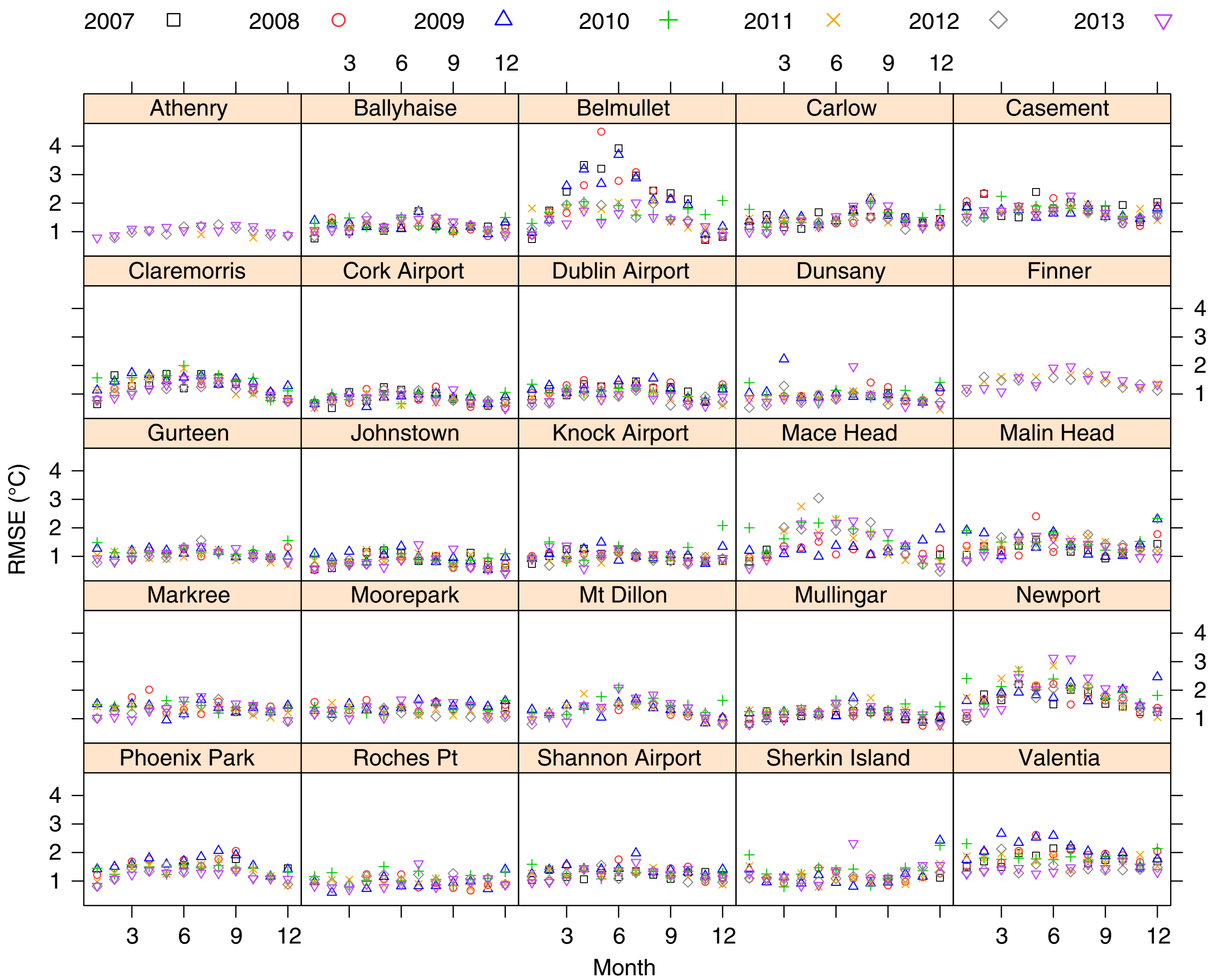

Figure 2. Monthly maximum temperature root mean square error (RMSE) (°C) of day 1 European Centre for Medium-Range Weather Forecasts forecasts for each of the Met Éireann synoptic stations in each year from 2007 to 2013. [Colour figure can be viewed at wileyonlinelibrary.com].

