Journal of the Royal Statistical Society Series C: Applied Statistics, 2024, **00**, 1–49
https://doi.org/10.1093/jrsssc/qlae044

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS
Journal of the Statistics Society
Series C
Applied Statistics

Original Article

A noisy-input generalized additive model for relative sea-level change along the Atlantic coast of North America

Maeve Upton[1], Andrew Parnell[1], Andrew Kemp[2], Erica Ashe[3], Gerard McCarthy[4] and Niamh Cahill[5]

[1]Hamilton Institute, Department of Mathematics & Statistics, ICARUS, Maynooth University, Maynooth, Ireland
[2]Department of Earth and Climate Sciences, Tufts University, Boston, MA, USA
[3]Department of Earth and Planetary Sciences, Rutgers University, New Brunswick, NJ, USA
[4]ICARUS, Maynooth University, Maynooth, Ireland
[5]Department of Mathematics and Statistics, ICARUS, Maynooth University, Maynooth, Ireland

Address for correspondence: Maeve Upton, Department of Mathematics & Statistics, Maynooth University, Room 207 Logic House, South Campus, Maynooth, Co. Kildare, Ireland. Email: uptonmaeve010@gmail.com

Abstract

We propose a Bayesian, noisy-input, spatial–temporal generalized additive model to examine regional relative sea-level (RSL) changes over time. The model provides probabilistic estimates of component drivers of regional RSL change via the combination of a univariate spline capturing a common regional signal over time, random slopes and intercepts capturing site-specific (local), long-term linear trends and a spatial–temporal spline capturing residual, non-linear, local variations. Proxy and instrumental records of RSL and corresponding measurement errors inform the model and a noisy-input method accounts for proxy temporal uncertainties. Results highlight the decomposition of regional RSL changes over 3,000 years along North America's Atlantic coast. The physical process glacial isostatic adjustment prevailed before 1800 CE, with anthropogenic forcing dominating after 1900 CE.

Keywords: Bayesian, sea-level, generalized additive models, uncertainty

1 Introduction

The Intergovernmental Panel for Climate Change (IPCC) in 2021 reported with 'high confidence' that global mean rates of sea-level rise increased from approximately 1.3 mm/year between 1901 and 1971 to 3.7 mm/year between 2006 and 2018, with a further increase in rates predicted for the remainder of the 21st century (Masson-Delmotte et al., 2021, p. 5). In contextualizing the socio-economic risk that this sea-level rise poses for coastal communities, it is necessary to place historic and predicted changes in a longer term (pre-anthropogenic) context and to recognize that local sea level can diverge sharply from the global average.

Relative sea level (RSL) is the height of the ocean surface at any given location and time, measured relative to the adjacent land (Church & Clark, 2013). Direct measurements of RSL (typically considered to be high accuracy and with low uncertainty) are made by a network of coastal tide gauges whose spatial distribution is highly uneven and whose temporal duration is typically limited to the past ~100 years or less (Church & White, 2011). Understanding RSL before tide-gauge measurements began requires proxies (physical, biological, or chemical features with an 'observable and systematic relationship to tidal elevation'; Horton et al., 2018) that are preserved in dated

Received: January 23, 2023. **Revised:** May 14, 2024. **Accepted:** July 30, 2024
© The Royal Statistical Society 2024.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

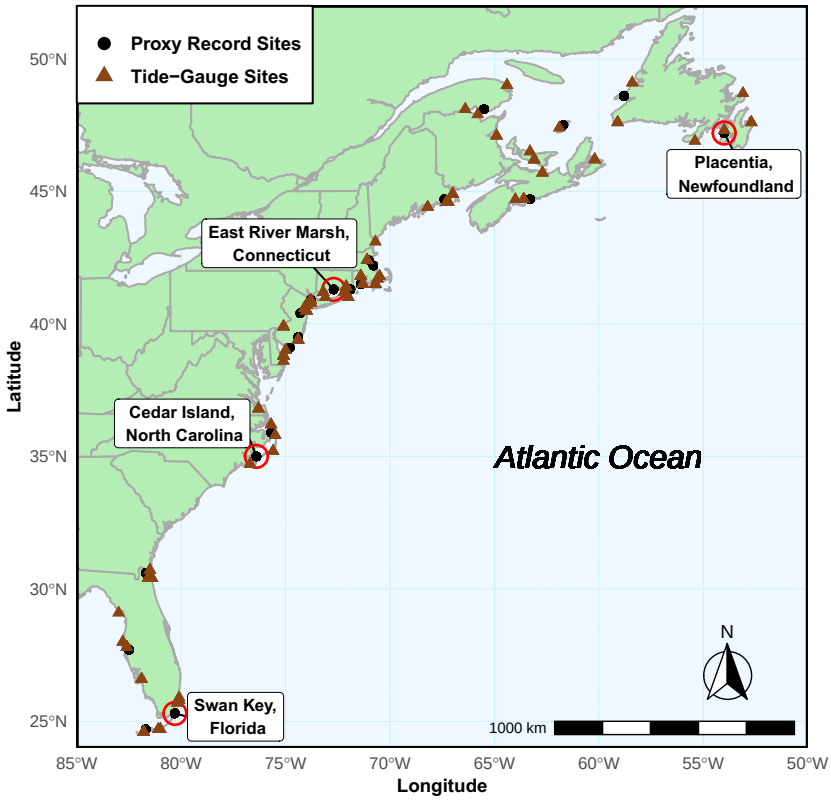

Figure 1. Location of the 66 tide gauge sites and 21 proxy record data sites along the Atlantic coast of North America with four proxy record sites chosen as case studies to present results of our model.

geological archives such as coastal sediment (e.g. Gehrels, 1994) or corals (e.g. Meltzner et al., 2017). For the past 3,000 years (a period in Earth's history called the late Holocene), it is possible to generate near-continuous proxy RSL reconstructions that overlap tide-gauge measurements (Kemp et al., 2013). The suite of late Holocene RSL proxy reconstructions is growing, but their global distribution is highly uneven (Ashe et al., 2019). However, the Atlantic coast of North America has a relatively large number of datasets (Figure 1) generated from sediment that accumulated in salt-marsh (e.g. Kemp et al., 2018) and mangrove environments (e.g. Khan et al., 2022). We therefore focus on this region to develop a new statistical model for quantifying patterns, rates, associated uncertainties and possible causes of late Holocene RSL change from a combination of proxy reconstructions and tide-gauge measurements concurrently. A general discussion on how proxy records are developed is provided in Section 2, and we point the reader in the direction of Shennan et al. (2015) for a more detailed account of the methodologies employed by the paleo sea-level community. In this article, we focus on analysing published data arising from proxy RSL reconstructions. The proxy records contain RSL estimates throughout time for different locations, specifically along the Atlantic coast of North America, and have associated bivariate uncertainties, i.e. uncertainty in time and vertical uncertainty in RSL (Figure 2).

Tide gauges and proxy records can only capture RSL, which is the net outcome of a complex combination of physical processes operating on characteristic temporal (years to millennia) and spatial (site-specific to global) scales. These physical processes often act simultaneously and serve to reinforce or mask one another; they can change both the height of the sea-surface and that of the land differently through time and across space (Khan et al., 2022). Consequently, RSL measurements can display a rich variety of spatio-temporal patterns. A principal goal of sea-level research is to interrogate these patterns to identify and quantify the contribution from specific physical processes including (but not limited to) the multi-millennial and regional response of the solid

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

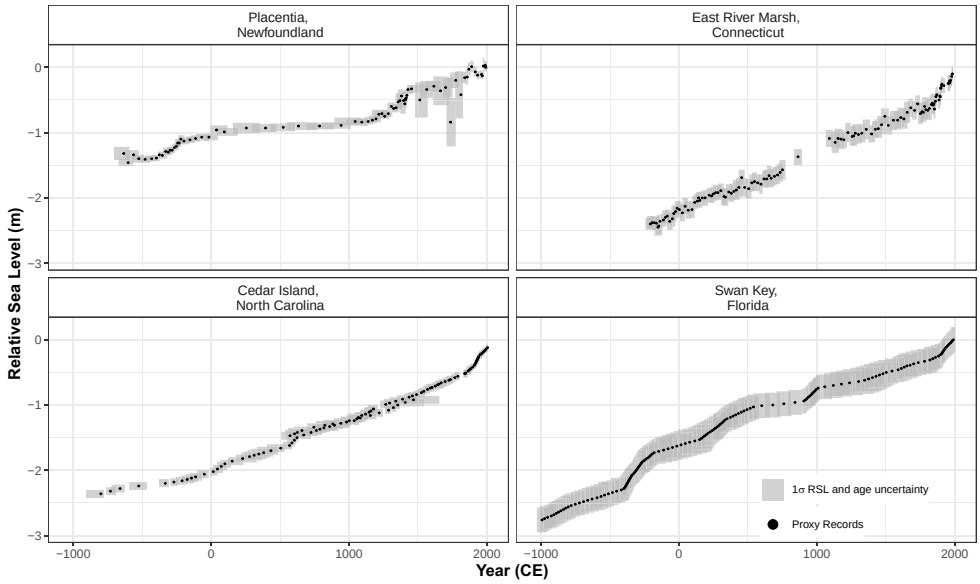

Figure 2. Proxy records from four proxy sites along the Atlantic coast of North America used as illustrative case studies. The *y*-axis is relative sea level (RSL) in metres, where 0 m is present sea level and negative values indicate RSL below present. Each proxy record observation consists of paired age and RSL estimate at the corresponding site. The black dot represents the midpoint of the proxy sea-level reconstruction and the grey boxes of 1 standard deviation represent vertical and horizontal (temporal) uncertainty.

Earth to de-glaciation, decadal to centennial redistribution of ocean water by changing currents, and the recent (multi-decadal) acceleration of global average sea-level rise in response to a warming climate (Church & Clark, 2013). This goal requires a means to decompose the site-specific RSL signal at each locality in the network into contributions at different temporal and spatial scales, while accounting for uncertainties in the underlying data.

The most widely used tool for decomposing late Holocene RSL is a model developed by Kopp *et al.* (2016, hereon K16) and its various extensions (Kemp et al., 2018; Walker et al., 2021). The K16 model decomposes RSL into three categories: (i) a non-linear signal common to all records in the dataset being analysed (termed global, irrespective of the geographic range of input data); (ii) a regional signal characterized by a linear rate of change over the past ~2,000 years, and (iii) a local (site-specific) signal that operates in a non-linear fashion. Rather than representing specific physical processes, these categories serve to represent groups of processes that operate at similar spatial and temporal scales informed by the data. K16 employs Gaussian Processes (GPs) for each component and, due to the associated computational burden that grows in proportion with the cube of the number of data points, relies on a maximum likelihood approach to estimate and fix model hyperparameters. These modelling decisions, that aim to reduce the computational burden of GPs, can impact uncertainty quantification (Ashe et al., 2019). In this article, we aim to propose an alternative method for estimating these complex, interdependent components, that improves uncertainty quantification whilst remaining computational feasible.

Our new spatio-temporal statistical approach to modelling RSL change uses Generalized Additive Models (GAMs). A GAM is a generalized linear model where 'the linear predictor depends linearly on a sum of smooth functions of the predictor variable' (GAMs; Wood, 2017, p. 161). GAMs flexibly model non-linear relationships using smooth functions (most commonly splines) and can reduce computational complexity when compared with GPs (for example those used in K16) as they do not require large matrix inversions. We place our model in a Bayesian framework that allows for the estimation of parameters conditioned on the RSL data with full accounting for, and propagation of, uncertainty. Similar to K16, our model partitions the total RSL signal into components that characterize distinctive spatial and temporal scales, which (to varying degrees) are associated with specific physical processes. These components are (i) a regional

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

component, a non-linear signal common to all sites along the Atlantic coast of North America, and equivalent to the global term in K16; (ii) a linear local component that contains unstructured random effects and is comparable the regional linear term in K16; and (iii) a non-linear local component, which is site-specific and varies smoothly in space and time. Similar to K16, any variation not addressed by the model is captured by a residual term.

Since the data points at each site have a bivariate error structure, and the decomposition required involves differing structures, the simple application of default GAMs does not work in our case study. Previous methods, such as Cahill et al. (2015), provides guidance for how to model RSL with bivariate uncertainty. We follow K16 in accounting for the time error using the Noisy-Input uncertainty method of McHutchon and Rasmussen (2011). This method inflates the residual variance by a corrective term to compensate for noisy-input measurements using a smooth process. Whilst the original paper uses the method exclusively for GPs, we extend the approach to spline terms. The RSL error is captured via a standard measurement error term added to the residual variance.

The structure of our article is as follows. Section 2 addresses the proxy records and tide-gauge data used in our analysis. Section 3 describes the main physical processes driving RSL changes and Section 4 discusses the previous modelling strategies employed by Kopp et al. (2016). Section 5 gives a detailed description of our statistical model with different splines representing each driver of RSL change and introduces the noisy-input method. The model validations are shown in Section 6, and the results for different drivers of RSL change and their associated rates are presented in Section 7. Section 8 provides concluding remarks for our approach for the Atlantic coast of North America.

2 Data

We use a combination of instrumental data from tide gauges and proxy records. This section discusses the different sea-level data sets, including their collection methods and their associated uncertainties.

2.1 Tide-gauge data

Tide gauges are fixed to the land and regularly measure (for example hourly or to higher frequency) the height of the adjacent sea surface (Pugh & Woodworth, 2014a). For understanding RSL change, these observations are usually expressed as annual averages and held in the database maintained by the Permanent Service for Mean Sea Level (PSMSL; Holgate et al., 2013; Woodworth & Player, 2003). The ~1,500 stations in this global network display highly uneven distribution of data across space and through time and in addition individual records may have temporal gaps (Church & White, 2011). The earliest tide gauges records began in the late 17th or early 18th centuries in northwestern Europe (Wöppelmann et al., 2006). Along the Atlantic coast of North America, the longest tide-gauge record in the PSMSL database is The Battery from New York City (since 1856 CE) (Holgate et al., 2013). Annual tide-gauge data from the PSMSL are treated as having fixed and known ages without uncertainty in elevation measurements (Holgate et al., 2013).

In our analysis, we use 66 tide-gauge sites along the Atlantic coast of North America (Figure 1). Tide gauges meeting at least one of the following criteria were included in our analysis; (i) record length exceeding 150 years; (ii) the nearest tide gauge to proxy site; (iii) within 1 degrees distance to a proxy site and longer than 20 years (Kopp et al., 2016; Walker et al., 2021). The addition of tide-gauge data supplements the long-term proxy records and provides additional insight into recent changes in RSL. Annual data for each tide gauge were downloaded from the PSMSL and expressed in metres relative to the average over 2000–2018 CE. This time window captures variability resulting from the 19-year cycle in astronomical tides (Pugh & Woodworth, 2014b) and serves to make proxy and tide-gauge data comparable since the sediment cores used to develop proxy reconstructions were recovered since ~2000 CE. In addition, we further average tide-gauge data by decade to increase comparability with proxy reconstructions that are developed from 1 cm thick slices of core sediment that accumulated over a period of several years (depending on sedimentation rate) and are therefore inherently time averaged. At this step, we include an uncertainty in the tide-gauge data (± 5 years for age and $\pm 1\sigma$ for RSL). See Appendix A for additional information.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

2.2 Proxy records

Proxy-based reconstructions provide estimates of pre-anthropogenic RSL (Kemp et al., 2013). On the Atlantic coast of North America, these near-continuous proxy-based reconstructions are generated using buried sequences of salt-marsh (at mid to high latitudes; Gehrels et al., 2020) or mangrove (low latitudes; Khan et al., 2022) sediment. Samples of this sediment are recovered in a core (a column of sediment extracted from the ground, where the oldest material is at the bottom and the youngest material is at the top) and interrogated in subsequent laboratory analysis to determine the age of the sample and the tidal elevation (height above a tide level) at which it accumulated (Horton & Edwards, 2006).

A history of sediment accumulation is established using an age-depth model (e.g. the Bchron (Parnell et al., 2008), Bacon (Blaauw & Christen, 2011), or Rplum (Aquino-López et al., 2018) packages in R). The input for an age-depth model are dates from a subset of depths in the sediment core, typically using radiocarbon measurements, an estimate of each core sample's age is obtained (Törnqvist et al., 2015). In addition, the shallowest (i.e. most recent) part of the core can be dated by recognizing historic pollution and land use changes of known age in down-core profiles of elemental abundance, isotopic activity and isotopic ratios (Marshall, 2015). The output of the age-depth model is and estimate the age (with uncertainty) of every 1 cm thick sediment sample in the core with uncertainty. Comparisons indicate that sediment accumulation histories have little dependence on the specific age-depth model used (Wright et al., 2017).

A sea-level proxy is required to reconstruct RSL. A sea-level proxy is any physical, biological or chemical feature with an observable and systematic relationship to tidal elevation (Shennan et al., 2015). Salt marshes and mangrove environments are vegetated by distinctive plant communities that are adapted to inundation by salt water, resulting in distinct and narrow elevation ranges (Redfield, 1972). This distribution makes salt-marsh vegetation a valuable sea-level proxy. Through reasoning by analogy, the observable distribution of plants in modern salt marshes enables interpretation of their analogous counterparts preserved in core material (Kemp & Telford, 2015). In this way, the paleo-marsh elevation (elevation with respect to tidal elevation at the time of formation) is reconstructed. Another sea-level proxy preserved in salt-marsh sediment is the remains of micro-fossils (e.g. foraminifera) that form distinctive assemblages with a strong relationship to elevation (Edwards & Wright, 2015). When using micro-fossils to reconstruct RSL a transfer function is required that relates the abundance of specific micro-fossil families to tidal elevation using a dataset that is representative to the modern environment (Kemp & Telford, 2015). There are various transfer functions available using Frequentist (Horton & Edwards, 2006; Kemp et al., 2011; Sachs et al., 1977) and Bayesian approaches (Cahill et al., 2016), which all estimate paleo-marsh elevation with uncertainty. The age of each core sample with a paleo-marsh elevation reconstruction is provided by the age-depth model. Resulting in a single proxy RSL record comprised of stratigraphically ordered data points of age (with 1 sigma uncertainty) and RSL (with 1 sigma uncertainty).

We analyse 21 RSL proxy records (totalling 1,731 data points) located on the Atlantic coast of North America from the Florida Keys, USA to Newfoundland, Canada (Figure 1). There are 66 tide-gauge records that meet our criteria for inclusion (Figure 1). The spatial scope of our analysis is restricted to this coastline because it has (by a considerable margin) the greatest concentration of available records. Results presented in this article are generated from all the proxy and tide-gauge records, yet we present four of these sites (Placentia Newfoundland Canada, East River Marsh Connecticut USA, Cedar Island North Carolina USA, and Swan Key Florida) as illustrative case studies throughout the remainder of the manuscript. The four sites were selected to provide diversity of location and therefore the processes causing RSL change during the past ~3,000 years.

3 Drivers of RSL change

Spatio-temporal models recognize sea-level variability characteristic of different spatial and temporal scales rather than from specific processes. Each component estimated in the model may capture several contributing processes depending on the location and time interval under examination. These processes may act simultaneously and in directions that mask or exaggerate contributions from other drivers.

Transfer of mass between land-based ice and the ocean drives RSL change. Ice melt/growth returns or removes mass to the ocean as liquid water, which causes a rise/fall in global mean sea level (this process is termed barystatic) (Gregory et al., 2019). This contribution varies in magnitude

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

across timescales, but is common to all locations. In addition, changes in global temperatures alter the density of ocean water resulting in a sea-level change (rise/fall when water warms/cools becoming less/more dense); this process is known as a thermosteric contribution (Grinsted, 2015). The global term in the K16 model and the regional component in our model attempts to capture influences from these processes.

Along the Atlantic coast of North America, the principal driver of RSL change during the pre-industrial late Holocene is glacial isostatic adjustment (GIA) (Engelhart et al., 2011; Walker et al., 2021). GIA is the response of the Earth, the gravitational field, and the ocean to the growth or decay of ice sheets (Whitehouse, 2018). GIA can be reasonably approximated as a linear contribution through time on this relatively short timescale, but with considerable variability along the coast (Engelhart et al., 2009). There are a family of physical models known as Earth-ice models that use a representation of the physical Earth structure (such as lithospheric thickness and properties such as mantle viscosity) to predict changes in GIA that occur through loading and unloading of ice, and provide estimates of GIA rates. One such example of an Earth-ice physical model is the ICE5G VM2-90 (Peltier, 2004). It is important to recognize that other processes (e.g. tectonically driven vertical land motion) can mimic the linear trend of GIA. However, along the passive margin of the Atlantic coast of North America these non-GIA drivers are likely modest in magnitude (Kopp et al., 2015). As a result, the linear local component in our model and the linear regional term in the K16 may capture contributions from processes other than GIA that drive RSL changes during the Common Era at a linear rate (positive or negative).

There are processes with a spatially coherent structure where the signal is shared by some but not all sites (Stammer et al., 2013). One such process that can cause RSL to vary on decadal to multi-century timescales is the redistribution of existing ocean mass by shifts in prevailing patterns and strength of atmospheric and oceanic circulation (termed dynamic sea-level change) (Gregory et al., 2019). Dynamic sea level varies by site, but the magnitude of the difference from one site to the next is too small to be detected using proxy data due to the resolution. Some processes (e.g. sediment compaction: that can impact the height of the solid Earth surface with changes in sediment volumes for each site; Horton et al., 2018) can drive RSL changes that are site-specific. Consequently, contributions from these processes lack spatial coherence and display an unpredictable spatial structure. Therefore, site-specific RSL changes can vary markedly across closely spaced sites. More often than not, RSL proxy reconstructions are not generated with the goal of understanding site-specific processes (Walker et al., 2021). It remains important to quantify this component as a means to distil the contribution from processes acting at larger spatial scales. In our model, the structured (common to some, but not all sites) and unstructured (unique to one site) RSL variability on century timescales is captured by the non-linear, local component.

4 Previous statistical models for RSL change

In this section, we review previous work on modelling RSL change, focusing in particular on K16. The model was further extended in Kemp et al. (2018) and Walker et al. (2021), here, we focus on the simpler K16 model. We first review the structure of this model, which decomposes RSL into component parts before discussing how the model might be fitted to the data and the potential influence of optimizing hyperparameters using maximum likelihood. K16 forms the basis upon which we build our new approach in Section 5.

The RSL measurements are recorded in units of height; with metres used by default. In cases where the scale of the change is relatively small, we use cm or mm instead for some plots and discussion in the text. We write $y_{ij} = y(\mathbf{x}_j, t_{ij})$ for the RSL height at location $\mathbf{x}_j$ (latitude and longitude) and time t_{ij}. These observations arise from the proxy records and tide gauges with j indexing the data site and i the observation. For the resolution of the data, the time is expressed in years CE. The K16 model can be written as:

$$y_{ij} = f(\mathbf{x}_j, t_{ij}) + w(\mathbf{x}_j, t_{ij}) + y_0(\mathbf{x}_j) + \epsilon^y_{ij}, \tag{1}$$

where f is the full RSL spatio-temporal field, w is a white noise process representing sub-decadal trends unexplained by the data due to resolution of data, y_0 is a site-specific spatially variable vertical offset, and ϵ^y_{ij} is residual error. In K16 all the structured terms above are given Gaussian

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Process prior distributions with stationary covariance functions.