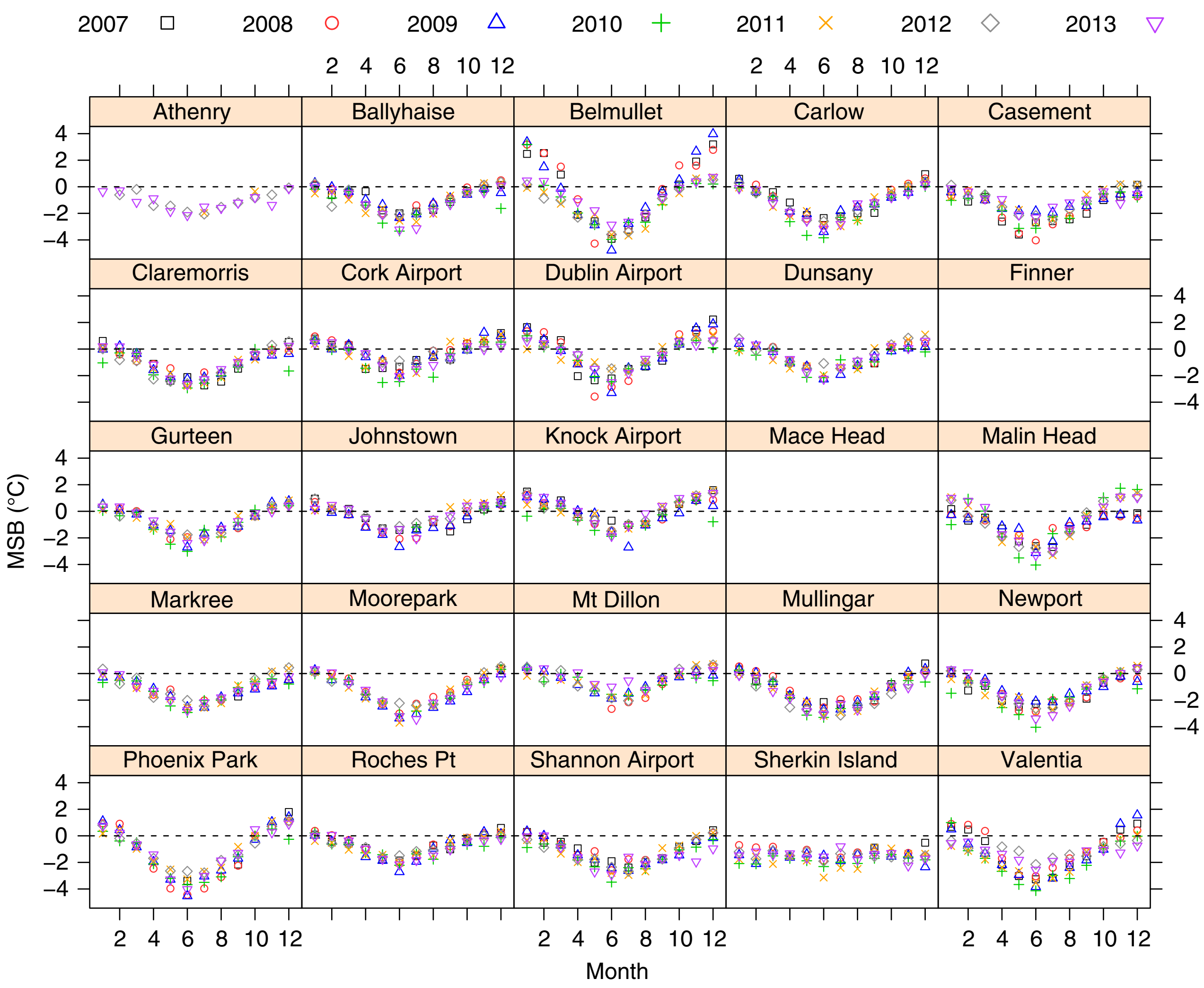

Figure 3. Monthly mean systematic bias (MSB) (°C) of STL1 (0–7 cm) forecasts and 5 cm soil temperature observations for day 1 forecasts in each year from 2007 to 2013 at 23 of the 25 Met Éireann synoptic stations. No soil observations were available at Finner or Mace Head. [Colour figure can be viewed at wileyonlinelibrary.com].

stations than the other air temperature variables (between −1.86 and 1.00 °C for forecast period 1).

Soil temperature observations at each depth (5, 10, 20, 30, 50 and 100 cm) may not be most accurately forecasted by the forecast range into which they fall (STL1 0–7, STL2 7–28, STL3 28–100 and STL4 100–289 cm). When data from all stations and all forecast periods were included, the RMSE statistic identified that depths 5, 20, 50 and 100 cm were best forecast by their corresponding STL range but that 10 cm was best predicted by STL1 and 30 cm was best predicted by STL2; the most accurate range was used as the ECMWF forecast in each case. For all six soil temperature depths, the MSB was reasonably consistent across forecast periods at each station. However, there was a strong within year difference in MSB, with the forecast consistently under-estimating the observations in summer at all stations (shown for 5 cm, Figure 3). This could be attributed to the fact that the forecasts are not directly predicting the observation depths; rather, they predict ranges, and therefore exhibit strong systematic biases. Bias correction should successfully reduce these biases (Joliffe and Stephenson, 2011).

Short term ECMWF soil temperature forecasts performed well, with RMSE values below 1.71 °C at every depth at forecast period 1 when data from all stations were considered (Figure 4(a)). As expected, at all depths forecast accuracy declined as forecast period increased; for example, the RMSE for 10 cm soil temperatures across all stations was 1.62 °C at forecast period 1 and increased to 2.46 °C at forecast period 10 (Figure 4(a)). At greater depths, the forecasts were usually more accurate and the differences in RMSE values between forecast periods 1 and 10 were smaller (Figure 4(a)). This can be explained by the tendency of greater soil temperature depths to be less sensitive to changes in air temperature, making them more homogeneous over time and therefore easier to predict. For example, at Dublin Airport between 1 June and 31 August 2012, the correlation between daily 5 cm soil temperature observations and corresponding 2 m mean air temperature observations was 0.83; the correlation between daily 100 cm soil temperature observations and corresponding 2 m mean air temperature observations was 0.52. There was a gradual overall warming effect at 100 cm, while there were persistent increases and decreases in 5 cm soil temperature over the same period (Figure S3). Since depths between 0 and 10 cm are most important for grass growth (Hurtado-Uria *et al.*, 2013a), these results suggests that, in the presence of accurate air temperature forecasts, soil temperature forecasts may not contribute appreciably further to the ability of a model to predict grass growth. However, ECMWF soil

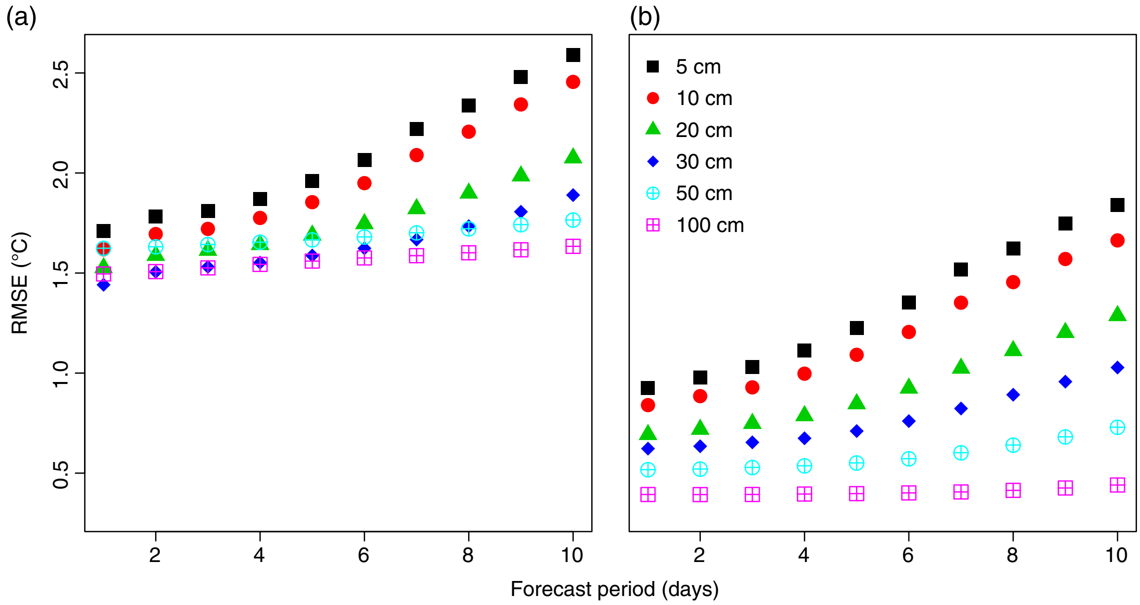

Figure 4. Root mean square error (RMSE) (°C) of (a) European Centre for Medium-Range Weather Forecasts and (b) model bias corrected by station soil temperature forecasts and observations at six depths across all stations at forecast periods 1–10. [Colour figure can be viewed at wileyonlinelibrary .com].

temperature forecasts have been shown for the first time to give accurate predictions of observations at a range of depths up to 100 cm; soil temperature forecasts have been verified only at 5 cm previously (Albergel *et al.*, 2015). Although the soil temperature RMSE values were low, soil persistence forecasts generally outperformed ECMWF forecasts: 83.9% of yearly RMSE values for day 1 10 cm forecasts at individual stations were higher than equivalent persistence RMSE values. This rose to 100% for day 1 100 cm forecasts.