A key complication is that the times t_{ij} associated with the proxy records are observed with uncertainty. Thus, the observed values $\tilde{t}_{ij}$ have measurement error, defined as:

$$\tilde{t}_{ij} = t_{ij} + \epsilon^{t}_{ij}. \tag{2}$$

Usually ϵ^{t}_{ij} is assumed iid normally distributed with known variance, though in reality the age-depth model through which the ages are estimated often provides skewed distributions. A previous attempt at resolving this issue can be found in Cahill et al. (2015), though across large, multi-site datasets the imposition of this assumption is believed to have minor effects on the outcome of the model (as shown in Parnell & Gehrels, 2015).

4.1 Decomposing the RSL field *f*

For the RSL process defined above as f, K16 use a spatio-temporal empirical Bayesian hierarchical model to partition the influence of the components into global, regional, and local scales. The fields that make up f are, as above, given stationary GP priors that can vary in time and space as controlled by the covariance functions (Ashe et al., 2019). The standard decomposition of f is written:

$$f(\mathbf{x}_j, t_{ij}) = c(t_{ij}) + g(\mathbf{x}_j)(t_{ij} - t_0) + l(\mathbf{x}_j, t_{ij}), \tag{3}$$

where $c(t_{ij})$ is termed the global term, the temporal non-linear signal common across all sites, designed to capture changes such as barystatic sea level rise and thermosteric changes. $g(\mathbf{x}_j)$ is a spatially varying term that captures slower processes such as long-term land level change (GIA) and vertical land motion driven by plate tectonics. The g term is multiplied by time t differenced from a reference point t_0 to form a temporally linear field. Unlike the other components in K16, g is given a univariate normal prior distribution with the mean centred on the value obtained from an Earth-ice physical model (ICE5G VM2-90; Peltier, 2004) that estimates the GIA rate. $l(\mathbf{x_j}, t_{ij})$ is the local spatio-temporal field that describes factors such as dynamic sea level change, sediment compaction and tidal regimes. These terms are explained in more detail in K16.

Without strong prior information, it is difficult to separate out the magnitudes of the model components. To address this in K16, the hyperparameters are initially determined by maximizing the likelihood of the model conditioned on the observations as detailed in the supplementary material of Kopp et al. (2016), albeit with a specific constraint. This constraint ensured that the two timescale hyperparameters for Gaussian process priors for the global term ($c(t_{ij})$) and the local term ($l(\mathbf{x_j}, t_{ij})$) are greater than 100 years. By allowing the two timescale hyperparameters to differ, it enables the global term to have a different characteristic time scale from the geographically varying local term. The model is then re-fitted using hyperparameters fixed at the values estimated via the maximum likelihood approach in order to estimate the components of the fields in an empirically Bayesian framework.

In our approach, we aim to avoid the empirical Bayesian approach of fixing parameter values to ensure that we capture and propagate the uncertainty associated with the hyperparameters. Instead, we opt for informed priors on model components, in contrast to the point estimate approach employed in K16. However, a modular modelling approach is unavoidable due to the complexity of the decomposition and the innate confounding of many of the key terms which are discussed in Section 5.

5 A new approach based on GAMs

In this section, we outline a new approach to evaluating the different drivers of spatio-temporal RSL using proxy records and tide gauge data. With careful choices of the prior distributions of the hyperparameters, we aim to recover the components of RSL change through the standard tools of Bayesian inference. Subsequently, we estimate rates of RSL change at sites along the Atlantic coast of North America. We build our model inspired by the standard decomposition of the

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

RSL field f as described above. Our approach contains four main differences compared to the Kopp et al. (2016) (and subsequent) models:

1. We focus on the high quality sites along North America's Atlantic coast and aim to produce a regional RSL curve. Thus we avoid making statements about global sea-level change.
2. We use splines instead of GPs to avoid the computationally challenging inversion of the GP covariance matrices. The model, at its simplest, thus falls under the standard generalized additive modelling paradigm.
3. We fit the model in two stages to maximize the regional variability that would otherwise be confounded with the local structure. This allows us to perform a more complete posterior analysis of the model hyperparameters that might otherwise have been fixed in K16.
4. We remove the spatial structure on the linear effect g in K16 and replace it with a univariate random effect on the slope. For the proxy records, we use an informed prior for the slope that is estimated using a linear regression for each site with data prior to 1800 CE (i.e. the pre-industrial time period defined by Neukom et al., 2019). This change is helpful because we have found the estimated values of the GIA rate from the Earth—ice physical model (e.g. Argus et al., 2014; Caron et al., 2018; Peltier, 2004) do not match the observed data well for the proxy record time period (Shennan et al., 2018; Vacchi et al., 2018). For the tide gauge records, the prior mean of the slope is taken from a physical Earth-ice model (ICE5G VM2-90; Peltier, 2004) with uncertainty taken from Engelhart et al. (2009). We refer back to this modelling choice in Section 8.

Below, we outline the full posterior distribution of the model and describe each term and its structure. The temporal uncertainty in the data causes difficulties in fitting the model in one step, and we resort to McHutchon and Rasmussen (2011)'s noisy-input method to account for this uncertainty. We then discuss the prior distributions assumed for the hyperparameters, and the computational details of our model. In Section 7, we showcase the successful implementation of our model.

5.1 Model notation

We now provide a full outline of our notation for reference:

- y_{ij} is an RSL observation in metres with $i = 1, \ldots, n_j$ observations at site j with $j = 1, \ldots, m$ sites. We vectorize the full set of observations as $\mathbf{y}$ and the observations for each site as $\mathbf{y}_j$, described in equation (6).
- t_{ij} are the ages of each RSL observation, indexed and vectorized as above, described in equation (6). We represent age in years of the Common Era (CE).
- $\mathbf{x}_j$ is the 2-vector of a latitude and longitude pair for each site j, described in equation (6).
- $z_{\mathbf{x}}$ is an index vector for the data sites that converts each site into a label, described in equation (7). Thus $z_{\mathbf{x}_j} = j$.
- $f(\mathbf{x}_j, t_{ij}) = f_{ij}$ is the mean sea-level process at site j and time t_{ij}, described in equation (7). We write $f(\mathbf{x}, t)$ as the mean process for a generic location and time, and continue with this notation below for brevity.
- $r(t)$ is the regional component at time t, described in equation (8).
- $l(\mathbf{x}, t)$ is the non-linear local component at location x at time t, described in equation (11).
- $g(z_{\mathbf{x}})$ is the linear local component at location x, described in equation (9).
- $h(z_{\mathbf{x}})$ is a site-specific vertical offset component at location x, described in equation (10).
- $b_r(t)$ and $b_l(\mathbf{x}, t)$ are sets of known b-spline basis functions corresponding to the regional (see equation (8)) and local components (see equation (11)), respectively.
- $\boldsymbol{m}_g$ and $\boldsymbol{s}_g$ are the mean and standard deviation parameters, respectively for the linear local correction component (see equation (13)). These are site specific and so each is a vector of length m.
- $\boldsymbol{\beta}^r, \boldsymbol{\beta}^l$ are the spline regression coefficient vectors of the regional (see equation (12)) and local (see equation (15)) components, respectively. $\boldsymbol{\beta}^r$ is of length k_r and $\boldsymbol{\beta}^l$ is of length k_l where k_r and k_l are the number of knots associated with each term.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

- $\boldsymbol{\beta}^g$ (see equation (13)), $\boldsymbol{\beta}^b$ (see equation (14)) are parameter vectors, each of length m, containing the random effect coefficients for each site.
- σ_r (see equation (12)) and σ_l (see equation (15)) are the smoothness parameters associated with the regional and local spline terms, respectively.
- σ_b is the standard deviation of the site-specific offset (see equation (14)).
- $s_{y_{ij}}$ is the known standard deviation of the RSL data point ij (see Section 5.5).
- $s_{t_{ij}}$ is the known standard deviation of the age of data point ij (see Section 5.5).
- σ is a residual standard deviation parameter to capture any remaining variability in y (see Section 5.2).
- ϵ_t and ϵ_y are the error terms for time and RSL, respectively, which are assumed to be i.i.d and normal distributed (see Section 5.5).

5.2 Posterior distribution

The joint posterior distribution of our Bayesian hierarchical model is shown below:

$$\begin{aligned}
&\underbrace{p(\sigma^2, \boldsymbol{\beta}^r, \boldsymbol{\beta}^l, \boldsymbol{\beta}^g, \boldsymbol{\beta}^b, \sigma_r^2, \sigma_l^2, \sigma_b^2 \mid \mathbf{y}, \boldsymbol{b}_r, \boldsymbol{b}_l, m_g, s_g^2, \mathbf{s}_y^2, \mathbf{s}_t^2)}_{\text{posterior}} \propto \underbrace{p(\mathbf{y} \mid \boldsymbol{f}, \sigma^2, \mathbf{s}_y^2, \mathbf{s}_t^2)}_{\text{likelihood}} \\
&\times \underbrace{p(\boldsymbol{\beta}^r \mid \sigma_r^2)}_{\text{prior on regional parameters}} \times \underbrace{p(\sigma_r^2)}_{\text{prior on regional smoothness parameter}} \\
&\times \underbrace{p(\boldsymbol{\beta}^l \mid \sigma_l^2)}_{\text{prior on non-linear local parameters}} \\
&\times \underbrace{p(\sigma_l^2)}_{\text{prior on non-linear local smoothness parameter}} \times \underbrace{p(\boldsymbol{\beta}^g \mid \boldsymbol{m}_g, \boldsymbol{s}_g^2)}_{\text{prior on linear local parameters}} \\
&\times \underbrace{p(\boldsymbol{\beta}^b \mid \sigma_b^2)}_{\text{prior on site-specific vertical offset parameters}} \times \underbrace{p(\sigma_b^2)}_{\text{prior on variance site-specific vertical offset parameters}} \\
&\times \underbrace{p(\sigma^2)}_{\text{prior on error variance}}
\end{aligned} \tag{4}$$

The likelihood $p(\mathbf{y} \mid \boldsymbol{f}, \sigma^2, \mathbf{s}_y^2, \mathbf{s}_t^2)$ can be deconstructed thus:

$$p(\mathbf{y} \mid \boldsymbol{f}, \sigma^2, \mathbf{s}_y^2, \mathbf{s}_t^2) = \prod_{j=1}^{m} \prod_{i=1}^{n_j} \mathcal{N}(y_{ij} \mid f_{ij}, \sigma^2 + s_{y_{ij}}^2 + s_{t_{ij}}^2). \tag{5}$$

5.3 A fully specified GAM for decomposing the RSL field

The data level for the NI-GAM is described as:

$$y_{ij} = f(\mathbf{x}_j, t_{ij}) + \epsilon_y, \tag{6}$$

where y_{ij} is the RSL, $f(\mathbf{x}_j, t_{ij})$ the mean sea level process and $\epsilon_y \sim N(0, \sigma^2 + s_{y_{ij}}^2)$ is the error vector with additional information provided in Section 5.1. Our version of the decomposition of the mean sea level field can be written as:

$$f(\mathbf{x}, t) = r(t) + g(z_{\mathbf{x}}) + h(z_{\mathbf{x}}) + l(\mathbf{x}, t). \tag{7}$$

All terms are as defined above: $r(t)$ is the regional component. $g(z_{\mathbf{x}})$ is the linear local component represented by a random effect with $z_{\mathbf{x}}$ representing each data site. $h(z_{\mathbf{x}})$ is the spatial vertical offset

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

for each data site. $l(\mathbf{x}, t)$ is the non-linear local component. We represent $r(t)$ using a spline:

$$r(t) = \sum_{s=1}^{k_r} b_{r_s}(t)\beta_s^r, \tag{8}$$

where β_s^r is the sth spline coefficient, k_r is the number of knots, and $b_{r_s}(t)$ is the sth spline basis function at time t.

The linear local component, $g(z_{\mathbf{x}})$, is an unstructured random effect for each site that is formulated as:

$$g(z_{\mathbf{x}_j}) = \beta_j^g t, \tag{9}$$

where β_j^g is a slope parameter specific for each site j. This specification is in contrast to K16 where the linear effect, g, varies smoothly in space and is informed through the prior by GIA model-derived values. We found such a restriction to adversely affect model performance due the lack of agreement between the data and the provided GIA values, and the wide variation in values between proximal sites (Engelhart et al., 2009).

The site-specific vertical offset h is a random effect used to capture vertical shifts associated with measurement variability between sites, ensuring the alignment of the 0 m RSL across all proxy records, facilitating comparability among them. It is formulated as:

$$h(z_{\mathbf{x}_j}) = \beta_j^h, \tag{10}$$

where β_j^h contains the random effect coefficients for site j.

The non-linear local component $l(\mathbf{x}, t)$ is described with a spatio-temporal spline function formulated by:

$$l(\mathbf{x}, t) = \sum_{s=1}^{k_l} b_{l_s}(\mathbf{x}, t)\beta_s^l, \tag{11}$$

where β_s^l is the sth spline coefficient, k_l is the number of knots and $b_{l_s}(\mathbf{x}, t)$ is the sth spline basis function at time t and location $\mathbf{x}$.

We use B-splines (de Boor, 1978) for both the regional and non-linear local terms. Our B-splines are constructed as piece-wise polynomials that join together at equidistant knots such that the first derivatives are equal (Eilers & Marx, 1996). For the regional component and the non-linear component, we use cubic B-splines as we are interested in the behaviour of the first derivatives. We can simply calculate these by differentiating the cubic B-splines and multiplying them with the posterior spline parameters to provide a posterior distribution for the derivative. For the non-linear local component, a tensor product is required to capture the variability over time and space (represented with longitude and latitude), which means that the individual covariates are combined product-wise (Wood, 2006). Many other basis function types and options are available (see, e.g. Dierckx, 1995; Wood, 2017), but we believe our approach balances both parsimony and computational efficiency for our application area. In Appendix B, we undertake a GAM sensitivity study to examine different basis function options and the optimal number of knots.

5.4 Prior distributions

Within the process level each component is given a prior distribution. Our prior for the spline coefficients of the regional component, β_s^r is:

$$\beta_s^r \sim \mathbb{N}(0, \sigma_r^2), \tag{12}$$

where σ_r is the standard deviation of the spline coefficient and fundamentally controls the

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

smoothness of the model fit.

Our prior for the linear local component for the proxy records is:

$$\beta_j^g \sim \mathbb{N}(m_{g_j}, s_{g_j}^2), \tag{13}$$

where m_{g_j} and $s_{g_j}^2$ are the empirically estimated rate and associated variance, respectively, obtained using a linear regression for each site j using data predating 1800 CE (this timeframe is defined as the pre-industrial period by Neukom et al., 2019). For the tide-gauge records, we obtain m_{g_j} from a physical model (ICE5G VM2-90; Peltier, 2004) and $s_{g_j}^2$ from previous studies (Engelhart et al., 2009).

Our prior distribution for the site-specific vertical offset is:

$$\beta_j^b \sim \mathbb{N}(0, \sigma_b^2), \tag{14}$$

where σ_b^2 is the variance of the random intercept across data sites.

Our prior on the spline coefficient for the non-linear local component is given as:

$$\beta_s^l \sim \mathbb{N}(0, \sigma_l^2), \tag{15}$$

where σ_l^2 is the variance of the spline coefficients over space and time. This parameter fundamentally controls the smoothness of the local non-linear effect.

The remaining hyperparameters of the model include σ_r^2, σ_b^2, and σ_l^2. The standard deviation parameter σ_b represents the variability in the site specific vertical shift after taking account of the local linear trend. As it is measured in metres, it is more interpretable in a physical context and so we place an informative prior here. The vertical shifts can be quite variable with some sites sitting many metres above or below others. From revisiting the publications associated with our data (see Appendix A for the full list), shifted values spanning more than 5 m seem unlikely. As a result, we specify the standard deviation to have a Cauchy distribution with mode 2.5 m but with a wide scale of a further 2 m (Gelman, 2006). For the variability of the spline coefficients across the knots, i.e. σ_r^2 and σ_l^2, we expect considerably smaller variation but we have less information, thus we use a truncated Cauchy distribution centred on zero and with scale value 1.

5.5 Noisy-input uncertainty method

Our data are corrupted with measurement error in the RSL values and that arising from the temporal uncertainty associated with radiocarbon dating the fossil layers of sediment. McHutchon and Rasmussen (2011) describe an assumption for GPs that avoids the need for complex errors-in-variables models (e.g. Cahill et al., 2015; Dey et al., 2000) and instead adds an extra measurement uncertainty on the response variable. We adapt this noisy-input (NI) approach for our RSL GAM that we now term an NI-GAM.

The response variable y is assumed to be a noisy measurement with the true output given as $\tilde{y}$:

$$y = \tilde{y} + \epsilon_y, \tag{16}$$

where the error term is given by $\epsilon_y \sim \mathbb{N}(0, s_y^2 + \sigma^2)$ with s_y being the known measurement standard deviation of the RSL data, and σ^2 is the residual standard deviation of the model. Similarly, for the input measurements, t is assumed to be a noisy estimate of the true time value $\tilde{t}$:

$$t = \tilde{t} + \epsilon_t \tag{17}$$

with the error term given by $\epsilon_t \sim \mathbb{N}(0, s_t^2)$ where s_t is the known standard deviation of the date

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

obtained from the age-depth model (described in Section 2). As a result, a function for the response variable is formed in the following way:

$$y = f(\mathbf{x}, \tilde{t} + \epsilon_t) + \epsilon_y. \tag{18}$$

Following McHutchon and Rasmussen (2011), we use a first-order Taylor expansion about the latent state $\tilde{t}$ to obtain:

$$f(\mathbf{x}, \tilde{t} + \epsilon_t) \approx f(\mathbf{x}, \tilde{t}) + \epsilon_t^T \frac{\partial f(\mathbf{x}, \tilde{t})}{\partial \tilde{t}}. \tag{19}$$

Thus, the error in t can be approximated by an increase in the measurement error proportional to the derivative of f. McHutchon and Rasmussen (2011) calculate the derivative of the mean of the GP function, given as vector $\partial_{\tilde{f}}$ for the first order case and $\Delta_{\tilde{f}}$ for a D-dimensional matrix.

Analogously for our NI-GAM setting, the first-order terms are expanded to form a linear model with input noise:

$$y = f(\mathbf{x}, t) + \epsilon_t^T \frac{\partial f(\mathbf{x}, t)}{\partial t} + \epsilon_y. \tag{20}$$

The derivative of the posterior mean for f is obtained using a two-step method. First, the model is fitted ignoring the input uncertainty and then the slope of the posterior mean is calculated. From this, a corrective variance term can be calculated, $\tilde{s}_{t_{ij}}$, which is used as an additional model error term in our subsequent full model fit, described in Section 5.6.

Intuitively, the presence of noise in the inputs (the time uncertainty) of the function is directly related to the variance of the outputs (RSL). The variance of input noise is intricately linked to the output (RSL) variability through the square of the posterior mean function's gradient (McHutchon & Rasmussen, 2011). Consequently, corrupted input measurements have a stronger impact on rapidly changing output values than on those that remain constant, i.e. in areas with steep sea level gradients, greater time uncertainty results in increased output variability. The advantage of this method is that the noise remains the same whether the measurement is considered an input or output measurement, and so all the data informs the input noise variance ensuring that the output dimensions are met, reducing the chance of over-fitting.

5.6 Model fitting and computational details

In previous sections, we have described our Bayesian hierarchical model structure using GAMs and the manner in which we account for uncertainty. In this section, we address how to fit our NI-GAM model and the decisions that influenced our model fitting strategy. We are constrained because we have to fit the model twice as part of the noisy-input uncertainty method, described in Section 5.5. We also found that a single model fit yielded poor convergence due to the unavoidable confounding between the regional, vertical shift, linear, and non-linear local terms. Thus, we use the two-stage NI process to our advantage by fitting a slightly reduced model at the first stage to estimate the regional term, and using the posterior as strong prior information in the second stage to provide the estimate of the non-linear local term. Our approach has some similarities to that of cut feedback or modularized Bayesian models (Plummer, 2015), but we do not explore these avenues further here.

The two steps of our model fit are:

1. We first fit a simplified version of our process level model where we replace $f(\mathbf{x}, t)$ with $f^*(\mathbf{x}, t)$, defined as:

$$f^*(\mathbf{x}, t) = r(t) + g(z_{\mathbf{x}}) + h(z_{\mathbf{x}}). \tag{21}$$

This removes the non-linear local component and so avoids the confounding issue. From this

Downloaded from https://academic.oup.com/jrssc/advance-article/doi/10.1093/jrssc/qlae044/7756353 by Maynooth University user on 23 September 2024

model fit, we calculate the first derivative of the posterior mean of $f^*(\mathbf{x}, t)$. The resulting slope estimate for each observation provides a corrective variance term, $\tilde{s}^2_{t_{ij}} = s^2_{t_{ij}} \left(\frac{\partial f^*(\mathbf{x}_j, t_{ij})}{\partial t_{ij}} \right)^2$. This term is added to the other model error variances for the fit in the second stage.

2. In the second step, we fit the complete process model as defined in Section 5.5. The only changes being (i) the addition of the new noisy-input measurement variance term; (ii) the prior distribution on the regional spline terms now being $\beta^r_s \sim \mathbb{N}(m^r_s, (s^r_s)^2)$ where m^r_s and s^r_s are estimated in the first model run; and (iii) the prior distribution on the vertical offset term being $\beta^h_s \sim \mathbb{N}(m^h_s, (s^h_s)^2)$, where as above m^h_s and s^h_s are estimated in the first model run.

The second stage of model fitting integrates comprehensive error uncertainty, decomposing residuals into both a pure error component and a non-linear local space-time effect. Our subsequent results rely on this refined model fit, offering insights into model error and, critically, the extent to which local factors influence deviations from the main regional effect.

At each stage, our models are written using the Just Another Gibbs Sample (JAGS: Plummer, 2003) software, which in turn is based on techniques described in Spiegelhalter et al. (2002). The JAGS language uses Gibbs sampling and the Markov Chain Monte Carlo (MCMC) algorithm to draw samples from the posterior distribution of the unknown parameters. We implement our approach using the `rjags` package in R (Plummer et al., 2016). For our models, we used 5,000 iterations with a burn-in value of 1,000, thinning at 5 and 3 chains. Convergence diagnostics for the parameters are investigated using the coda package (Plummer et al., 2006) and the ShinyStan app that provides an interactive visualization tool for investigating model convergence (Gabry & Goodrich, 2017). All convergence diagnostics were checked and ensured to be satisfactory before the model results were interpreted. The code and data for our model are available on Github and on Zenodo, and the computational run time was 39 min on a 2019 MacBook Pro with a 1.4 GHz Quad-Core Intel Core i5 processor. Furthermore, NI-GAM is incorporated into the `reslr` package in R, offering users a range of statistical models for analysing RSL data within Bayesian hierarchical frameworks (Upton et al., 2023).

6 Model validation

We examine the validity of our model using 10-fold cross-validation (10-CV). We run the 10-CV on 20 proxy sites and 7 tide gauge sites that have data exceeding 10 years of observations. We do not include the remaining sites because they are short in duration and provide relatively weak information about model performance. We present the site-specific results for only the four case study sites with the remaining sites shown in Appendix B. We evaluate the model performance based on the out-of-sample empirical coverage and the root mean squared error (RMSE). The prediction intervals are created using posterior predictive simulations with the full error structure, i.e. $\hat{y}_{ij} \sim N(\hat{f}_{ij}, \sigma^2_{y_{ij}} + \sigma^2_{t_{ij}} + \sigma^2)$. The empirical coverage provides the percentage of occasions that the true RSL observation is within the model prediction interval (PI) for RSL. The RMSE provides insight into prediction performance in the same units as the response (metres).

The 10-CV for our full model obtained overall empirical coverage of 98.8% with the 95% prediction interval and 79.9% with the 50% prediction interval. The RMSE for the model is 0.14 m. An average out-of-sample error of 14 cm is reasonable given the scale and variety of the data set. These are satisfactory for a model fitted to complex data such as ours, especially with the addition of the model error. The conservative empirical coverage values, which exceed the expected 50% and 95%, can be attributed to the consideration of the observed measurement errors in estimating the prediction intervals. By accounting for these errors, our model tends to generate wider prediction intervals, resulting in a greater proportion of true observations falling within the predicted uncertainty range.

Figure 3 presents the true RSL observations vs. the model-based RSL point estimates with 95% prediction intervals at each site and Table 1 provides a site-specific insight into the empirical coverage for the model and the size of the prediction intervals. Three out of four sites have a coverage of 100% due to the large prediction intervals arising from the bivariate uncertainties associated with the proxy data (Figure 2). Based on the RMSE, the best fitting case study site is Cedar Island in North Carolina, where the RMSE is 7 cm (Table 1). At the other end of the spectrum is Swan Key Florida, where the RMSE is larger at 19 cm (Table 1).

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

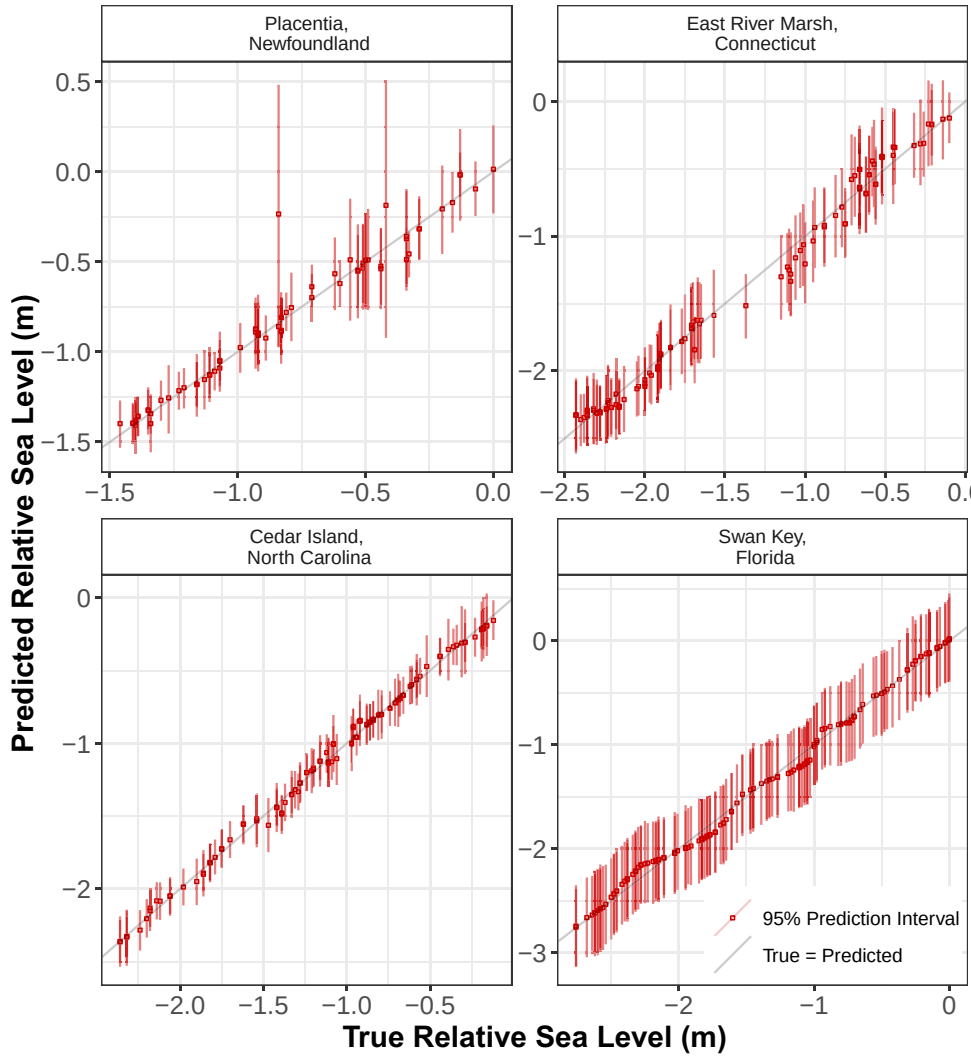

Figure 3. True vs. predicted RSL values using 10-fold cross-validation (CV) using the 20 proxy records and 7 tide gauge sites that have data exceeding 10 years of observations. Illustrated are the outcomes from our case study locations: Placentia Newfoundland, East River Marsh Connecticut, Cedar Island North Carolina and Swan Key Florida. The predicted means are the red points with the vertical red lines representing the 95% prediction interval for each point. The identity line is shown in grey.

Table 1. Empirical coverage from the 10-fold cross-validation and the corresponding size of the prediction intervals (PI) used for model validation for our 4 chosen sites

Site name	95% Empirical coverage	95% Average PI width (m)	50% Empirical coverage	50% Average PI width (m)	RMSE (m)
Placentia, Newfoundland	0.99	0.36	0.81	0.12	0.10
East River Marsh, Connecticut	1.00	0.53	0.74	0.18	0.14
Cedar Island, North Carolina	1.00	0.29	0.85	0.10	0.07
Swan Key, Florida	1.00	0.77	1.00	0.27	0.19

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

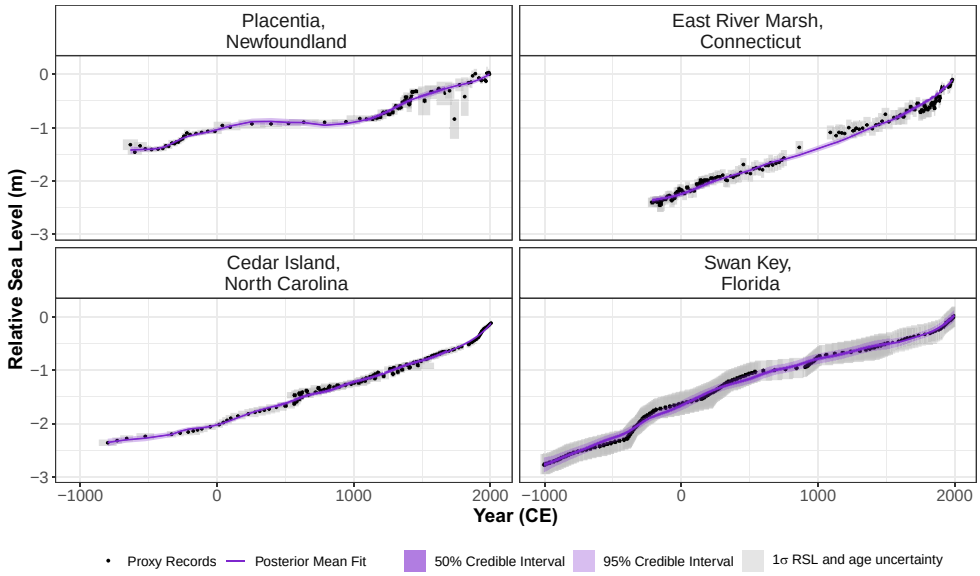

Figure 4. The noisy-input generalized additive model (NI-GAM) fit ($f(\mathbf{x}, t)$) for four selected sites along the Atlantic coast of North America. The four sites include: Placentia, Newfoundland, Canada; East River Marsh, Connecticut, USA; Cedar Island, North Carolina, USA; and Swan Key, Florida, USA. The black dots and grey boxes represent the midpoint and associated uncertainty, respectively, for each proxy record. The solid purple line represents the mean of the model fit with a 95% credible interval denoted by shading.

7 Results

In this section, we present the results from our Bayesian hierarchical RSL model. We consider the full model fit and its decomposition into the different components of RSL, i.e. regional component, linear local component and non-linear local component.

7.1 Full model fit and rate of change

The full model fit for the four case study sites are shown in Figure 4 (results from all 21 proxy sites are included in Appendix B). The model demonstrates how the RSL varies in time at each site. Overall, the model fits the data well. The 95% credible intervals for Swan Key Florida are larger due to the relatively large observation uncertainties at this site and the fit is notably smoother than the others. The data in Placentia, Newfoundland experience additional variability in the observations compared with the other sites and this is reflected in a more variable total model fit.

Figure 5 shows the site-specific rates of change for the case study locations calculated as described in Section 5. The remaining sites are shown in Appendix B. Late Holocene rates of RSL change display century to multi-century scale variability around a stable mean at each site until the 19th and 20th centuries since when the rate of rise appears unprecedented for 3 out of 4 case study sites. Rates fluctuate throughout the last 2,000 years but remain below 1.5 mm/year until the late 1800s in East River Marsh, Connecticut, and the early to mid-1900s in Cedar Island, North Carolina and Swan Key, Florida. The late 20th and 21st century rates at these sites are unprecedented in the last 2,000 years with the most recent rates of change being 3.06 ± 0.3, 2.9 ± 0.5, and 2.9 ± 0.7 mm per year in East River Marsh Connecticut, Cedar Island North Carolina, and Swan Key Florida, respectively. Placentia Newfoundland does not experience the same uptick in rates that the other sites do with the most recent rate being 1.21 ± 0.4 mm per year.

7.2 Examining the decomposition of RSL

The RSL process level f consists of the regional component, the linear local component, the site-specific vertical offset, and a non-linear local component, all as described in Section 5. Figure 6 illustrates the decomposition in our case study sites and provides an insight into how the

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

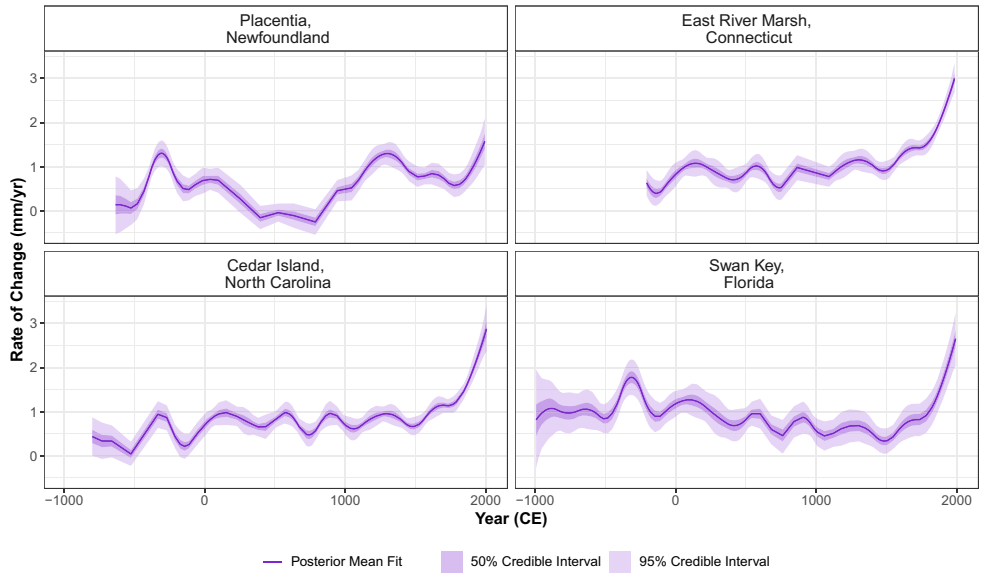

Figure 5. Rate of relative sea change found by taking the first derivative of the total model fit ($f'(\mathbf{x}, t)$) for four sites along the Atlantic coast of North America. The mean of the fit is the solid purple line with the dark shaded area being the 50% credible interval and the light shaded area being the 95% credible interval.

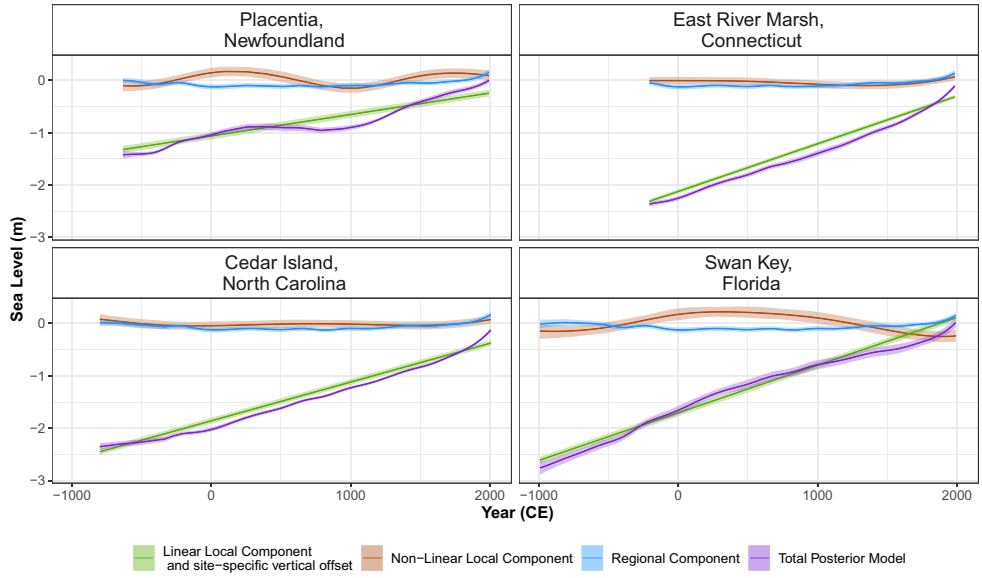

Figure 6. The decomposition of the relative sea level process level for the four sites, with solid lines indicating means and shaded areas 95% posterior credible intervals. The blue curve represents the regional component (r). The brown curve represents the non-linear local component (l). The green line represents the site-specific vertical offset plus the linear local component ($g + h$). The purple line is the sum of all three components and represents the full noisy-input generalized additive model fit ($f(\mathbf{x}, t)$).

components of RSL have varied over time for the Atlantic coast of North America by demonstrating the scale of the different components and how they interact over time. The total posterior model is obtained by the sum of these separate components as illustrated by the purple line in Figure 6.

Downloaded from https://academic.oup.com/jrssc/advance-article/doi/10.1093/jrssc/qlae044/7756353 by Maynooth University user on 23 September 2024

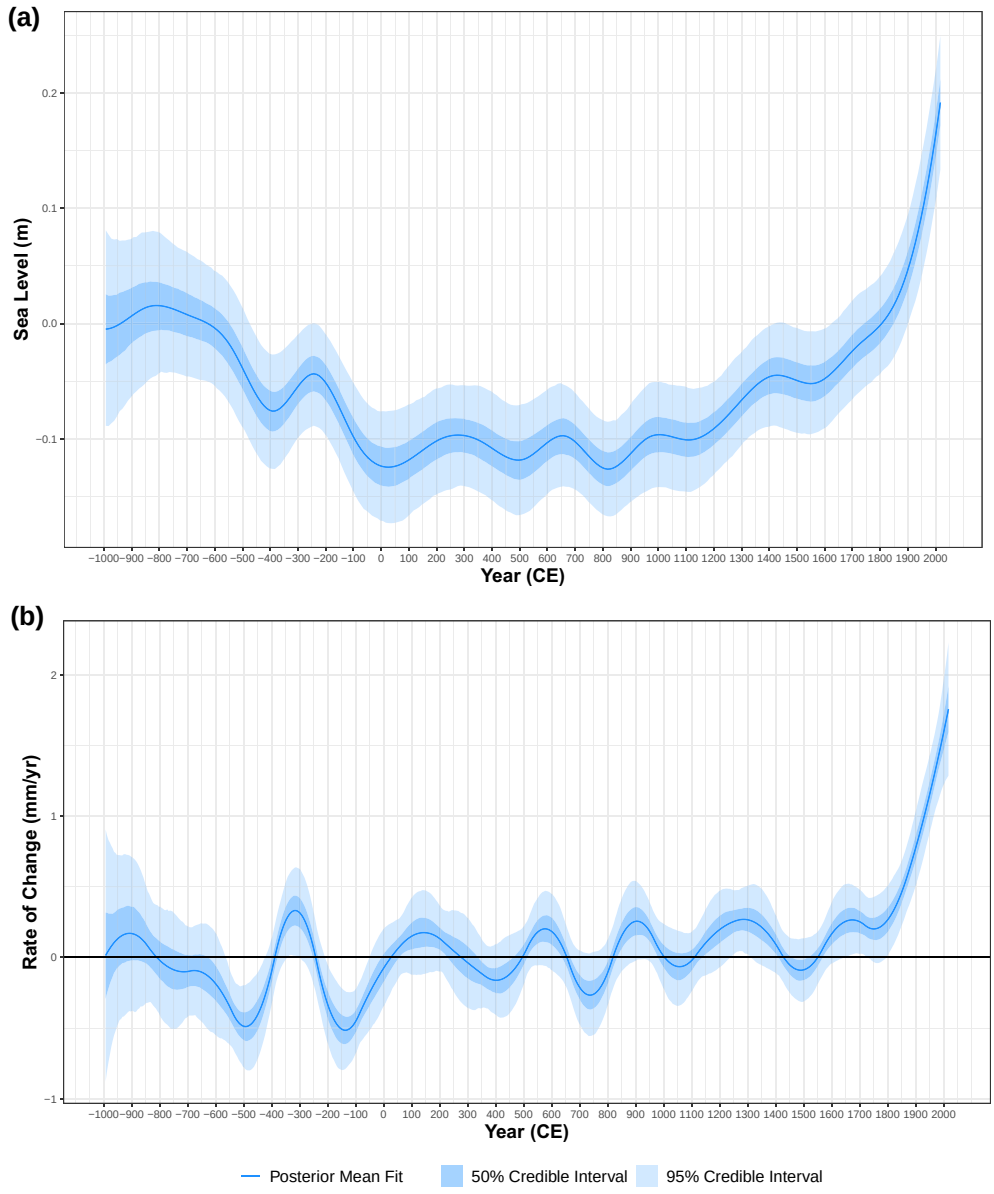

Figure 7. The noisy-input generalized additive model (NI-GAM) results for (a) the regional component ($r(t)$) and (b) the regional rate of change component ($r'(t)$). (a) The regional component mean model fit represented with a solid line and the shading indicating the 50% credible interval in dark blue and 95% credible interval in light blue. The *y*-axis is the sea level in m with the *x*-axis representing the time across the last 3,000 years for the Atlantic coast of North America. (b) Rate of change for the regional component for the Atlantic coast of North America with the solid line representing the mean of the fit, the dark blue shaded area representing the 50% credible interval and the light blue shaded area representing the 95% credible interval. The *y*-axis is the instantaneous rate of change of regional sea level in mm per year.

It is evident that the dominant driver of RSL change for these four sites until the late 1800s is the linear local component. After this interval, regional variability along the Atlantic coast of North America appears to take over and the total RSL trends at each site tend to reflect the RSL rise seen in the regional component. The non-linear local component is picking up the remaining variability and demonstrates that non-linear local effects on RSL variability are more apparent in Placentia and Swan Key compared to Cedar Island and East River Marsh.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Table 2. The rate associated with the linear local component (*g*) for our four sites along the Atlantic coast of North America is given in mm per year

	Empirical rate prior to 1800 CE (mm/year)	ICE5G-VM2-90 Earth-ice GIA rate (mm/year) (Peltier, 2004)
Placentia, Newfoundland	0.41	0.21
East River Marsh, Connecticut	0.91	0.96
Cedar Island, North Carolina	0.74	0.69
Swan Key, Florida	0.90	0.11

Note. These empirical rates were obtained using a linear regression through the data prior to 1800 CE (the time period marking a notable shift in global temperatures defined by: Neukom et al., 2019), which is used to inform the priors for the linear local component. ICE5G-VM2-90 Earth-ice GIA rate is from an Earth-ice physical model (Peltier, 2004).

Figure 7a shows the regional component (common to all sites) for the 21 proxy sites and 66 tide gauge sites along the Atlantic coast of North America. As a reminder, the regional component is represented with a spline in time that is common across all sites. Prior to 0 CE, sea level fluctuated from 0.01 m to −0.13 m. From 0 CE to 1200 CE, variability of sea level reduced ranging from −0.12 m to −0.09 m. Following 1200 CE, a sharp increase in sea level can be seen with brief periods of stability from 1410 CE to 1560 CE and from 1800 CE to 1840 CE. After 1800 CE, sea levels are consistently rising and the most dramatic increase can been seen from the mid-1800s until the present day. Figure B1 in Appendix B demonstrates the underlying behaviour of the posterior samples for the regional component. Figure 7b shows the rate of change for the regional component along the Atlantic coast of North America. Rates fluctuated around 0 mm per year between −990 CE and approximately 1800 CE, after that a continuous increase can be observed from around 1800 CE onwards. The rate from the late 20th century is unprecedented when compared with the last 3,000 years and is estimated to be 1.8 ± 0.5 mm per year. The regional sea-level rate observed in the 20th century reflects patterns across multi-decadal to centennial timescales, influenced by the resolution constraints and natural time-averaging inherent in proxy reconstructions, as well as the decadal averaging of tide-gauge data.

The linear local component is represented with a random slope effect as described in Section 5. As stated in Section 3, this parameter removes a long-term variation driven principally (but perhaps not exclusively) by GIA. Table 2 compares our empirically estimated values of this parameter for the proxy record sites in mm per year prior to 1800 CE. It is evident that there is significant variability in these values among different sites. For instance, areas like Swan Key, Florida, and East River Marsh, Connecticut, experience rates of 0.91 mm/year, whereas Placentia, Newfoundland, has a rate of just 0.41 mm/year. To show a comparison with physical model-based GIA rates, Table 2 presents values obtained from the ICE5G - VM2-90 Earth-ice model (Peltier, 2004). It is evident that the data-driven rates and the GIA-model rates differ, with Swan Key Florida experiencing the greatest difference of 0.8 mm/year. Whereas, East River Marsh, Connecticut has similar rates with a difference of only 0.05 mm/year.