Day 5 climatology forecasts gave lower RMSE values than equivalent ECMWF forecasts in 64.9% of years at individual stations for daily rainfall, while day 7 climatology forecasts were better than ECMWF forecasts in all cases. Thus, after 7 days, ECMWF rainfall forecasts are no more useful in a grass growth model than mean climatology values. The yearly rainfall MSB values were often non-zero and were quite consistent within station across all forecast periods indicating that there were systematic biases in ECMWF rainfall forecasts that could possibly be eliminated by bias correction (Figure S2). The MSB values in northwest Ireland at Belmullet, Finner and Mace Head were higher than at the other stations: the ECMWF model over-predicted rainfall at these stations more than other stations. Perhaps due to their proximity to the Atlantic Ocean, high rainfall values were predicted here which were not reflected by the observations. It was also the case that the decrease in forecast accuracy of maximum temperature during the summer was most extreme at coastal stations such as Belmullet, Mace Head and Newport (Figure 2). The topography of the station could be different to the general area described by the ECMWF forecast. It could also be that the grid boxes used to forecast these stations are overly marine influenced. These biases should be reduced after bias correction.

As one might expect, high rainfall observations led to a decrease in forecast accuracy (Roberts and Lean, 2008). For example, the MAE for day 5 forecasts from all stations with observations between 40 and 50 mm inclusive was 31.3 mm compared to an MAE of 2.2 mm for all observations between 0 and 10 mm inclusive. The MSB for the day 5 forecasts from all stations with observations between 40 and 50 mm inclusive was −31.3 mm, so these high rainfall events were all under-predicted at a lead time of 5 days. The MSB at Valentia Observatory in 2009 was lower than all other years at that station, particularly for day 5 forecasts and longer (Figure S2). This was probably due to unusually high summer rainfall values causing the forecast to under-predict the observations; the total rainfall at Valentia in summer 2009 was 619.7 mm, compared to a mean total summer rainfall of 351.0 mm at the same station in the other 6 years. Inaccurate ECMWF forecasts of high rainfall events during the growing season can cause poor grazing management, resulting in reduced use of available grass and subsequently impacting negatively on future grass growth. Due to poor use, farms might be forced to supplement with concentrate, increasing variable costs. There are also risks of a decline in ground conditions and poor use of fertilizer (Shalloo *et al.*, 2004).

3.2. Bias correction results

Model BC by station gave the greatest reductions in RMSE (compared to ECMWF forecasts) for almost all of the air and soil temperature variables for both forecast periods 1 and 10 (Table 1). It generally performed better than model BC since it does not assume common month parameters for each station and was therefore more effective at eliminating monthly biases (such as those in Figure 3). Although COM BC gave similar and sometimes slightly higher reductions, model BC by station is recommended as it requires less computation time. Model BC, model BC by station and COM BC reduced the bias at every station to magnitudes of 0.05 °C or less for all air temperature variables at all forecast periods. Although MAV BC reduced bias in forecasts, it did not improve the RMSE in general and was generally the worst bias correction method (Table 1). The relatively high percentage reductions in RMSE for predictions of soil temperatures could be because there were often large systematic biases attributable to the fact that the forecasts predicted ranges rather than directly forecasting the observations. Model BC by station was also generally the most effective method at reducing

Table 1. The percentage reduction in RMSE resulting from the different bias correction methods compared to ECMWF forecasts across all stations for each of the rainfall and air and soil temperature forecasts at forecast periods 1 and 10.

	Year BC	Season BC	Month BC	MAV BC	Model BC	Model BC by station	COM BC
Min temp day 1	8.6	8.7	8.2	−42.7	9.5	12.7	12.3
Min temp day 10	1.6	1.6	1.5	−12.5	13.9	13.8	14.0
Max temp day 1	26.8	27.5	27.4	−49.0	27.1	29.0	28.8
Max temp day 10	5.5	5.7	5.5	−11.7	15.3	15.6	15.7
Mean temp day 1	19.5	21.0	21.0	−102.4	21.0	24.8	24.5
Mean temp day 10	2.4	2.7	2.7	−14.8	13.0	13.2	13.3
5 cm day 1	18.7	38.0	43.2	−2.6	42.1	45.9	45.6
5 cm day 10	11.2	16.7	17.7	−0.8	28.0	29.0	28.9
10 cm day 1	22.3	38.1	42.0	−4.2	43.7	48.1	47.8
10 cm day 10	13.0	17.3	17.9	1.3	31.1	32.3	32.2
20 cm day 1	26.2	43.8	48.4	1.1	48.3	54.6	54.2
20 cm day 10	19.4	25.4	26.4	3.7	35.8	38.0	37.8
30 cm day 1	27.5	34.2	35.0	0.9	48.1	56.8	55.9
30 cm day 10	21.8	24.6	24.9	8.2	42.1	45.6	45.2
50 cm day 1	33.5	56.8	64.0	22.5	57.4	68.1	67.2
50 cm day 10	34.4	51.1	55.6	20.8	51.5	58.7	58.0
100 cm day 1	47.0	54.4	55.5	30.6	59.4	73.7	72.4
100 cm day 10	47.3	54.6	56.0	35.2	61.0	73.0	71.8
Rainfall day 1	3.1	3.3	3.1	−12.7	7.8	9.0	8.9
Rainfall day 10	1.2	1.3	1.3	−1.5	23.6	23.5	23.6

The COM BC included in this table is the composite of model BC and model BC by station since it performed best of all of the COM BC forecast combinations. The best bias correction technique, identified by the highest RMSE reduction, is highlighted by grey shading in each row.

the RMSE (Table 1) and eliminating bias in rainfall forecasts: across all forecast periods the MSB values at individual stations ranged from −0.00 to 0.06 mm. This was compared to methods such as month BC which gave forecasts that over-predicted the daily rainfall on average by 0.82 mm at Mace Head.