Figure 8 shows our non-linear local component that represents the spatially structured behaviour specific to each site. There are clearly different patterns of non-linear local sea-level change, which is to be expected given that the common source of variation across all sites has been captured by the regional component. Placentia Newfoundland and Swan Key Florida show non-linear local variations in sea level ranging from 0.19 m to −0.12 m and 0.21 m to −0.25 m, respectively. On the other hand, Cedar Island North Carolina and East River Marsh Connecticut do not experience this level of variability with sea levels in the non-linear local component fluctuating close to zero.

8 Discussion

RSL change is the net result of multiple physical processes within the oceans, atmosphere, and solid Earth, that can alter the height of the land and/or sea surface (Church et al., 2001). The importance of specific processes varies markedly across space and through time, giving rise to a complex

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

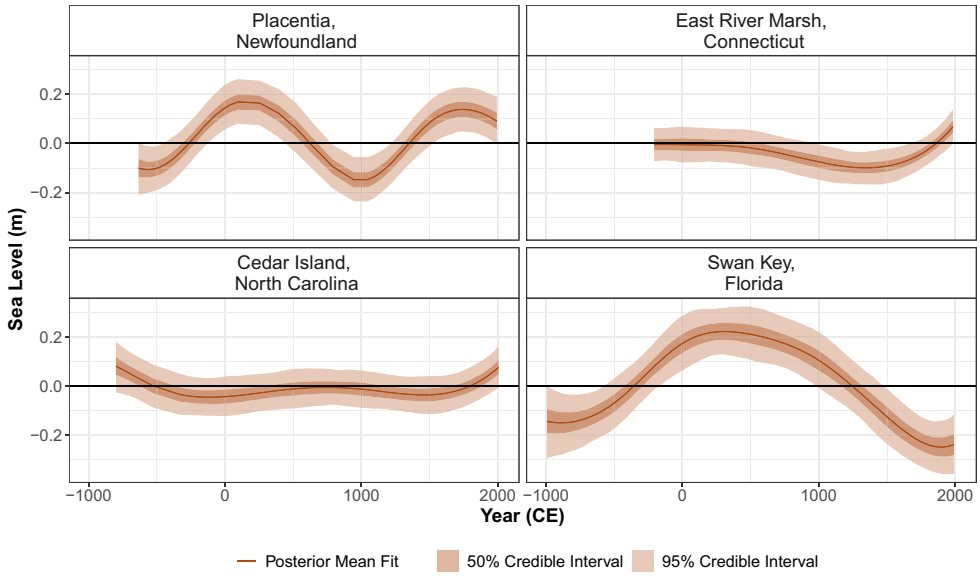

Figure 8. The non-linear local component ($l(\mathbf{x}, t)$) for our four sites along the Atlantic coast of North America. The y-axis represents sea level in metres. The brown solid line represents the mean of the model fit with the 50% credible interval in dark brown shading and 95% credible interval in the light brown shading.

and evolving pattern of RSL change. Tide-gauge data and proxy records contain information about these processes and the sea-level community requires advanced statistical tools to decompose the net RSL signal into contributions from physical processes while accounting for uncertainties in the underlying data. The need for RSL decomposition motivated current modelling strategies such as the K16 model and our approach presented here.

Our approach provides a more computationally efficient method for decomposing the RSL signal. The process level of our model utilizes a spatio-temporal field decomposed into a regional component, a linear local component, and a non-linear local component. In contrast to K16 that uses GPs, we use splines to examine the different drivers of RSL. This is due to the computational complexity associated with the likelihood computation for a Gaussian Process being of $O(n^3)$ where n is the number of data points. In contrast, the likelihood computation for the equivalent spline with pre-computed basis functions is just $O(n)$ (Wood, 2017). Thus, our model is fitted quicker than K16, allowing for further checks on the performance of our model, such as 10-fold cross-validation across extensive temporal records and numerous spatial sites. As the database expands, this discrepancy in run-time will magnify, resulting in our approach becoming increasingly more efficient. The model validations presented in Section 6 highlight that the NI-GAM is effectively and efficiently capturing the different components of RSL along the Atlantic coast of North America. In addition, the construction of our GAM using spline basis functions and random effects allows for easy interpretability without the need for covariance matrices and correlation functions (Porcu et al., 2021). We can accurately model late Holocene RSL changes along the Atlantic coast of North America and the interpretability of GAMs allows for these changes to be easily examined (Figure 4).

Our approach attempts to deviate from the Empirical Bayesian framework as implemented by K16 and related models. Piecuch et al. (2017) demonstrated that Empirical Bayesian methodologies can underestimate uncertainty when examining historic sea-level change along the Atlantic coast of North America. However, we recognize the difficulty of a fully Bayesian approach due to the confounding nature of the regional, linear local, and non-linear local components. Instead, we opted to take advantage of the two-step fitting required by our use of the noisy-input method to take account of age errors. The first step of the modelling procedure obtains posterior distributions for the regional component and the site-specific vertical offset. The second step uses the resulting posterior estimates and uncertainties to inform the priors for the remaining linear local and non-linear local components, and the extra measurement variance contribution from the age uncertainties. The first step can be

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

thought of as estimating the main component of our model: the regional RSL curve, with the second step designed to decompose the residuals and ensure the uncertainty is properly calibrated. Our modelling strategy avoids fixing process model parameters, and severe confounding that would occur were we to fit the model in one step.

Considering the individual RSL components, pre-anthropogenic (before 1800 CE; Neukom et al., 2019) RSL change along the Atlantic coast of North America is dominated by the linear local component that is principally capturing the contribution from ongoing GIA. However, there are some notable differences between the empirically estimated rates obtained from our models and the GIA rates obtained from the ICE5G-VM2-90 Earth-ice physical model (Table 2). There are several possible explanations for these discrepancies. First, a single Earth-ice model generates GIA predictions from a specific representation of the solid Earth (e.g. mantle viscosity and lithospheric thickness parameters) and history of deglaciation. It is unlikely that any single Earth-ice model will perfectly estimate GIA at all places and all times because the parameters are uncertain and may vary by location (Roy & Peltier, 2015). In particular, locations close to the margins of former ice sheets (such as Newfoundland) may exhibit particularly pronounced differences in GIA estimated by different Earth-ice models. Systematic difference between RSL predicted by specific Earth-ice models and proxy reconstructions on Holocene timescales is well documented in eastern North America (Vacchi et al., 2018) and elsewhere (Shennan et al., 2018). One such example is Placentia Newfoundland where our empirically estimated rate and the Earth-ice physical model GIA rate differ by 0.1 mm/year (Table 2). Second, physical Earth-ice models only estimate the contribution from GIA, while the empirical approach captures contributions from other processes such as vertical land motion from tectonic processes that may also be a linear driver of RSL change on the timescales under consideration. Although these non-GIA processes may be small on the passive margin of the Atlantic coast of North America, they are also unlikely to be zero at all sites. For example, Khan et al. (2022) reconstructed RSL at two sites (Snipe Key and Swan Key) in southern Florida where the predicted (by ICE5G-VM2-90 Peltier et al., 2015) rate of long-term change is 0.1 mm/year and the two sites were anticipated to share a common sea-level history given their proximity to one another and no indication of an unusual contribution from local-scale processes. The reconstructed rates of RSL rise were ~0.75 mm/year at Snipe Key and ~1.1 mm/year at Swan Key. This suggests that long-term RSL trends in Florida exceeded those predicted by an Earth-ice model. Similar rates of rise were reconstructed by previous studies (e.g. Hawkes et al., 2016; Scholl & Stuiver, 1967). This regional difference may be attributed to a contribution from non-GIA processes that are not included in Earth-ice models (e.g. tectonics), or limitations of the model (e.g. not including 3D Earth structure). The faster rate of rise at Swan Key compared to Snipe Key was interpreted by Khan et al. (2022) as evidence for a contribution to RSL rise from a local-scale process, possibly dissolution of the underlying carbonate bedrock.

After ~1900 CE, the regional component dominates, and we see the regional rate of change increase markedly from 0.7 ± 0.5 mm/year in 1902 to 1.8 ± 0.5 mm/year at the end of the 20th century (Figure 7b). This change is the result of anthropogenic forcing of the climate system (Neukom et al., 2019), which drove sea-level rise through thermosteric and barystatic processes (Frederikse et al., 2020). Our estimate of regional sea-level rise during the 20th century represents trends sustained on multi-decadal to centennial timescales because of the natural time-averaging and resolution limits of the proxy reconstructions and our decadal average of tide gauge measurements. Despite our analysis using different sources of data, being limited in geographical scope to the Atlantic coast of North America and using a new approach, our estimated century-scale rate of historic sea-level rise is similar to those proposed in previous studies. Using global compilation of tide gauge-records (and no proxy data), Hay et al. (2015) concluded that global mean sea level rose at 1.2 ± 0.2 mm/year during the 20th century, while Frederikse et al. (2020) estimated a rate of 1.56 ± 0.33 mm/year since 1900 CE. Using the an extended version of the K16 model (proxy reconstructions from around the world supplemented with nearby tide gauge data), Walker et al. (2022) estimated global mean sea-level rise to be 1.4 ± 0.2 mm/year since 1940CE.

The regional component in our model is the term common among all sites (which are limited to the Atlantic coast of North America). This term cannot be directly compared to the global term in the K16 model which is common to sites in a dataset drawn (unevenly) from all ocean basins. Processes such as dynamic sea level change from ocean currents and the spatial fingerprint of ice melt could cause sea level in the western North Atlantic to depart from the global mean. For example, mass loss from the West Antarctic ice sheet would cause many of the sites used in our

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

study to experience sea-level rise approximately 1.2 times greater than the global average (e.g. Hay et al., 2017). The similarity between our estimate rate of regional rise during the 20th century and those thought to represent global rise in the K16 model may suggest that basin-scale sea-level changes in the North Atlantic Ocean were comparable to global trends, or that the distribution of proxy data biases efforts to estimate global trends toward the data-rich Atlantic Ocean.

The diverse trends captured by the non-linear local component, as shown in Figure 8, highlight the important influence site-specific processes can have on the RSL. At Placentia Newfoundland, the non-linear local component experiences large fluctuations with maximum peaks reaching values of ∼0.19 m at around 250 CE and 1775 CE and minimum troughs of −0.12 m at around −450 CE and 1050 CE. This is a particularly pronounced degree of variability. The original study of the site by Kemp et al. (2018) recognized that the geomorphology at Placentia rendered it sensitive to site-specific RSL change due to the position of the salt marsh. The salt marsh is separated from the open ocean by a narrow inlet that is likely prone to opening and closing of the dynamic sediment barrier. In contrast, the East River Marsh record was generated exclusively through sediment in direct contact with bedrock to negate the potential influence of sediment compression as a driver of RSL change (Kemp et al., 2015). This contrast is reflected in our estimate of the non-linear local component where variability is present with a slight increase in sea level followed by a fall at around 100 CE and from 650 CE onwards an increase. Thus the component is non-zero due to the presence of other processes that can affect individual sites or groups of sites (e.g. dynamic sea level change; Kemp et al., 2015).

There are a number of potential extensions to the NI-GAM model that have not been addressed in our paper. A future aim is to extend NI-GAM further to larger regions, e.g. North Atlantic, or potentially to examine global RSL trends. By encompassing a broader geographical scope, we gain more understanding of additional drivers of sea-level change, for example dynamic sea level change arising from ocean dynamics (as defined in Gregory et al., 2019). Our extension would require an additional parameter to capture this important driver. This term could incorporate insights into ocean basin structure through a hierarchical prior, offering a comprehensive understanding of oceanic variations. By adopting this approach, we could explore the implications of individual coastlines separately within the new hierarchical term, alongside the collective influence of the ocean. However, this poses a challenge as the network of proxy records and tide gauges is non-uniformly spread and biased to coastal regions in the Northern hemisphere. Previous attempts to resolve this spatial bias have used a variety of techniques, (e.g. Berrett et al., 2020; Church & White, 2011; Dangendorf et al., 2017; Hay et al., 2015; Jevrejeva et al., 2008; Wenzel & Schröter, 2010), yet have mostly focused on instrumental data from tide gauges and satellites. Models like K16 and its extensions (e.g. Kemp et al., 2018; Khan et al., 2017; Walker et al., 2021) created a global component that may give insight into the changes in sea level common across many sites. Therefore, our model would require more components and further solutions to additional potential confounding issues. Yet, NI-GAM is an extendable modelling approach due the flexible structures of spline-GAMs and the Bayesian framework that allows for the inclusion of informed priors from future RSL analyses. Our modelling strategy is of course not limited to RSL changes. Rather it has the potential to be expanded to other areas of research that require the decomposition of a signal into different components that vary in time and space with complex measurement errors. One such example would be investigating historic temperature trends at a local and regional level to gauge the components that alter temperature spatially and temporally.

Acknowledgments

We express our gratitude to the Editor, Associate Editor, and the two Reviewers for their insightful comments and valuable suggestions.

Conflict of interest: None declared.

Funding

M.U.'s and G.M.'s work is supported by A4 (Aigéin, Aeráid, agus athrú Atlantaigh) project is funded by the Irish Marine Institute (grant: PBA/CC/18/01). A.P.'s work is supported by the Science Foundation Ireland awards 17/CDA/4695; 16/IA/4520; 12/RC/2289P2. A.K. is supported by a U.S. National Science Foundation CAREER award (OCE-1942563). E.A.'s research is funded by the U.S. National Science Foundation grant OCE-2002437 and OCE-2103754.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

N.C.'s research is conducted with the financial support of Science Foundation Ireland and co-funded by Geological Survey Ireland under Grant number 20/FFP-P/8610.

Data availability

The code and data for our model are available on Github and on Zenodo. The corresponding R package, `reslr` can be found on CRAN and on Github. The print for this manuscript is available on arXiv.

Table A1. The 21 data sites used in our model and the associated reference for each location

Reference and proxy indicator	Site name	Empirical rate prior to 1800 CE (mm/year)	ICE5G GIA rate (mm/year) (Peltier, 2004)
Donnelly et al. (2004): Plants (Gehrels et al., 2020): Foraminifera	Barn Island, Connecticut	1.00	0.99
Kemp et al. (2015) and Stearns and Engelhart (2017): Foraminifera	East River Marsh, Connecticut	0.91	0.96
Kemp et al. (2014): Foraminifera and plants	Nassau, Florida	0.42	0.28
Donnelly (2006): Plants	Revere, Massachusetts	0.58	0.40
Kemp et al. (2011): Plants	Wood Island, Massachusetts	0.52	0.40
Kemp et al. (2017, 2011): Foraminifera	Sand Point, North Carolina	0.97	0.69
Kemp et al. (2017, 2011): Foraminifera	Cedar Island, North Carolina	0.74	0.69
Kemp et al. (2013) and Cahill et al. (2016): Foraminifera, δ13C	Cape May Courthouse, New Jersey	1.19	1.24
Kemp et al. (2013) and Cahill et al. (2016): Foraminifera, δ13C	Leeds Point, New Jersey	1.69	1.41
Barnett et al. (2017): Foraminifera, testates	Les Sillons, Magdelen Islands	1.23	2.20
Gerlach et al. (2017): Foraminifera	Little Manatee River, Florida	0.28	0.14
Kemp et al. (2018): Foraminifera	Big River Marsh, Newfoundland	0.87	0.60
Kemp et al. (2018): Foraminifera, testates	Placentia, Newfoundland	0.41	0.21
Barnett et al. (2019): Foraminifera	Saint Simeon, Quebec	0.93	2.33
Gehrels et al. (2020): Foraminifera	Chezzetcook Inlet, Nova Scotia	1.76	0.63
Gehrels et al. (2020): Foraminifera	Sanborn Cove, Maine	3.23	0.08
Khan et al. (2022): Mangrove Peat	Snipe Key, Florida	0.66	0.13
Khan et al. (2022): Mangrove Peat	Swan Key, Florida	0.90	0.11
Walker et al. (2021): Foraminifera	Cheesequake, New Jersey	0.85	1.31
Kemp et al. (2017) and Stearns and Engelhart (2017): Foraminifera, δ13C	Pelham Bay, New York	0.8	1.31
Stearns and Engelhart (2017): Foraminifera, δ13C	Fox Hill Marsh, Rhode Island	1.00	1.07

Note. The site name is a combination of the site-specific name and the corresponding state. In addition, a comparison is made between the GIA rates we use from the data and GIA rates from physical models such as the ICE5G—VM2-90 Earth-ice model (Peltier, 2004).

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

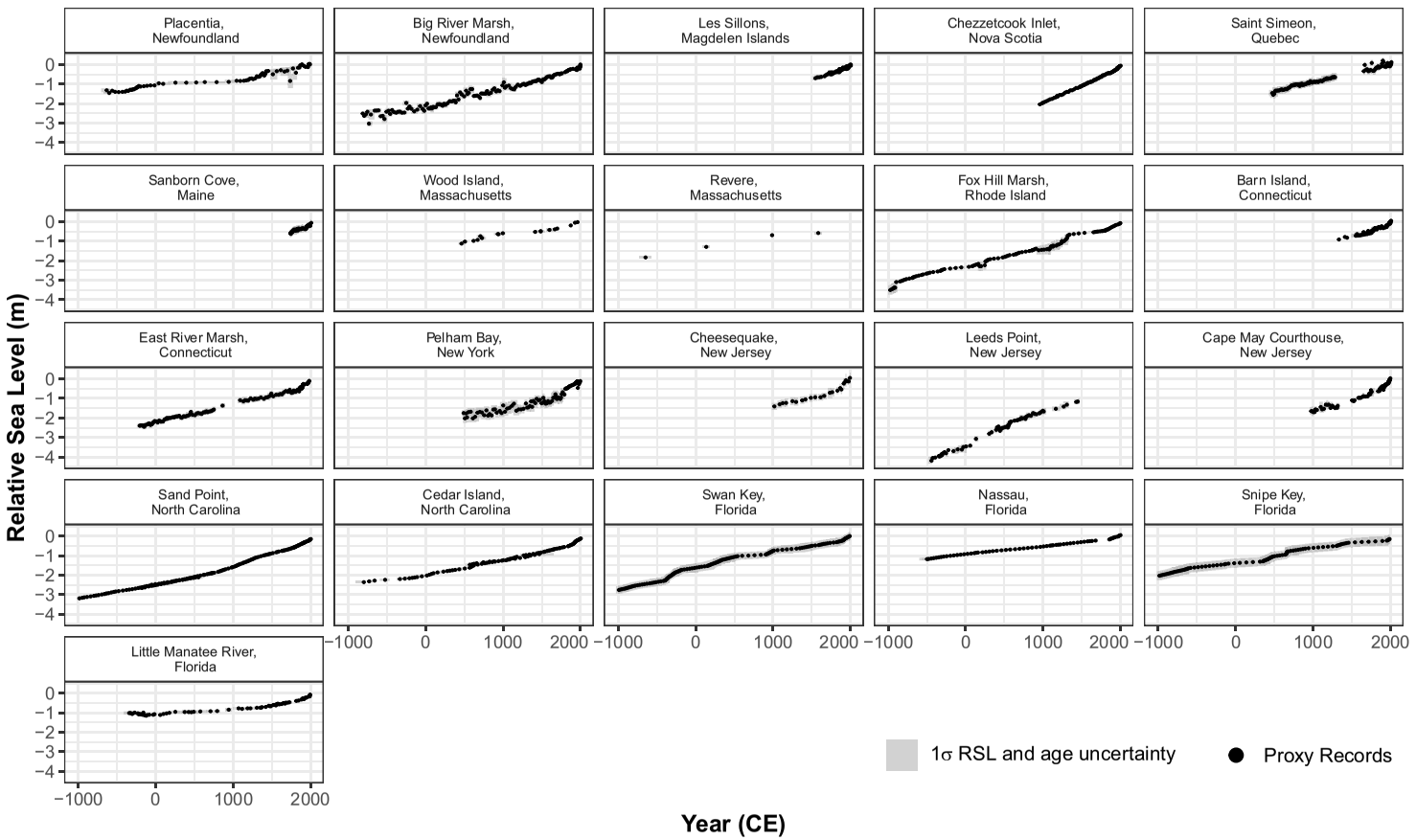

Figure A1. Proxy records for 21 sites along Atlantic coast of North America. The grey boxes represents the 1 σ uncertainty in RSL and age. The black points represents the midpoint of the uncertainty boxes that we use as the input of our data.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Appendix A

Data

Table A1 provides a list of all the proxy record sites used along the Atlantic coast of North America in our model and Figure A1 represents the proxy data associated with the 21 proxy record sites used in our model with the grey boxes the 1 σ uncertainty in the age and RSL value and the black dots the mid-point of the uncertainty box. Table A1 gives the references associated with each proxy data site and more information in regard to data collection can be sourced here. In addition, Table A1 contains the GIA rate used to inform the prior for the linear local component calculated using a linear regression for the data prior to 1800 CE for each site (Neukom et al., 2019). This provides an estimate for the rebounding effect of the tectonic plate after a glacier melts (Whitehouse, 2018). In previous models,

Table A2. The 66 tide-gauge data sites and their geographical coordinates used in our model from Holgate et al. (PSMSL 2013)

Longitude	Latitude	Site name	ICE5G-VM2-90 GIA rate (mm/year) (Peltier, 2004)
−54.00	47.30	ARGENTIA	0.21
−74.40	39.40	ATLANTICCITY	1.41
−68.20	44.40	BARHARBOR,FRENCHMANBAY,ME	−0.11
−76.70	34.70	BEAUFORT,NORTHCAROLINA	0.61
−63.60	44.70	BEDFORDINSTITUTE	0.49
−65.80	47.90	BELLEDUNE	1.98
−74.10	40.60	BERGENPOINT,STATENIS.	1.31
−53.10	48.70	BONAVISTA	0.36
−71.10	42.40	BOSTON	0.40
−64.00	44.70	BOUTILIERPOINT	0.49
−73.20	41.20	BRIDGEPORT	0.96
−70.60	41.70	BUZZARDSBAY	1.07
−61.90	47.40	CAPAUXMEULES	2.20
−75.60	35.20	CAPEHATTERAS,NORTHCAROLINA	0.69
−75.00	39.00	CAPEMAY	1.24
−83.00	29.10	CEDARKEYI	0.24
−63.10	46.20	CHARLOTTETOWN	1.67
−82.80	28.00	CLEARWATERBEACH	0.25
−67.20	44.60	CUTLER	0.08
−67.30	44.60	CUTLERII	0.08
−66.40	48.10	DALHOUSIE	2.12
−75.70	36.20	DUCKPIEROUTSIDE	0.67
−67.00	44.90	EASTPORT	0.08
−81.50	30.70	FERNANDINABEACH	0.42
−81.90	26.60	FORTMYERS	0.13
−70.70	43.10	FORTPOINT,NEWCASTLEISLAND	−0.30
−63.60	44.70	HALIFAX	0.49
−80.10	25.90	HAULOVERPIER	0.11
−75.10	38.60	INDIANRIVERINLET	1.13
−81.60	30.40	JACKSONVILLE	0.28
−81.00	24.70	KEYCOLONYBEACH	0.07

(*continued*)

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Table A2. Continued

Longitude	Latitude	Site name	ICE5G-VM2-90 GIA rate (mm/year) (Peltier, 2004)
−81.80	24.60	KEYWEST	0.13
−73.80	40.80	KINGSPOINT,NEWYORK	1.31
−58.40	49.10	LARKHARBOUR	−0.82
−75.10	38.80	LEWES(BREAKWATERHARBOR)	1.13
−64.90	47.10	LOWERESCUMINAC	2.24
−81.00	24.70	MARATHONSHORES	0.07
−81.40	30.40	MAYPORT	0.42
−81.40	30.40	MAYPORT(BARPILOTSDOCK),FLORIDA	0.42
−80.10	25.80	MIAMIBEACH	0.11
−72.00	41.00	MONTAUK	0.99
−76.70	34.70	MOREHEADCITY	0.61
−72.10	41.40	NEWLONDON	0.99
−71.30	41.50	NEWPORT	1.07
−73.80	40.90	NEWROCHELLE	1.31
−74.00	40.70	NEWYORK(THEBATTERY)	1.31
−60.20	46.20	NORTHSYDNEY	1.51
−75.50	35.80	OREGONINLETMARINA,NORTHCAROLINA	0.91
−75.10	39.90	PHILADELPHIA(PIER9N)	1.24
−62.70	45.70	PICTOU	0.93
−72.20	41.20	PLUMISLAND	0.99
−59.10	47.60	PORTAUXBASQUES	1.19
−73.10	41.00	PORTJEFFERSON	0.96
−76.30	36.80	PORTSMOUTH(NORFOLKNAVYYARD)	0.67
−71.40	41.80	PROVIDENCE(STATEPIER)	1.07
−64.40	49.00	RIVIERE-AU-RENARD	1.12
−63.30	46.50	RUSTICO	1.67
−70.50	41.80	SANDWICHMARINA, CAPECODCANALENTRANCE	1.20
−74.00	40.50	SANDYHOOK	1.31
−52.70	47.60	ST.JOHN'S,NFLD.	0.41
−82.60	27.80	ST.PETERSBURG	0.26
−55.40	46.90	STLAWRENCE	0.59
−81.10	24.70	VACAKEY	0.07
−80.20	25.70	VIRGINIAKEY,FL	0.11
−73.80	40.80	WILLETSPOINT	1.31
−70.70	41.50	WOODSHOLE(OCEAN.INST.)	1.07

Note. Also, the GIA rate for each tide gauge site is provided using the ICE5G—VM2-90 Earth—ice physical model developed by Peltier (2004).

physical GIA models are used to inform the prior for the linear local component. Peltier (2004) developed the ICE5G—VM2-90 Earth—ice physical model that provides a GIA rate for each site. We carried out a comparison between these techniques however, the data driven GIA rates were our preferred choice for the proxy records.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

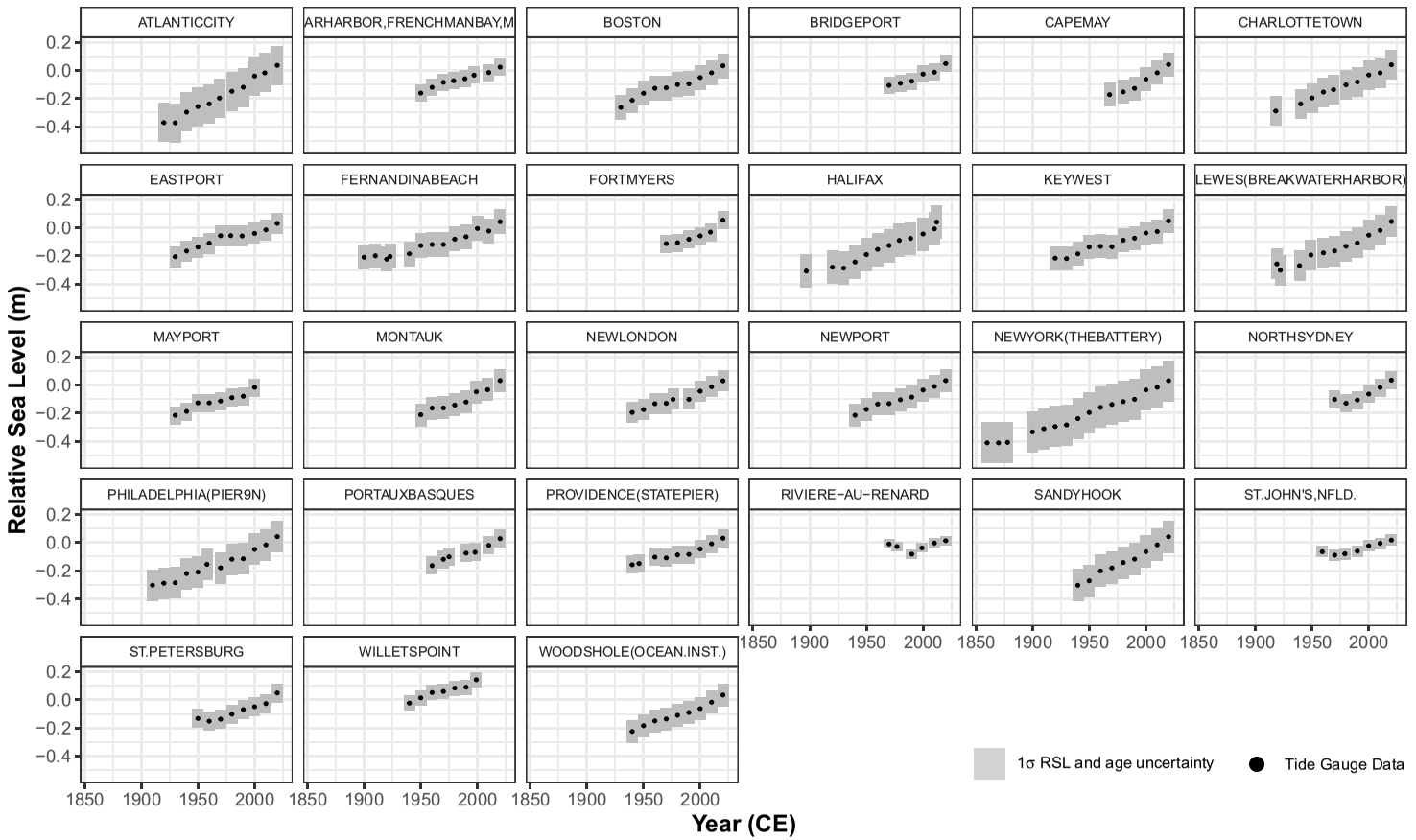

Figure A2. Tide gauge data for a subset of locations with observations greater than 50 years along the Atlantic coast of North America. The grey boxes represents the 1 σ uncertainty in RSL and age. The black points represents the midpoint of the uncertainty boxes that we use as the input of our data.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Table A2 gives the reference name associated with each tide-gauge data site and its location along North America's Atlantic coast using PSMSL database (Holgate et al., 2013). For each tide gauge site, the associated GIA rate is provided which is obtained using the ICE5G—VM2-90 Earth—ice physical model developed by Peltier (2004). These GIA rates serve as the mean values, denoted as m_{g_j}, for each site j in the prior distribution of the linear local component, as described in Section 5.4, equation (13). Additionally, following the findings of Engelhart et al. (2009), we set the uncertainty associated with the GIA rate values for the tide gauge data to be 0.3 mm per year. This uncertainty value is used as the known variance, $s^2_{g_j}$, for the prior distribution for the linear local component as described in Section 5.4, equation (13). In Figure A2, we plotted a subset of 27 tide gauge sites with more than 50 observations each. The tide gauge data is averaged over a decade in order to match the sediment accumulation rate of the proxy records which is approximately every 10 years along the Atlantic coast of North America.

Appendix B

Model Results and Validations for the Full Data Set

Section 7 showcases the results of four illustrative case studies and it is important to note that the model was executed on the complete dataset. In this section, we delve into the comprehensive results derived from our extensive dataset comprising 21 proxy sites and 66 tide gauges. In addition, we present the results from the 10-fold cross-validation using the 21 proxy sites. Finally, we explore the effects of altering the number of knots and types of basis functions in the GAM Sensitivity test section.

Results for Full Dataset

The model is run using 21 proxy sites and 66 tide gauge sites, yet we present the results of the proxy record sites as their long temporal trend provides insight into long-term changes in RSL along the Atlantic coast of North America. Figure B2 provides the total model fit for the 21 proxy sites along the Atlantic coast of North America, and Figure B3 provides the rates of change for the

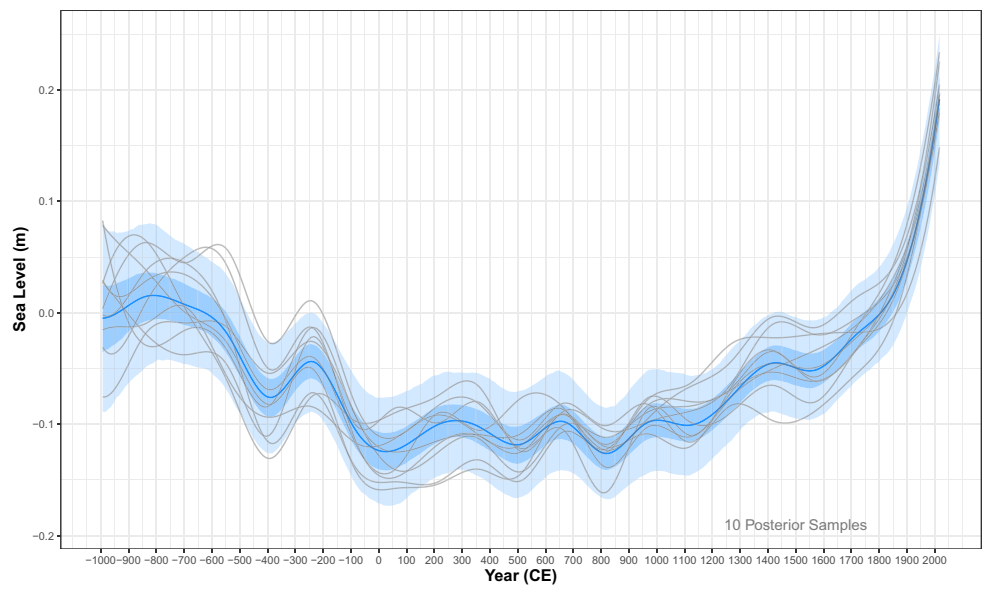

Figure B1. Regional component ($r(t)$) of the noisy-input generalized additive model using 21 proxy sites and 66 tide gauge sites along the Atlantic coast of North America. The dark blue line highlights the mean posterior model fit and the dark blue shading indicated the 50% credible interval and the lighter blue shading is the 95% credible interval. The grey lines represent 10 posterior samples.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

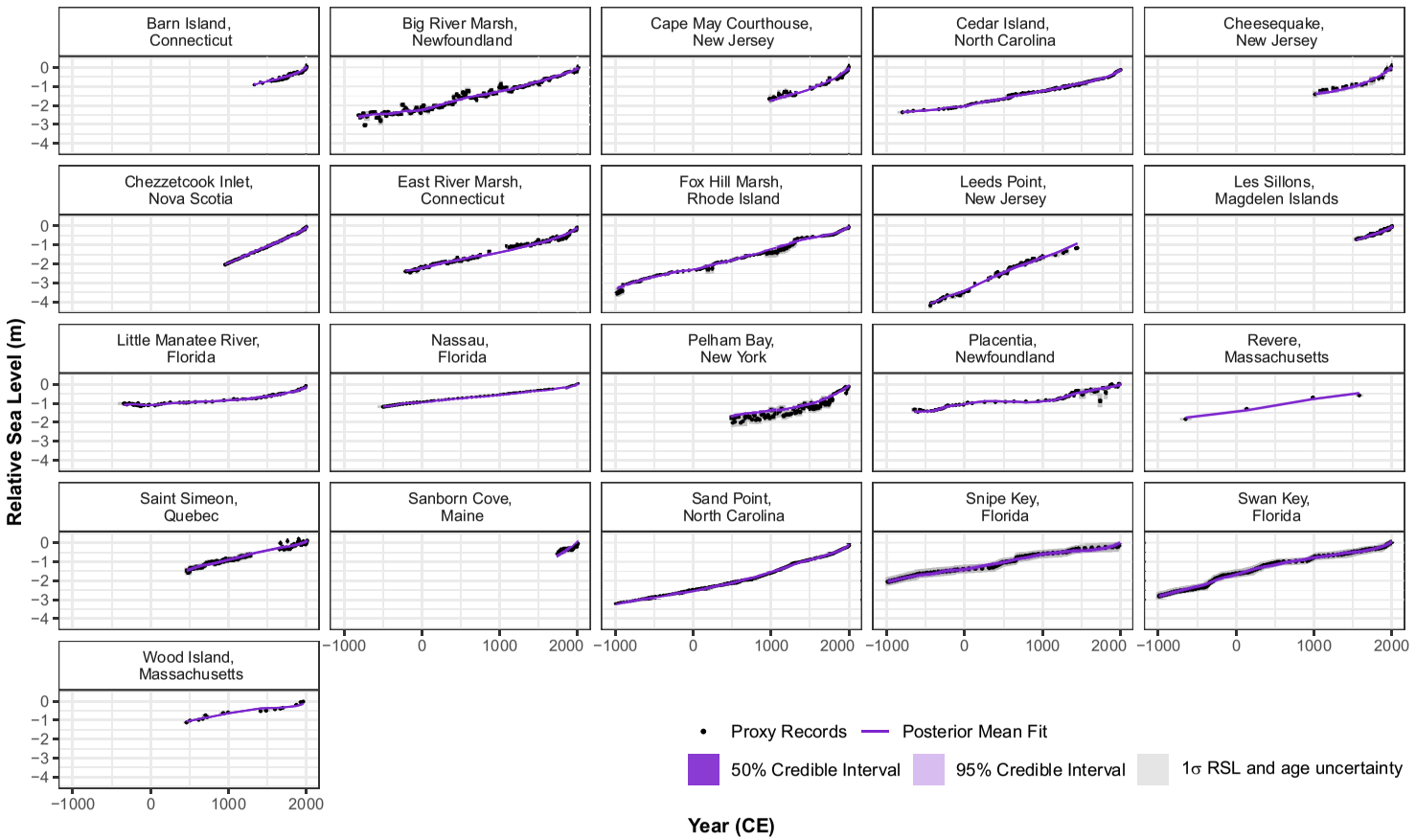

Figure B2. The noisy-input generalized additive model (NI-GAM) fit ($f(\mathbf{x}, t)$) for 21 proxy sites along the Atlantic coast of North America. The black dots and grey boxes represent the midpoint and associated uncertainty, respectively, for each proxy record. The solid purple line represents the mean of the model fit with a 95% credible interval denoted by shading.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

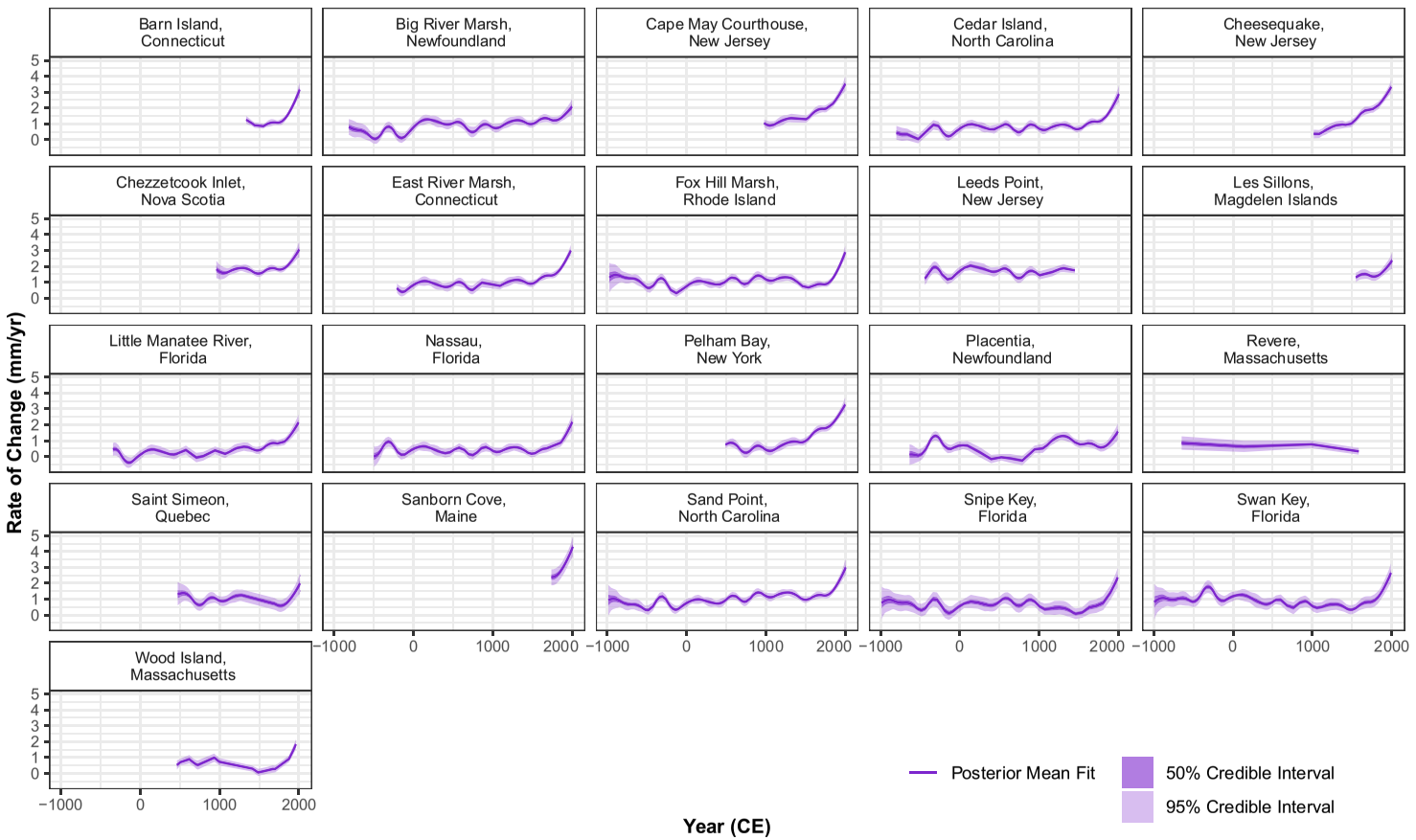

Figure B3. Rate of relative sea change found by taking the first derivative of the total model fit ($f'(\mathbf{x}, t)$) for 21 proxy sites along the Atlantic coast of North America. The mean of the fit is the solid purple line with the dark shaded area being the 50% credible interval and the light shaded area being the 95% credible interval.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qiae044/7756353 by Maynooth University user on 23 September 2024

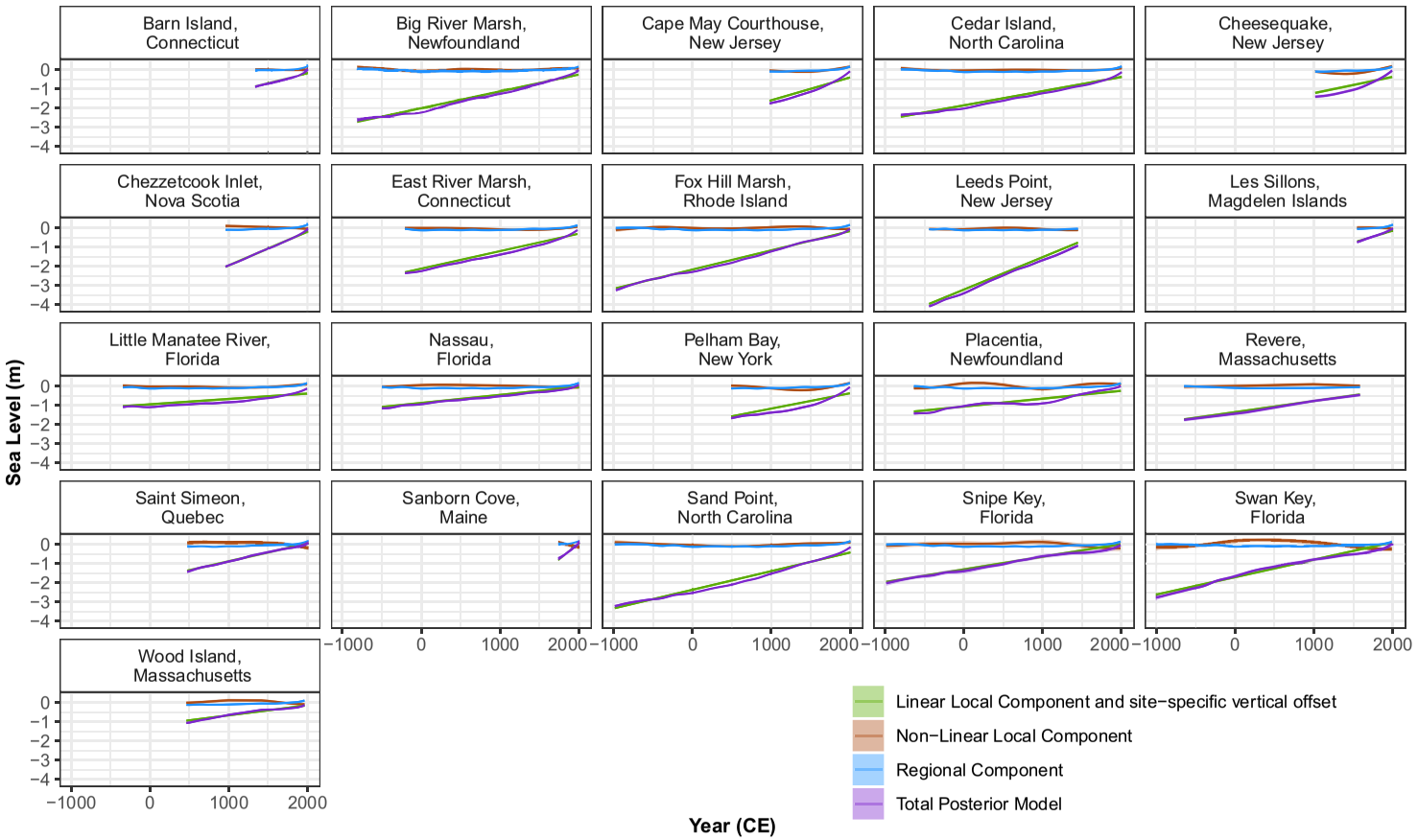

Figure B4. All components of the NI-GAM for the 21 proxy sites along the Atlantic coast of North America. The regional component (h) is in blue with a 95% credible interval. The linear local component and the site-specific vertical offset ($g + h$) are green with a 95% credible interval. The non-linear local component (l) is brown with a 95% credible interval. The total posterior model fit ($f(\mathbf{x}, t)$) is purple with 95% credible interval. The x-axis is in years and y axis is in metres.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