3.3. Accuracy assessment of best bias corrected forecasts

After model BC by station the day 1 MSB values were 0.002, 0.000 and 0.001 °C for maximum, minimum and mean temperature, respectively, across all stations. In comparison, the ECMWF day 1 MSB values were −0.770, 0.039 and −0.304 °C, respectively. Model BC by station reduced the range in MSB for day 1 maximum temperature ECMWF forecasts across stations from between −1.57 and 0.43 °C to between −0.004 and 0.006 °C. The RMSE values for the maximum temperature ECMWF forecast at forecast period 1 ranged from 0.86 to 1.96 °C at individual stations, and were reduced to range between 0.82 and 1.19 °C by model BC by station, with forecast improvements at almost all stations. Day 1 air temperature forecasts experienced decreases in RMSE of up to 49% at individual stations. Yearly RMSE values for day 7 persistence forecasts were preferable to day 7 model BC by station forecasts for maximum, minimum and mean temperature 47.1%, 48.4% and 58.0% of the time, respectively, a substantial improvement on the ECMWF forecasts. Thus, model BC by station improved the ECMWF air temperature forecasts, indicating that they should be of practical use in a grass growth model up to approximately a week in advance. Model BC by station did not generally improve the accuracy of the imprecise ECMWF air temperature forecasts of the extreme low temperature observations in 2010. Regression model approaches are not usually good at predicting extremes (Allen and DeGaetano, 2001; Zhai *et al.*, 2005), and the forecasts used as predictors in the model were not accurate to begin with. However, as noted in Section 3.1, this may not be too important provided forecasts for extremely low observations are below 5 °C.

Model BC by station soil temperature forecasts had RMSE values below 0.93 °C for day 1 forecasts and below 1.84 °C for day 10 forecasts at all depths (Figure 4(b)). The forecast accuracy was generally worst at 5 cm, improving at each subsequent depth (Figure 4(b)). Due to their homogeneity, the systematic biases were more constant at greater depths, meaning bias correction usually gave larger improvements at these depths. While ECMWF 30 cm soil temperature forecasts had lower RMSE values than all other depths at some forecast periods (Figure 4(a)), the systematic bias was not as large as at other depths. As a result, they did not have the lowest RMSE values after model BC by station. The reductions in RMSE from the ECMWF forecasts were similar across all forecast periods but tended to be higher in summer when the MSB values were of higher magnitude (Figure 3). Model BC by station soil temperature forecasts for depths such as 5 and 10 cm were generally more useful than persistence at forecast periods of less than a week, while forecasts at depths such as 50 and 100 cm were not. For example, day 5 model BC by station forecasts for 10 cm gave lower RMSE values than persistence in 70.5% of years at individual stations (a large improvement on ECMWF forecasts), but day 1 100 cm persistence forecasts outperformed equivalent model BC by station forecasts in 99.3% of cases. The accuracy of forecasts at the 5 and 10 cm depths is useful since the soil temperature at shallow depths will be the most influential in determining grass growth.

Model BC, model BC by station and the COM BC approaches gave the highest RMSE reductions for rainfall at both forecast periods 1 and 10 (Table 1); and although day 7 model BC by station forecasts outperformed climatology forecasts in 81.6% of cases (a large improvement on ECMWF forecasts), they may not be useful in practice. Because rainfall is difficult to forecast and predictive rainfall accuracy diminishes rapidly with forecast period, these bias correction methods tend to forecast conservative estimates close to the mean rainfall. Although these methods usually resulted in lower RMSE values than the original forecast, they did not follow the trend of the observations as closely as other forecasts (see Ballyhaise for example, Figure 5). At Ballyhaise in January 2012, the RMSE values of the day 10 ECMWF, month BC and model BC by station forecasts were 5.56, 5.60 and 4.40 mm, respectively. The standard deviation of