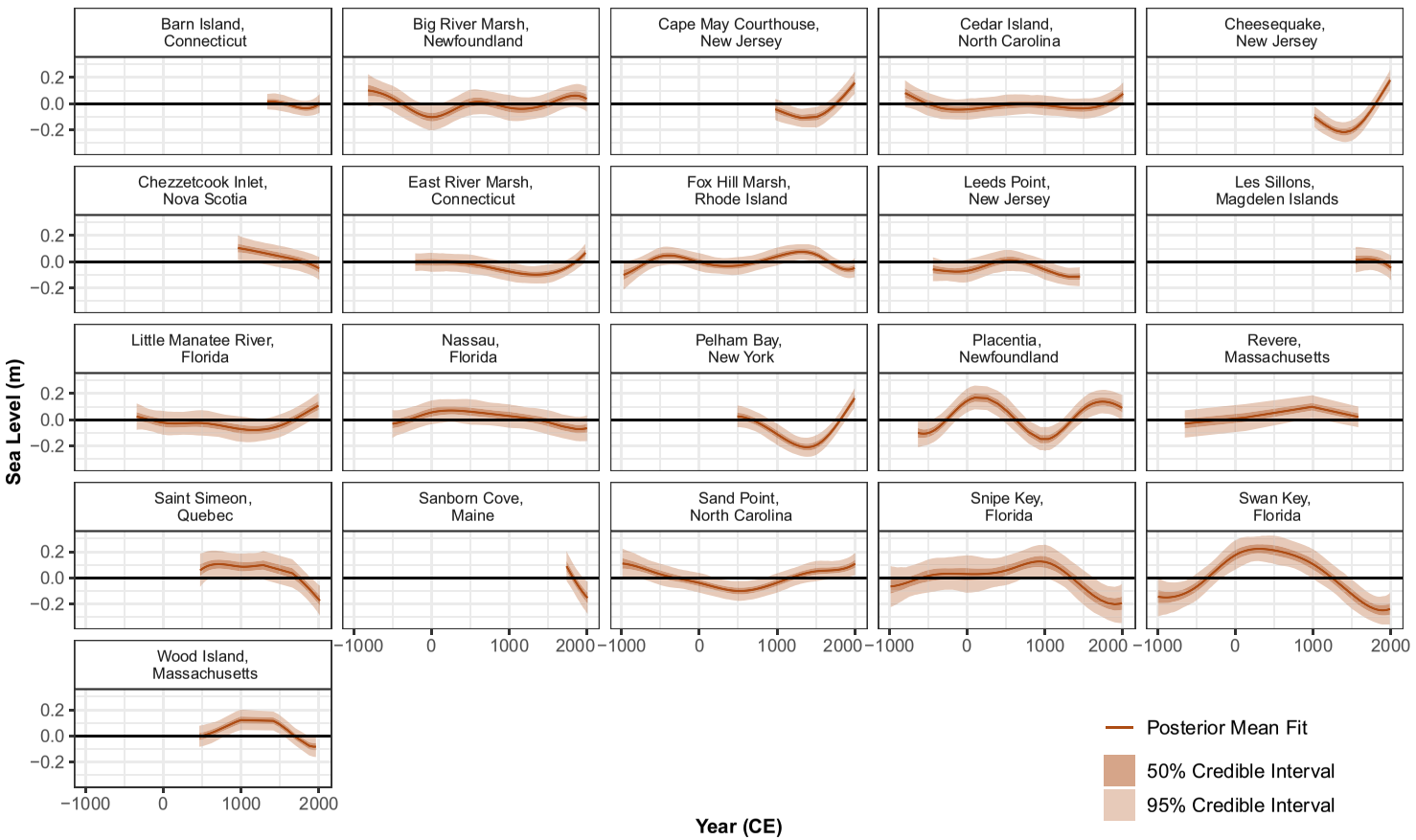

Figure B5. The non-linear local component ($l(\mathbf{x}, t)$) for all proxy sites along the Atlantic coast of North America. The *y*-axis represents sea level in metres. The brown solid line represents the mean of the model fit with the 50% credible interval in dark brown shading and 95% credible interval in the light brown shading.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

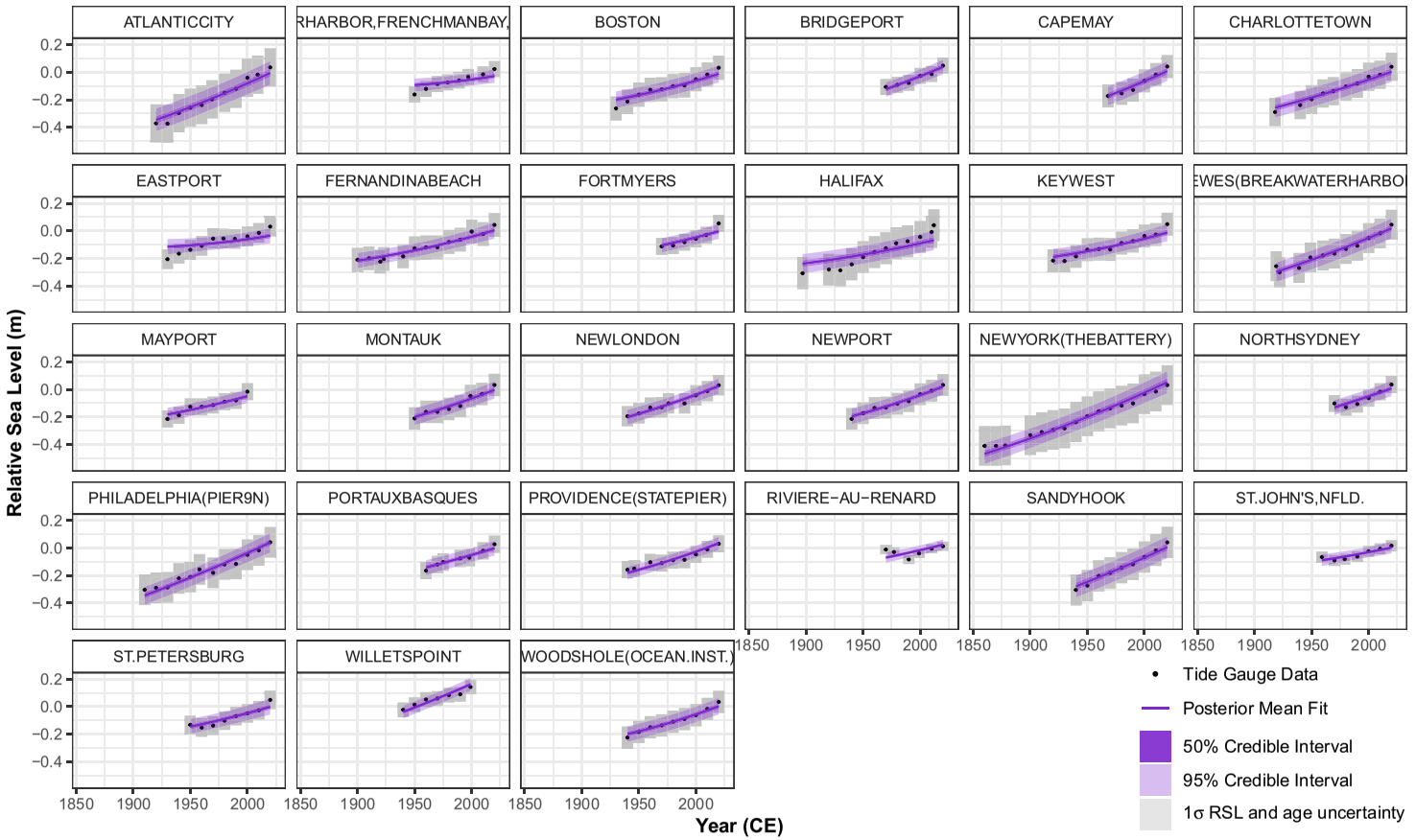

Figure B6. The noisy-input generalized additive model (NI-GAM) fit ($f(\mathbf{x}, t)$) for tide gauge sites, with data greater than 50 years of data, along the Atlantic coast of North America. The black dots and grey boxes represent the midpoint and associated uncertainty, respectively, for each proxy record. The solid purple line represents the mean of the model fit with a 95% credible interval denoted by shading.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

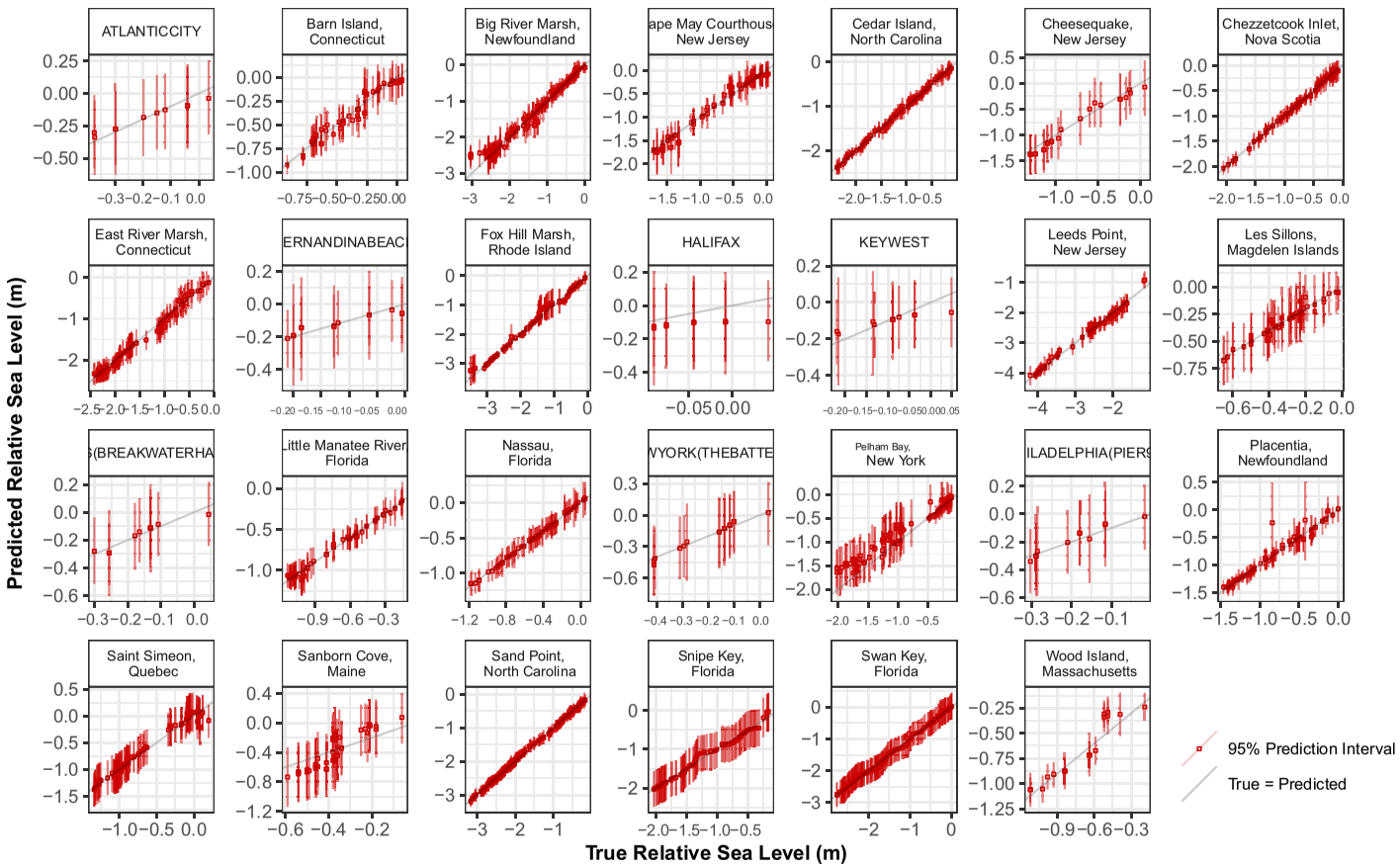

Figure B7. True vs. predicted RSL for all the 20 proxy and 7 tide gauge sites along the Atlantic coast of North America using 10-fold cross-validation. The grey line indicating the identity line. The red points are the predicted means and the red vertical lines represent the 95% prediction interval for each point.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Table B1. Empirical 95% coverage for the 20 proxy sites (head of table) and 7 tide gauge (tail of table) sites along the Atlantic coast of North America with the associated prediction interval (PI) using 10-fold cross-validation

Site name	Empirical 95% coverage	95% PI width (m)	Empirical 50% coverage	50% PI width (m)	RMSE (m)
Barn Island, Connecticut	1.00	0.33	0.53	0.11	0.09
Big River Marsh, Newfoundland	0.95	0.48	0.62	0.17	0.13
Cape May Courthouse, New Jersey	0.99	0.56	0.75	0.19	0.15
Cedar Island, North Carolina	1.00	0.29	0.85	0.10	0.07
Cheesequake, New Jersey	1.00	0.82	0.91	0.28	0.21
Chezzetcook Inlet, Nova Scotia	1.00	0.30	0.96	0.10	0.08
East River Marsh, Connecticut	1.00	0.53	0.74	0.18	0.14
Fox Hill Marsh, Rhode Island	0.96	0.44	0.66	0.15	0.13
Leeds Point, New Jersey	0.99	0.50	0.64	0.17	0.13
Les Sillons, Magdelen Islands	1.00	0.42	0.89	0.15	0.11
Little Manatee River, Florida	1.00	0.32	0.84	0.11	0.08
Nassau, Florida	1.00	0.33	1.00	0.12	0.08
Pelham Bay, New York	0.97	0.76	0.45	0.26	0.21
Placentia, Newfoundland	0.99	0.36	0.81	0.12	0.10
Saint Simeon, Quebec	1.00	0.65	0.90	0.22	0.16
Sanborn Cove, Maine	1.00	0.74	0.43	0.26	0.19
Sand Point, North Carolina	1.00	0.35	0.92	0.12	0.09
Snipe Key, Florida	1.00	0.93	0.98	0.32	0.23
Swan Key, Florida	1.00	0.77	1.00	0.27	0.19
Wood Island, Massachusetts	0.83	0.30	0.33	0.10	0.07
ATLANTICCITY	1.00	0.60	1.00	0.21	0.15
FERNANDINABEACH	1.00	0.45	1.00	0.15	0.11
HALIFAX	1.00	0.53	0.75	0.18	0.13
KEYWEST	1.00	0.41	0.91	0.14	0.10
LEWES(BREAKWATERHARBOR)	1.00	0.50	1.00	0.18	0.12
NEWYORK(THEBATTERY)	1.00	0.63	1.00	0.22	0.16
PHILADELPHIA(PIER9N)	1.00	0.52	1.00	0.18	0.13

Note. As a comparison, the prediction intervals are reduced to 50% intervals and the empirical coverage for the 50% is presented. The root mean square error (RMSE) is included in metres

corresponding 21 proxy sites. Figure B4 provides the decomposition of the NI-GAM into the total model fit and the three components; regional, linear local component with the site-specific vertical offset and non-linear local component for the 21 proxy sites. Figure B1 presents the regional component of the NI-GAM. The grey lines represent 10 randomly chosen posterior samples showing the underlying behaviour of the posterior for the regional component. Figure B5 provides the non-linear local component for each proxy record data site. In Figure B6, we illustrate the total model fit for tide gauge data where the data exceed 50 years of observations.

Model Validation

Model validation using 10-fold cross-validation was undertaken using data from 20 proxy sites and 7 tide gauge sites along the Atlantic coast of North America. The remaining data sites were not used in the 10-fold cross-validation as all sites had fewer than 10 data points. Figure B7 provides the true vs. predicted RSL for the 20 proxy sites and 7 tide gauge sites using 10-fold cross-validation. In Table B1, we provide the empirical coverage for each data site in

Downloaded from https://academic.oup.com/jrssc/advance-article/doi/10.1093/jrssc/qlae044/7756353 by Maynooth University user on 23 September 2024

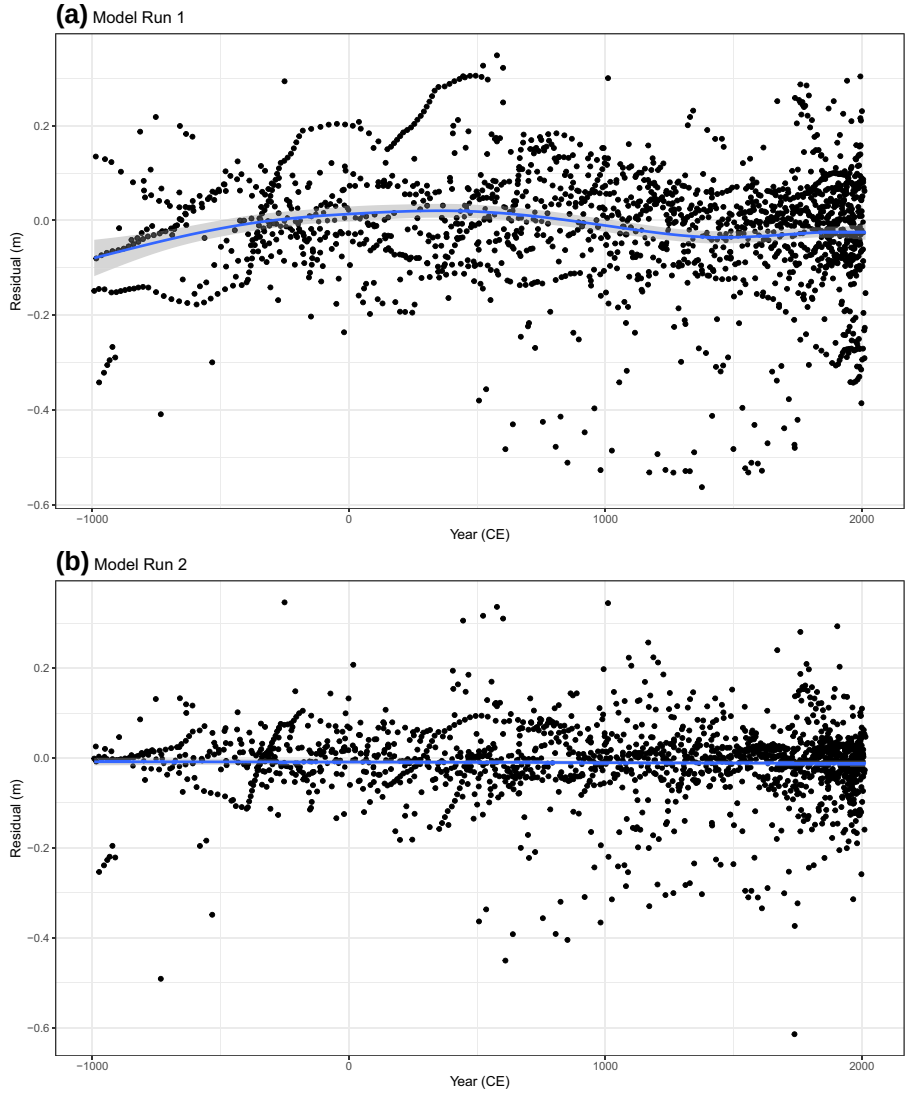

Figure B8. Residual analysis for the noisy-input generalized additive model highlighting. (a) Model run 1 represents the first step of our modelling approach and the residuals possess evidence for an underlying trend. (b) Model run 2 presents the second step of our model which removed the underlying trend, highlighting the requirement of the two step modelling approach.

the 10 CV analysis. The empirical coverage indicates the percentage of times the true observation lies within the prediction interval. A comparison is made between the 95% empirical coverage and the 50% empirical coverage. It is evident that the prediction intervals for our model are large resulting in 100% coverage in many sites. This is due to the large size of the prediction intervals resulting from the large bivariate uncertainty that arises from the proxy records. Also included in Table B1 is the root mean square error (RMSE) for the 20 proxy sites and 7 tide gauge sites along the Atlantic coast of North America that gives an insight into the prediction errors.

Residual Analysis

To evaluate the efficacy of our models, we analyse the residuals, representing the differences between observed and predicted values, through residual plots (Baddeley et al., 2005). In

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

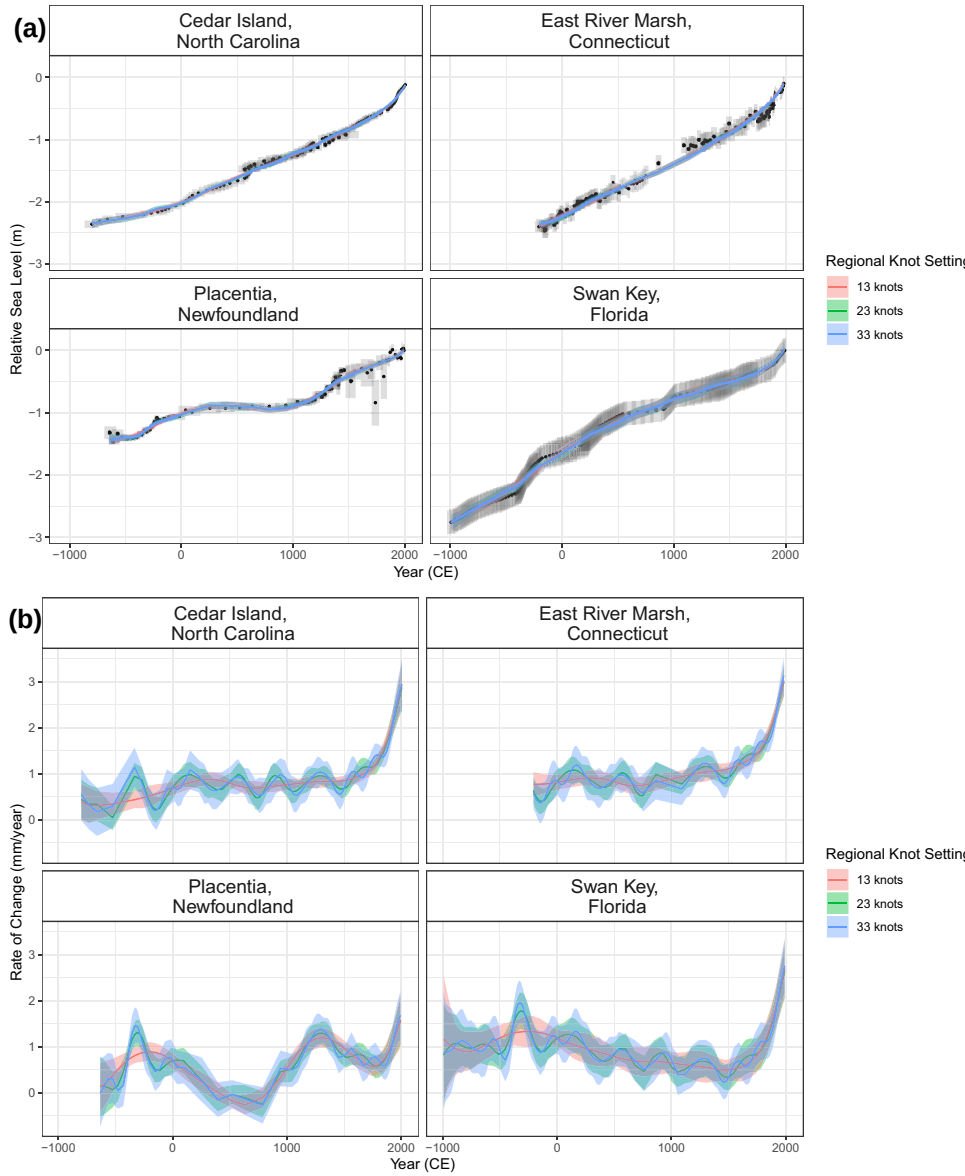

Figure B9. (a) Total model fit for the noisy-input generalized additive model using a range of knot values for the regional component. The grey boxes represent the bi-variate uncertainties associated with the proxy records and the black dot is the midpoint of the box. The coloured lines with corresponding 95% credible intervals represent a subset of sensitivity tests used to find the optimum regional knot setting. (b) The rate of change for the total model fit for the noisy-input generalized additive model using a range of knot values. The coloured lines with corresponding 95% credible intervals represent a subset of sensitivity tests used to find the optimum regional knot setting.

Figure B8, our residual plot provides insights into model performance, aiding in the identification of patterns, trends in what should be identified as random noise. Figure B8a Model run 1 represents the residuals after the first step of our modelling approach and demonstrates a discernible trend, indicating that the model has not fully captured underlying patterns in the data. In contrast, Figure B8b Model run 2 presents the residual plot after the second model run that eliminates the discernible trend, presenting residuals that are random and evenly spread around the horizontal axis. By comparing the residual plots after each step of the

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

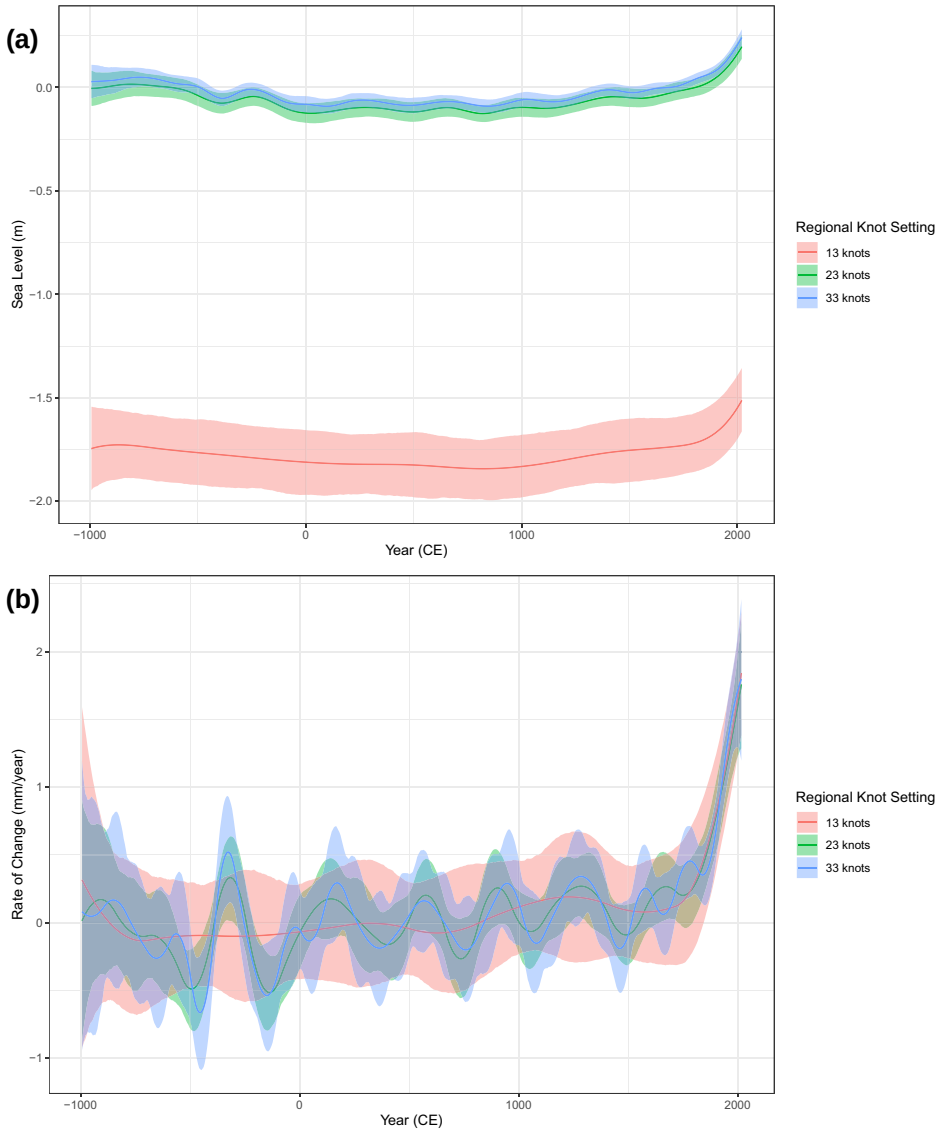

Figure B10. (a) The regional component for the noisy-input generalized additive model with the coloured lines and corresponding 95% credible intervals highlighted a subset of the knot settings examined during our sensitivity analysis. (b) The rate of change for the regional component for the noisy-input generalized model. The coloured lines represent 95% credible intervals for three different knot settings for the regional component, ranging from 13–3 knots. Through our GAM sensitivity analysis, we deemed 23 knots as the optimum value for the regional component knot setting.

modelling approach, we justify the adoption of a two-step modelling process, as outlined in Section 5.

GAM Sensitivity Study

As mentioned in Section 5, we use B-splines for our regional and non-linear local terms. In this section, we utilize a GAM sensitivity study to highlight the range of basis functions settings examined and how our model choices provided satisfactory model fits as well as computational efficiencies.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

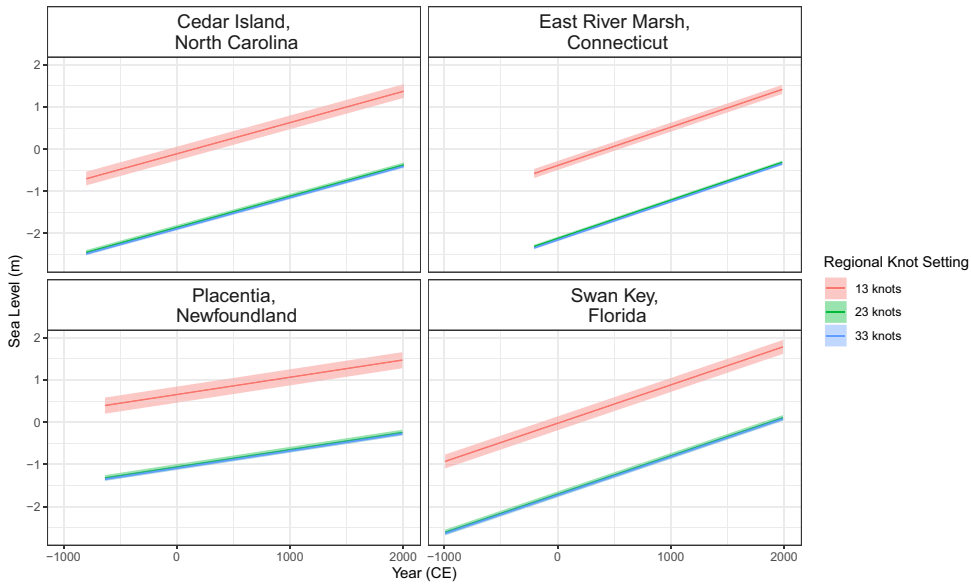

Figure B11. The linear local component for the noisy-input generalized additive model with the coloured lines and corresponding 95% credible intervals highlighted a subset of the knot settings for the regional component that were examined during our sensitivity analysis. The coloured lines represent three knot settings for the regional component; 13 knots in red, 23 knots in green, and 33 knots in blue.

Initially, we assessed the capabilities of the `mgcv` package developed by Wood (2015), which facilitates the fitting of various smoother functions and generalized additive models. However, this package lacks provisions for Bayesian analysis or the incorporation of uncertainties associated with input and output variables. Subsequently, we explored the `jagam` package by Wood (2016), leveraging the functionalities of `mgcv` while automatically generating corresponding JAGS model code, enabling Gibbs sampling. Nevertheless, challenges emerged when incorporating informed priors for the linear local component, leading to issues with model fits at certain locations. Consequently, to enhance flexibility in the model structure, we developed our own basis functions and JAGS scripts.

Following a comprehensive examination, we investigated both the individual components, the overall fit of the model and the corresponding rates for both. Our residual analysis demonstrated the importance of the diverse components within the RSL field. Additionally, it highlighted the necessity for conducting multiple model runs to effectively incorporate uncertainty and capture inherent patterns in the data. At the development stage of model, we commenced with the most basic model involving a spline in time and systematically progressed through a spline in space-time, eventually incorporating additional random effect terms. It is crucial to highlight that the structure and smoothness of basis functions associated with our regional (spline in time) and non-linear local components (spline in space-time) play a pivotal role in enhancing the performance and interpretability of statistical models. To create the basis functions, we draw inspiration from Eilers and Marx (2021). Central to our function are two key parameters: 'nseg' that determines the partitioning length for equidistant knots, and 'deg' that specifies the degree of the underlying basis function. We systematically tested both the number of segments ('nseg') and the degree, thereby adjusting the number of knots accordingly. Leveraging the bs() function from Hastie (2017), we constructed the spline basis function from these calculated knots. Subsequently, we enhanced model efficiency by removing columns with zero values. In this section, we present our GAM sensitivity tests, offering a representative sample of the extensive experimentation undertaken. We explore three distinct configurations: one with a small number of knots, a medium setting, and a scenario involving a large number of knots, applied to both the spline in time and the spline in space-time. Despite exploring various alternatives, our cross-validation analyses consistently favoured a moderate number of knots for both time and space-time spline models, achieving a balance between

Downloaded from https://academic.oup.com/jrssc/advance-article/doi/10.1093/jrssc/qlae044/7756353 by Maynooth University user on 23 September 2024

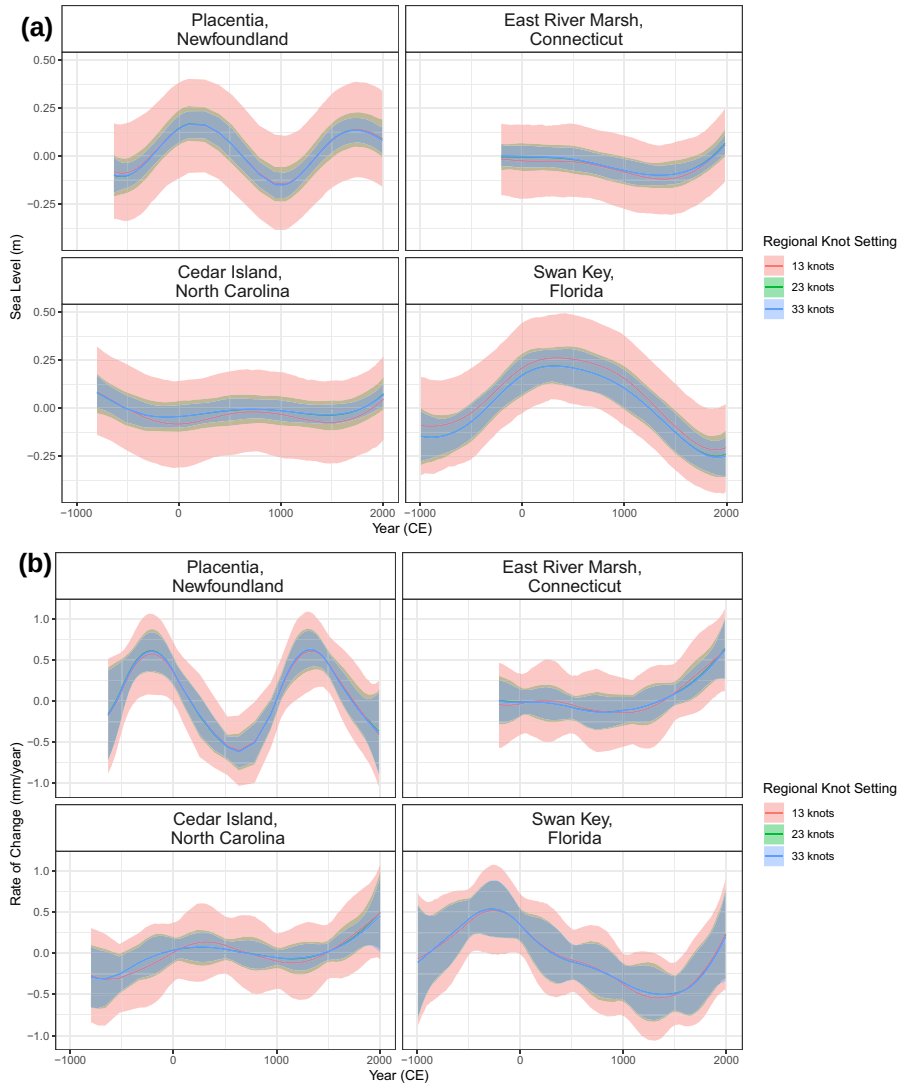

Figure B12. (a) The non-linear local component for the noisy-input generalized additive model with the coloured lines and corresponding 95% credible intervals highlighted a subset of the regional component knot settings. (b) The rate of change for the regional component for the noisy-input generalized model. The coloured lines represent 95% credible intervals for three different knot settings for the regional component, ranging from 13–33 knots.

model fit and computational efficiency. While the `mgcv` package (Wood, 2015) used a generalized cross-validation technique to reduce computational demands, we opted for the 10-fold cross-validation method to better capture the complexity inherent in our data.

Regarding the regional component, we conducted comparisons with B-splines (Hastie & Tibshirani, 1986) and P-splines (Eilers & Marx, 1996). While both spline types involve coefficient estimation, P-splines introduce the added complexity of estimating penalty parameters to regulate smoothness (Wood, 2017). Unfortunately, issues arose with model convergence due to the additional penalty parameter for the P-spline. This was especially notable for the non-linear component, which involved modelling 3D splines across varied spatial sites and time series lengths. In this case, uneven spacing between knots was required, leading to the estimation of numerous coefficients and penalties for each site and time frame, which was considerable when examining 21 proxy locations and 66 tide gauge sites. As suggested by Wood (2017), the efficiency of P-splines diminishes with unevenly spaced knots, and identifiability issues arose with the complex

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

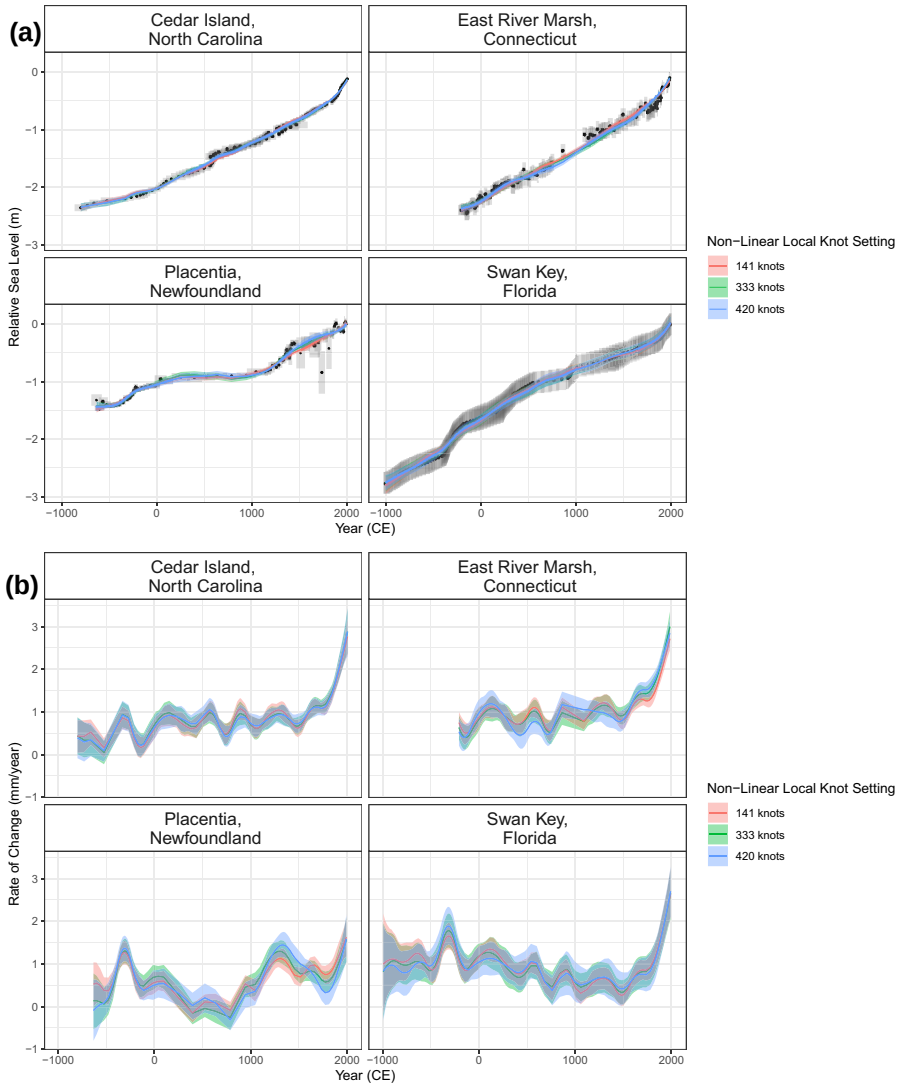

Figure B13. (a) Total model fit for the noisy-input generalized additive model using a range of knot values for the non-linear local component. The grey boxes represent the bi-variate uncertainties associated with the proxy records and the black dot is the midpoint of the box. The coloured lines with corresponding 95% credible intervals represent a subset of sensitivity tests used to find the optimum non-linear local knot setting. (b) The rate of change for the total model fit using range of knot values for the non-linear local component. The number of knots for the non-linear local component ranges from 141 (red) to 333 (green) to 420 (blue).

structure of the non-linear local component. Consequently, we reverted to using B-splines and focused on fine-tuning the number of knots and validated our approach using 10-fold cross-validation, as discussed earlier. Figure B9a displays the overall model fit for the NI-GAM, and Figure B9b shows the corresponding rate of change for the model fit for the NI-GAM. We demonstrate a subset of sensitivity tests that we conducted to explore the interplay between model smoothness and efficiency focusing on the regional component and the corresponding rate. We found that examining the rate provides a very important insight into the required smoothness of the total NI-GAM. The plots demonstrates how the optimal fit for the data is achieved by selecting a range of knots between 13 to 33 knots for the regional component and the corresponding rate of change. In this case, 13 knots (red) results in an overly smooth fit, while 33 knots (blue) lack

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

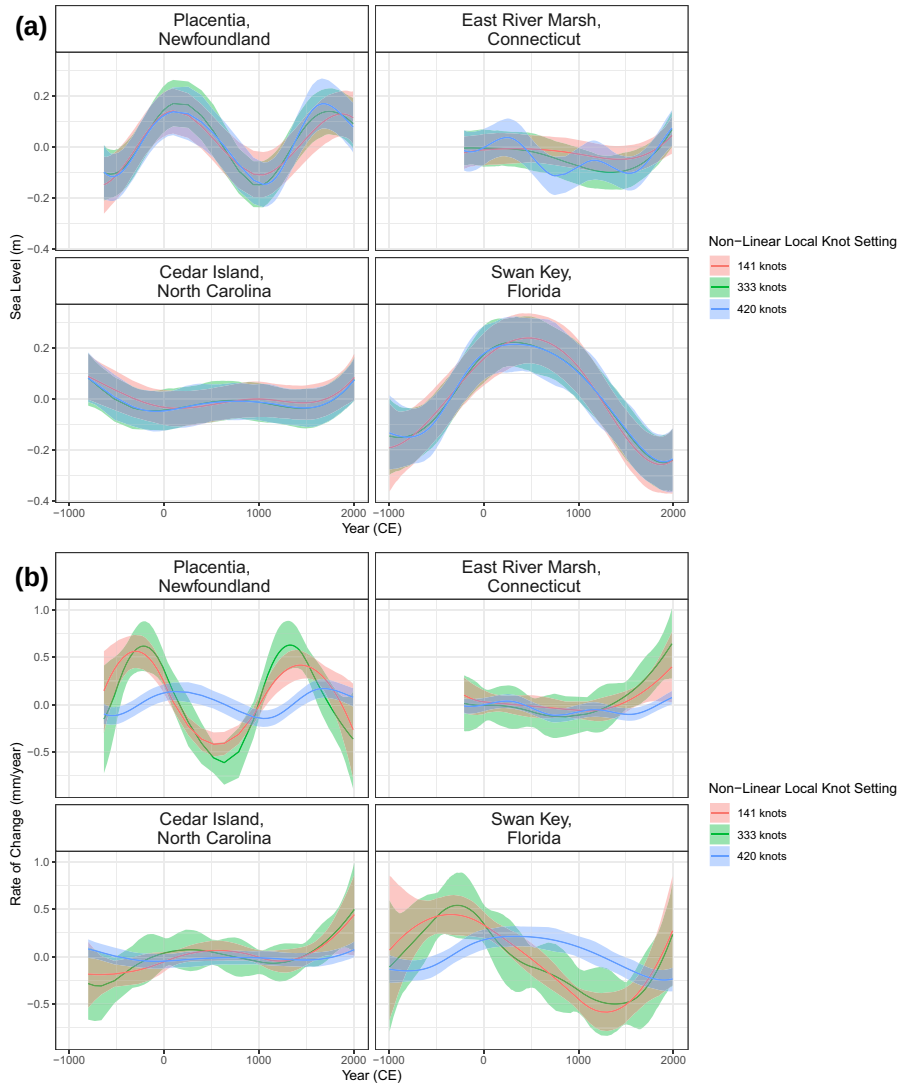

Figure B14. (a) The non-linear local component for the noisy-input generalized additive model with the coloured lines with corresponding 95% credible intervals representing a subset of knot settings. (b) The rate of change for the non-linear local component using range of knot values for the non-linear local component. The number of knots for the non-linear local component ranges from 141 (red) to 333 (green) to 420 (blue).

smoothness. With 23 knots (green), we achieve a balance, ensuring smoothness in the rates while effectively capturing trends in the data.