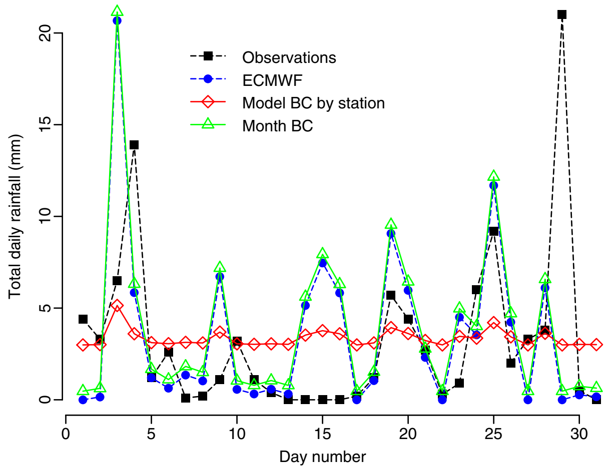

Figure 5. Daily observed rainfall (squares), day 10 European Centre for Medium-Range Weather Forecasts (ECMWF) forecast (circles, root mean square error (RMSE) = 5.56 mm), day 10 monthly bias correction (BC) forecast (triangles, RMSE = 5.60 mm) and day 10 model bias corrected by station forecast (diamonds, RMSE = 4.40 mm) at Ballyhaise in January 2012. [Colour figure can be viewed at wileyonlinelibrary.com].

the observations was 4.55 mm, and it was 4.51 mm for both the ECMWF forecast and the monthly BC forecasts. This fell to 0.47 mm for the model BC by station forecast. Thus, bias correction approaches such as monthly BC were preferable even though the reductions in RMSE were minimal (Table 1). Although most bias correction methods gave reductions in RMSE when using all observations, the reductions usually only occurred for low rainfall observations. As highlighted in Section 3.1, high rainfall values were predicted poorly by ECMWF forecasts. When daily observations greater than 10 mm were considered, none of the methods gave substantial reductions in RMSE at any forecast period. A possible reason for this is that, if the initial forecasts are generally imprecise, bias corrections cannot improve the forecasts. The inaccuracy of rainfall forecasts and their bias corrections is problematic since rainfall values greater than 10 mm are likely to be more influential on grass growth than lower rainfall values. Model BC by station air and soil temperature forecasts did not have the same problem with standard deviation as the rainfall forecasts, often giving standard deviation values closer to those of the observations than ECMWF forecasts.

3.4. Future work and implications for Irish agriculture

Future work will involve developing a grass growth model to simulate on-farm growth in Ireland that will incorporate weather forecast predictors. To date, grass growth prediction models have tested a wide range of predictors (e.g. soil type, fertilizer application, grazing events) but have only considered historical weather data (Hurtado-Uria *et al.*, 2013b). The verification of the ECMWF forecasts suggests that they will accurately predict weather conditions in the coming 5–7 days, and therefore will be valuable as predictors of grass growth. The best bias corrected forecasts were usually more accurate than the ECMWF forecasts. If they prove to give grass growth predictions preferable to the ECMWF forecasts, methods of applying bias corrections across the island at farm level could be useful. If the rainfall component of the grass growth model is not sufficient, the use of probabilistic rainfall forecasts could be explored (Hamill *et al.*, 2004; Wilks, 2009). The fact that soil temperature forecasts are useful will not only benefit grassland management, it also has positive implications for Irish crop production. Climatology and persistence forecasts both regularly outperforming ECMWF and bias corrected forecasts after 7 days suggests that a combination of these forecasts with recent observed weather and long term climate would work best when forecasting weather for on-farm grass growth prediction. Online resources allowing farmers to input local parameters and obtain grass growth predictions accounting for weather forecasts would increase grass use and on-farm profitability.

4. Conclusions

Forecast verification was undertaken to determine how accurate European Centre for Medium-Range Weather Forecasts (ECMWF) operational deterministic forecasts are at Irish synoptic weather stations. Air temperature forecasts were accurate for all forecast periods but often gave higher root mean squared error (RMSE) values than persistence after 7 days. ECMWF soil temperature forecasts have not been extensively verified internationally but here it is shown for the first time that in Ireland they can predict observations well, with prediction accuracy increasing as depth beneath the surface increases. This was due to the more conservative rate of change in temperatures at greater depths making them easier to model. However, this slow rate of change also meant that persistence forecasts gave lower RMSE values than ECMWF forecasts in almost all years and at all forecast periods. ECMWF rainfall forecasts showed skill for forecast periods of 6 days or less. After this climatology forecasts tended to give lower RMSE values.

Systematic biases were evident for all of the weather variables examined; their values varied in magnitude and orientation across locations. Thus, seven approaches to eliminate them were proposed. For air and soil temperatures, model bias correction by station was generally the most effective method of bias correction. Soil temperature forecasts were greatly improved by bias correction since the ECMWF forecasts exhibited strong systematic biases at many depths. Monthly bias correction was recommended for ECMWF rainfall forecasts, although none of the bias correction methods assessed gave large improvements because they did not improve forecast quality for observations greater than 10 mm, and some methods gave forecasts with much lower standard deviations than the observations.

Overall, it has been shown that weather forecasts have the potential to contribute to grass growth prediction for up to 1 week in Ireland. This knowledge can contribute to better efficiency in on-farm management of grass resources and help improve primary productivity and profit for Irish agriculture.

Acknowledgements

This work was funded by the Irish Department of Agriculture, Food and the Marine Research Stimulus Fund (RSF 11/S/132). The first author is in receipt of a Teagasc Walsh Fellowship.

Supporting information

The following material is available as part of the online article:

Table S1. Number of missing values at each station in each year for the 10 weather variables examined.
Table S2. RMSE values (°C) for minimum temperature, maximum temperature and mean temperature at forecast periods 1–10 for all years at all stations.

Figure S1. Yearly minimum temperature MSB of ECMWF forecasts for each forecast period from day 1 to day 10 for each of the Met Éireann synoptic stations.

Figure S2. Yearly total daily rainfall MSB of ECMWF forecasts for each forecast period from day 1 to day 10 for each of the Met Éireann synoptic stations.

Figure S3. 1 June to 31 August 2012 daily soil temperature observations at 5 cm and 100 cm and forecasts for STL1 and STL3 at Dublin Airport. Daily mean 2 m air temperature observations for the same time period are also shown.

References

Acharya N, Chattopadhyay S, Mohanty U, Dash S, Sahoo L. 2013. On the bias correction of general circulation model output for Indian summer monsoon. *Meteorol. Appl.* **20**: 349–356.

Albergel C, Dutra E, Munoz-Sabater J, Haiden T, Balsamo G, Beljaars A, *et al.* 2015. Soil temperature at ECMWF: an assessment using ground-based observations. *J. Geophys. Res. Atmos.* **120**: 1361–1373.

Allen R, DeGaetano A. 2001. Estimating missing daily temperature extremes using an optimized regression approach. *Int. J. Climatol.* **21**: 1305–1319.

Aligne T, McNally A, Dee D. 2007. Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. R. Meteorol. Soc.* **133**: 631–642.

Boi P. 2004. A statistical method for forecasting extreme daily temperatures using ECMWF 2-m temperatures and ground station measurements. *Meteorol. Appl.* **11**: 245–251.

Brereton A. 1995. Regional and year-to-year variation in production. In *Irish Grasslands – Their Biology and Management*, Jeffrey DW, Jones MB, McAdam JH (eds). Royal Irish Academy: Dublin, Ireland; 12–22.

Burke JI, Brereton AJ, O'Kiely P, Schulte R. 2004. Weather and crop production. In *Climate, Weather and Irish Agriculture*, 2nd edn, Keane T, Collins JF (eds). AGMET: Dublin, Ireland.

DAFM. 2015. *Food Wise 2025*. Department of Agriculture Food and the Marine: Dublin, Ireland. http://www.agriculture.gov.ie/media/migration/foodindustrydevelopmenttrademarkets/agri-foodandtheeconomy/foodwise2025/report/FoodWise2025.pdf (accessed 11 September 2017).

DAFM. 2016. *Grass and Clover, Recommended List Varieties for Ireland 2016*. Department of Agriculture, Food and the Marine: Dublin, Ireland. http://www.agriculture.gov.ie/media/migration/publications/2016/GrassWhiteCloverRecom2016080216.pdf (accessed 11 September 2017).

Dillon, P. 2011. The Irish dairy industry – Planning for 2020. *National Dairy Conference 2011*, 15 November 2011, Cork, Ireland; 1–24. http://www.teagasc.ie/media/website/publications/2011/Dairy_Conference_Proceedings_2011.pdf (accessed 11 September 2017).

Ebert E, McBride J. 2000. Verification of precipitation in weather systems: determination of systematic errors. *J. Hydrol.* **239**: 179–202.

Fan Y, van den Dool H. 2011. Bias correction and forecast skill of NCEP GFS ensemble week-1 and week-2 precipitation, 2-m surface air temperature, and soil moisture forecasts. *Weather Forecasting* **26**: 355–370.

Finneran E, Crosson P, O'Kiely P, Shalloo L, Forristal PD, Wallace M. 2010. Simulation modelling of the cost of producing and utilizing feeds for ruminants on Irish farms. *J. Farm. Manage.* **14**: 95–116.

Frame J. 1992. *Improved Grassland Management*, 1st edn. Farming Press Books: Ipswich, UK; 1–351.

Galanis G, Anadranistakis M. 2002. A one-dimensional Kalman filter for the correction of near surface temperature forecasts. *Meteorol. Appl.* **9**: 437–441.

Hamill T, Hagedorn R, Whitaker J. 2008. Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part II: precipitation. *Mon. Weather Rev.* **136**: 2620–2632.

Hamill T, Whitaker J, Wei X. 2004. Ensemble reforecasting: improving medium-range forecast skill using retrospective forecasts. *Mon. Weather Rev.* **132**: 1434–1447.

Harrison D, Driscoll S, Kitchen M. 2000. Improving precipitation estimates from weather radar using quality control and correction techniques. *Meteorol. Appl.* **7**: 135–144.

Hopkins A. 2000. *Grass : Its Production and Utilization*, 3rd edn. Published for the British Grassland Society by Blackwell Science: Oxford; Malden, MA.

Huang L, Isaac G, Sheng G. 2012. Integrating NWP forecasts and observation data to improve nowcasting accuracy. *Weather Forecasting* **27**: 938–953.

Hurtado-Uria C, Hennessy D, Shalloo L, O'Connor D, Delaby L. 2013a. Relationships between meteorological data and grass growth over time in the south of Ireland. *Irish Geogr.* **46**: 175–201.

Hurtado-Uria C, Hennessy D, Shalloo L, Schulte R, Delaby L, O'Connor D. 2013b. Evaluation of three grass growth models to predict grass growth in Ireland. *J. Agr. Sci.* **151**: 91–104.

Joliffe IT, Stephenson DB. 2011. Deterministic forecasts of continuous variables. In *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 2nd edn. John Wiley & Sons: Hoboken, NJ; 97–114.

Louka P, Galanis G, Siebert N, Kariniotaki G, Katsafados P, Pytharoulis I, *et al.* 2008. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *J. Wind Eng. Ind. Aerod.* **96**: 2348–2362.

Mass C, Baars J, Wedam G, Grimit E, Steed R. 2008. Removal of systematic model bias on a model grid. *Weather Forecasting* **23**: 438–459.

O'Donovan M. 2000. The relationship between the performance of dairy cows and grassland management on intensive dairy farms in Ireland, PhD thesis, University College Dublin.

O'Donovan M, Lewis E, O'Kiely P. 2011. Requirements of future grass-based ruminant production systems in Ireland. *Irish J. Agr. Food Res.* **50**: 1–21.

Roberts N. 2008. Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteorol. Appl.* **15**: 163–169.

Roberts N, Lean H. 2008. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Weather Rev.* **136**: 78–97.

Shalloo L. 2009. Milk production costs – can we compete? *Teagasc National Dairy Conference*, 18 November 2009, Cork, Ireland; 19–38.

Shalloo L, Dillon P, O'Loughlin J, Rath M, Wallace M. 2004. Comparison of a pasture-based system of milk production on a high rainfall, heavy-clay soil with that on a lower rainfall, free-draining soil. *Grass Forage Sci.* **59**: 157–168.

Stensrud D, Yussouf N. 2005. Bias-corrected short-range ensemble forecasts of near surface variables. *Meteorol. Appl.* **12**: 217–230.

Sun B, Yu L, Weller R. 2003. Comparisons of surface meteorology and turbulent heat fluxes over the Atlantic: NWP model analyses versus moored buoy observations. *J. Clim.* **16**: 679–695.

Sweeney C, Lynch P. 2011. Adaptive post-processing of short-term wind forecasts for energy applications. *Wind Energy* **14**: 317–325.

Sweeney C, Lynch P, Nolan P. 2011. Reducing errors of wind speed forecasts by an optimal combination of post-processing methods. *Meteorol. Appl.* **20**: 32–40.

Taylor A, Leslie L. 2005. A single-station approach to model output statistics temperature forecast error assessment. *Weather Forecasting* **20**: 1006–1020.

Thorvaldsson G, Bjornsson H, Hermannsson J. 2004. The influence of weather on early growth rate of grasses. *Icelandic Agric. Sci.* **16**: 65–73.

Vannitsem S, Hagedorn R. 2011. Ensemble forecast post-processing over Belgium: comparison of deterministic-like and ensemble regression methods. *Meteorol. Appl.* **18**: 94–104.

Vrac M, Friederichs P. 2015. Multivariate-intervariable, spatial, and temporal-bias correction. *J. Clim.* **28**: 218–237.

Walsh S. 2012. *A Summary of Climate Averages 1981–2010 for Ireland*. Climatological Note No.14. Met Éireann: Dublin.

Wilks D. 2009. Extending logistic regression to provide full-probability-distribution MOS forecasts. *Meteorol. Appl.* **16**: 361–368.

Yussouf N, Stensrud D. 2007. Bias-corrected short-range ensemble forecasts of near-surface variables during the 2005/06 cool season. *Weather Forecasting* **22**: 1274–1286.

Zhai P, Zhang X, Wan H, Pan X. 2005. Trends in total precipitation and frequency of daily precipitation extremes over China. *J. Clim.* **18**: 1096–1108.