When investigating the sensitivity of the regional knot setting, we assessed the total model fit and corresponding rate as described previously, but also the individual components of the RSL field and their corresponding rates. Through the combination of model fit and smoothness of rates, an optimum number of knots was selected. In Figure B10a, we illustrate the different knot setting for the regional component, and Figure B10b depicts the corresponding rate for the regional component. When the 13 knots is used, as depicted in read in Figure B10, the large credible intervals and the varying intercept reflect how the model is struggling to identify which component should pick up the variability. This demonstrates how selecting the number of knots depends on the individual component variability and how the components interact with each other. We identified that the prefer setting to be 23 knots as it significantly reduced uncertainty compared to the

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

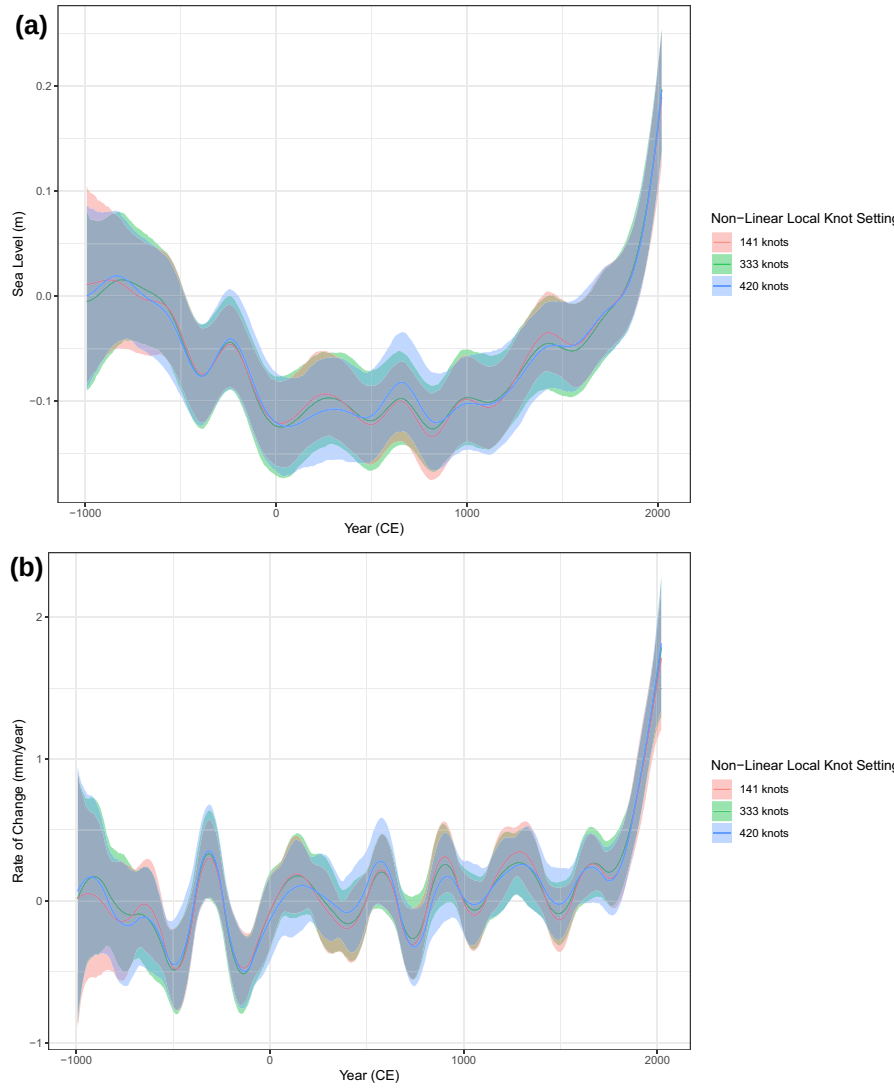

Figure B15. (a) The regional component for the noisy-input generalized additive model with the coloured lines with corresponding 95% credible intervals representing a subset of knot settings. (b) The rate of change for the regional component using range of knot values for the non-linear local component. The number of knots for the non-linear local component ranges from 141 (red) to 333 (green) to 420 (blue).

alternative with 13 knots and effectively balanced the variability among the remaining components of the NI-GAM. Following this, we investigate the impact of the regional component knot setting on the linear-local component. In Figure B11, we demonstrate how varying the number of knots in the regional component reduces the size of the credible intervals for the linear local component and how the intercept (i.e. y offset term) varies. This underscores the importance of effectively managing confounding variables while balancing individual components, highlighting the crucial role of optimizing knot settings, especially for the regional component.

The final step for the sensitivity test for the regional component knot setting is to investigate the varying knots for the non-linear local component. In Figure B12a and b, we show how 13 knots (red) results in large uncertainties for the non-linear local component and using 23 or 33 knots reduced this uncertainty. In instances where one component exhibits less flexibility, the remaining component tends to become more variable. Consequently, we conducted these tests by maintaining a fixed setting for the non-linear local component knots while varying the regional component

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

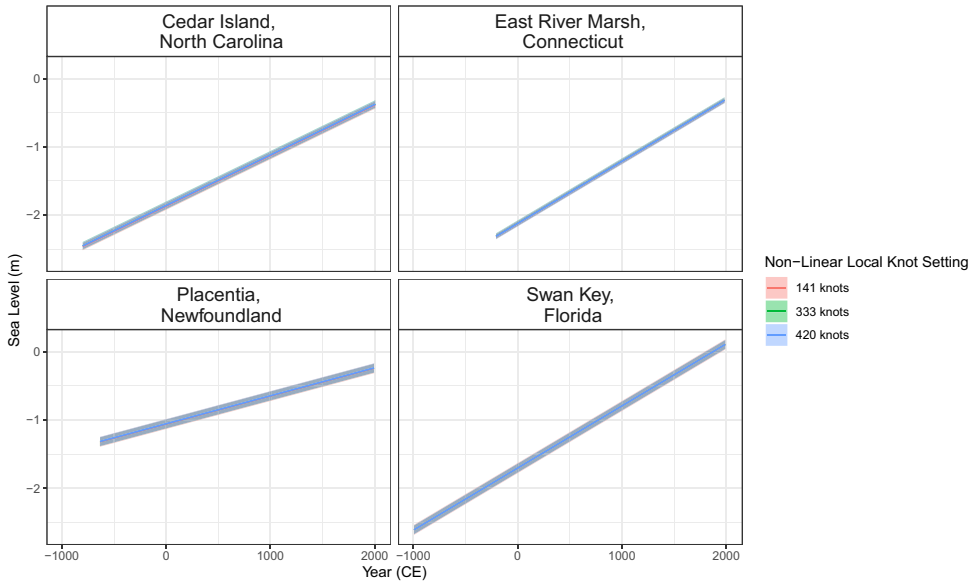

Figure B16. The linear local component for the noisy-input generalized additive model with the coloured lines and corresponding 95% credible intervals highlighted a subset of the knot settings for the non-linear local component that were examined during our sensitivity analysis. The coloured lines represent three knot settings for the regional component; 141 knots in red, 333 knots in green, and 420 knots in blue.

knot configuration. In the following section, we demonstrate the outcomes of modifying the non-linear local component knot setting and examine its effects on the overall model fit, as well as the other components and their corresponding rates.

A similar GAM sensitivity analysis was carried out for the knot settings associated with non-linear local component. This component requires a tensor product which results in significantly larger number of knots compared with the previous tests for the regional component knot setting. In Figure B13a and b, we show the overall model fit and rate for the NI-GAM with a subset of knot settings. Similar to the previous case, tests conducted to explore the interplay between model smoothness and efficiency focusing on the non-linear local component. Figure B13a and b demonstrates a subset of the knot settings examined for the non-linear local component. It is clear that varying the non-linear knots has minimal impact on the total model fit and the corresponding rate. This is in contrast to the impact of the regional knot setting shown in Figure B9b which impacts the smoothness greatly.

Figure B14a and b represents the impact the knot setting has on the non-linear local component. In this case, there are differing results based on the knot setting employed. We selected the knot setting 333 (green) based on a combination of model fit and smoothness in the rates.

In Figure B15a and b, highlights how the knot setting for the non-linear local component has minimal impact on the regional component. This results reflects the two step modelling procedure we employ and how the remaining variability is captured within the non-linear local component.

Similarly, we demonstrate the impact of the non-linear local component knot setting on the linear-local component. In Figure B16, we show how varying the number of knots in the non-linear local component does not impact the linear local component. Once again, this underscores our two-step modelling procedure and highlights how the remaining variability is contained within the non-linear local component.

In this GAM sensitivity study, we have demonstrated how basis functions serve as fundamental components, playing a crucial role in representing the relationship between predictors and the response variable in a flexible and non-linear manner. We have presented a subset of our sensitivity study, which helps address the modelling decisions we made to ensure the smoothness of the basis functions. We complemented this sensitivity study with the 10-fold CV techniques, as previously

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

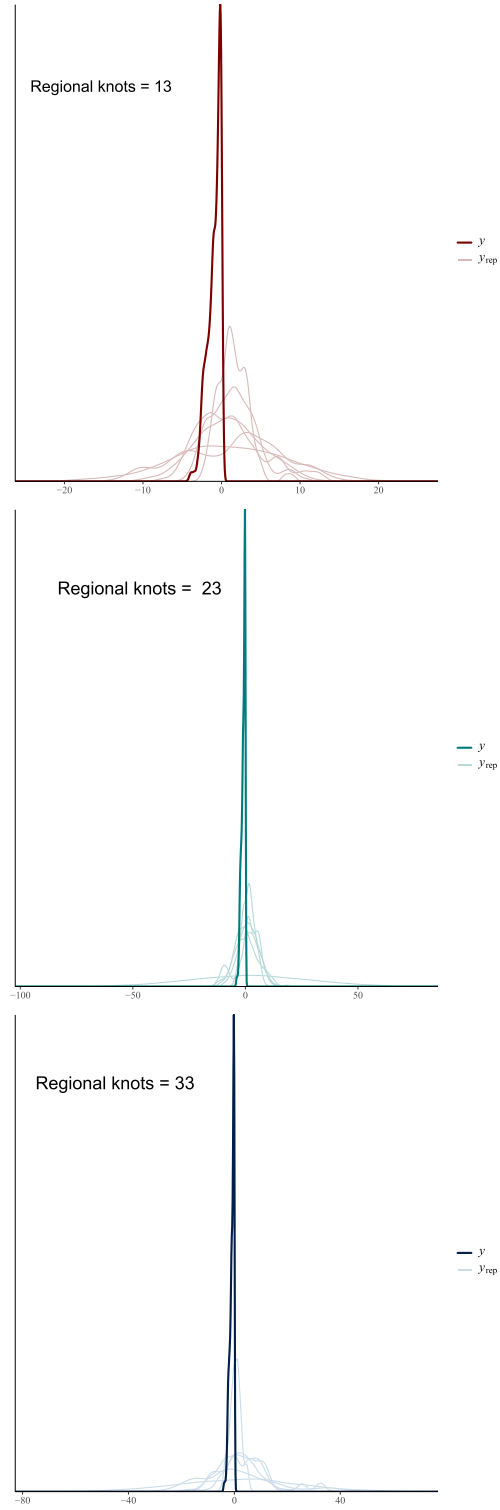

Figure B17. Prior predictive checks for the regional component varying the number of knots ranging from 13 knots (red), 23 knots (green), and 33 (knots). For each plot, the solid colour line denoted by y represents the real data, i.e. RSL. The y_{rep} lines represent draws from the prior predictive distribution.

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

mentioned, to assess the goodness-of-fit for different numbers of knots and determine the optimal number of knots that balance model complexity and efficiency.

Prior Predictive Tests

In this section, we employ prior predictive tests to validate our choice of priors, as recommended by Rubin (1984) and Gelman et al. (1996). Prior predictive checks can be used to assess the suitability of chosen prior distributions by generating data according to those priors and the model structure (Gabry et al., 2019). The observations play no role in the prior predictive checks, other than for comparison purposes. For the prior predictive checks, we use the 'data' block within `JAGS` to simulate data from the likelihood based solely on our prior assumptions (Plummer, 2003). During our prior sensitivity analysis phase, we systematically changed the priors to better comprehend their impact and assess the robustness of the results across various prior specifications and confirmed our decisions using prior predictive checks (Winter & Depaoli, 2023). In Figure B17, we demonstrate the impact of changing the number knots for the regional component on the prior. This plot demonstrates that the priors are conservative, as the data falls within the intervals. If the plot were excessively tight, it would indicate overfitting; however, our models strike a balance, being slightly informed yet sufficiently conservative.

References

Aquino-López M. A., Blaauw M., Christen J. A., & Sanderson N. K. (2018). Bayesian analysis of 210 Pb dating. *Journal of Agricultural, Biological and Environmental Statistics*, *23*(3), 317–333. https://doi.org/10.1007/s13253-018-0328-7

Argus D., Peltier W., Drummond R., & Moore A. (2014, May). The Antarctica component of postglacial rebound model ICE-6GC (VM5a) based on GPS positioning, exposure age dating of ice thicknesses, and relative sea level histories. *Geophysical Journal International*, *198*, 537–563. https://doi.org/10.1093/gji/ggu140

Ashe E. L., Cahill N., Hay C., Khan N. S., Kemp A., Engelhart S. E., Horton B. P., Parnell A. C., & Kopp R. E. (2019). Statistical modeling of rates and trends in Holocene relative sea level. *Quaternary Science Reviews*, *204*, 58–77. https://doi.org/10.1016/j.quascirev.2018.10.032

Baddeley A., Turner R., Møller J., & Hazelton M. (2005). Residual analysis for spatial point processes (with discussion). *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(5), 617–666. https://doi.org/10.1111/j.1467-9868.2005.00519.x

Barnett R., Bernatchez P., Garneau M., Brain M., Charman D., Stephenson D., Haley S., & Sanderson N. (2019). Late Holocene sea-level changes in eastern Quaébec and potential drivers. *Quaternary Science Reviews*, *203*, 151–169. https://doi.org/10.1016/j.quascirev.2018.10.039

Barnett R. L., Bernatchez P., Garneau M., & Juneau M.-N. (2017). Reconstructing late Holocene relative sea-level changes at the Magdalen Islands (Gulf of St. Lawrence, Canada) using multi-proxy analyses. *Journal of Quaternary Science*, *32*(3), 380–395. https://doi.org/10.1002/jqs.v32.3

Berrett C., Christensen W. F., Sain S. R., Sandholtz N., Coats D. W., Tebaldi C., & Lopes H. F. (2020). Modeling sea-level processes on the U.S. Atlantic Coast. *Environmetrics*, *31*(4), e2609. https://doi.org/10.1002/env.v31.4

Blaauw M., & Christen J. A. (2011). Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Analysis*, *6*(3), 457–474. https://doi.org/10.1214/11-BA618

Cahill N., Kemp A. C., Horton B. P., & Parnell A. C. (2015). Modeling sea-level change using errors-in-variables integrated Gaussian process 1. *The Annals of Applied Statistics*, *9*(2), 547–571. https://doi.org/10.1214/15-AOAS824

Cahill N., Kemp A. C., Horton B. P., & Parnell A. C. (2016). A Bayesian hierarchical model for reconstructing relative sea level: From raw data to rates of change. *Climate of the Past*, *12*(2), 525–542. https://doi.org/10.5194/cp-12-525-2016

Caron L., Ivins E. R., Larour E., Adhikari S., Nilsson J., & Blewitt G. (2018). GIA model statistics for GRACE hydrology, cryosphere, and ocean science. *Geophysical Research Letters*, *45*(5), 2203–2212. https://doi.org/10.1002/grl.v45.5

Church J. A., & Clark P. U. (2013). Climate change 2013: The physical science basis. In *Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. IPCC 2013* (Vol. 13, pp. 1150–1186). Cambridge University Press.

Church J. A., Gregory J. M., Huybrechts P., Kuhn M., Lambeck K., Nhuan M. T., Qin D., & Woodworth P. L. (2001). Changes in sea level: Climate change 2001: The scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel.

Church J. A., & White N. J. (2011). Sea-level rise from the late 19th to the early 21st century. *Surveys in Geophysics*, *32*(4), 585–602. https://doi.org/10.1007/s10712-011-9119-1

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Dangendorf S., Marcos M., Wöppelmann G., Conrad C. P., Frederikse T., & Riva R. (2017). Reassessment of 20th century global mean sea level rise. *Proceedings of the National Academy of Sciences*, *114*(23), 5946–5951. https://doi.org/10.1073/pnas.1616007114

de Boor C. (1978). *A practical guide to spline*, Applied mathematical sciences (Vol. 27). Springer. https://doi.org/10.2307/2006241

Dey D. K., Ghosh S. K., & Mallick B. K. (2000). *Generalized linear models: A Bayesian perspective*. CRC Press.

Dierckx P. (1995). *Curve and surface fitting with splines*. Oxford University Press.

Donnelly J. P. (2006). A revised late Holocene sea-level record for Northern Massachusetts, USA. *Journal of Coastal Research*, *22*(5), 1051–1061. https://doi.org/10.2112/04-0207.1

Donnelly J. P., Cleary P., Newby P., & Ettinger R. (2004). Coupling instrumental and geological records of sea-level change: Evidence from southern New England of an increase in the rate of sea-level rise in the late 19th century. *Geophysical Research Letters*, *31*(5), 1–4. https://doi.org/10.1029/2003GL018933

Edwards R., & Wright A. (2015). Chapter 13: Foraminifera. In *Handbook of sea-level research* (pp. 191–217). John Wiley & Sons Ltd. https://doi.org/10.1002/9781118452547.ch13

Eilers P., & Marx B. (1996). Flexible smoothing with B-splines and penalities. *Statistical Science*, *11*, 89–121. https://doi.org/10.1214/ss/1038425655

Eilers P., & Marx B. (2021). *Practical smoothing: The joys of P-splines*. Cambridge University Press. https://books.google.ie/books?id=XpgjEAAAQBAJ.

Engelhart S. E., Horton B. P., Douglas B. C., Peltier W. R., & Törnqvist T. E. (2009, December). Spatial variability of late Holocene and 20th century sea-level rise along the Atlantic coast of the United States. *Geology*, *37*(12), 1115–1118. https://doi.org/10.1130/G30360A.1

Engelhart S. E., Horton B. P., & Kemp A. C. (2011). Holocene sea level changes along the United States' Atlantic coast. *Oceanography*, *24*(2), 70–79. https://doi.org/10.5670/oceanog

Frederikse T., Landerer F., Caron L., Adhikari S., Parkes D., Humphrey V. W., Dangendorf S., Hogarth P., Zanna L., Cheng L., & Wu Y.-H. (2020). The causes of sea-level rise since 1900. *Nature*, *584*(7821), 393–397. https://doi.org/10.1038/s41586-020-2591-3

Gabry J., & Goodrich B. (2017). Rstanarm: Bayesian applied regression modeling via Stan.

Gabry J., Simpson D., Vehtari A., Betancourt M., & Gelman A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *182*(2), 389–402. https://doi.org/10.1111/rssa.12378

Gehrels W. R. (1994). Determining relative sea-level change from salt-marsh foraminifera and plant zones on the coast of Maine, USA. *Journal of Coastal Research*, *10*(4), 990–1009. https://www.jstor.org/stable/4298291

Gehrels W. R., Dangendorf S., Barlow N. L. M., Saher M. H., Long A. J., Woodworth P. L., Piecuch C. G., & Berk K. (2020). A preindustrial sea-level rise hotspot along the Atlantic coast of North America. *Geophysical Research Letters*, *47*(4), 1–9. https://doi.org/10.1029/2019GL085814

Gelman A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, *1*(3), 515–534. https://doi.org/10.1214/06-BA117A

Gelman A., Meng X.-L., & Stern H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, *6*(4), 733–760. https://www.jstor.org/stable/24306036

Gerlach M. J., Engelhart S. E., Kemp A. C., Moyer R. P., Smoak J. M., Bernhardt C. E., & Cahill N. (2017). Reconstructing Common Era relative sea-level change on the Gulf Coast of Florida. *Marine Geology*, *390*, 254–269. https://doi.org/10.1016/j.margeo.2017.07.001

Gregory J. M., Griffies S. M., Hughes C. W., Lowe J. A., Church J. A., Fukimori I., Gomez N., Kopp R. E., Landerer F., Cozannet G. L., Ponte R. M., Stammer D., Tamisiea M. E., & van de Wal R. S. (2019). Concepts and terminology for sea level: Mean, variability and change, both local and global. *Surveys in Geophysics*, *40*(6), 1251–1289. https://doi.org/10.1007/s10712-019-09525-z

Grinsted A. (2015). Projected change–sea level. In *Second assessment of climate change for Baltic Sea basin* (pp. 253–263). Springer. https://doi.org/10.1007/978-3-319-16006-1

Hastie T., & Tibshirani R. (1986). Generalized additive models. *Statistical Science*, *1*(3), 297–318. https://www.jstor.org/stable/2245459

Hastie T. J. (2017). Generalized additive models. In *Statistical models in S* (pp. 249–307). Routledge.

Hawkes A. D., Kemp A. C., Donnelly J. P., Horton B. P., Peltier W. R., Cahill N., Hill D. F., Ashe E., & Alexander C. R. (2016). Relative sea-level change in northeastern Florida (USA) during the last 8.0 ka. *Quaternary Science Reviews*, *142*, 90–101. https://doi.org/10.1016/j.quascirev.2016.04.016

Hay C. C., Lau H. C. P., Gomez N., Austermann J., Powell E., Mitrovica J. X., Latychev K., & Wiens D. A. (2017). Sea level fingerprints in a region of complex earth structure: The case of WAIS. *Journal of Climate*, *30*(6), 1881–1892. https://doi.org/10.1175/JCLI-D-16-0388.1

Hay C. C., Morrow E., Kopp R. E., & Mitrovica J. X. (2015). Probabilistic reanalysis of twentieth-century sea-level rise. *Nature*, *517*(7535), 481–484. https://doi.org/10.1038/nature14093

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Holgate S. J., Matthews A., Woodworth P. L., Rickards L. J., Tamisiea M. E., Bradshaw E., Foden P. R., Gordon K. M., Jevrejeva S., & Pugh J. (2013). New data systems and products at the permanent service for mean sea level. *Journal of Coastal Research*, *29*(3), 493–504. https://doi.org/10.2112/JCOASTRES-D-12-00175.1

Horton B. P., & Edwards R. J. (2006). Quantifying Holocene sea level change using intertidal foraminifera: Lessons from the British Isles. Departmental Papers (EES) (p. 50).

Horton B. P., Kopp R. E., Garner A. J., Hay C. C., Khan N. S., Roy K., & Shaw T. A. (2018). Mapping sea-level change in time, space, and probability. *Annual Review of Environment and Resources*, *43*(1), 481–521. https://doi.org/10.1146/annurev-environ-102017-025826

Jevrejeva S., Moore J. C., Grinsted A., & Woodworth P. L. (2008). Recent global sea level acceleration started over 200 years ago? *Geophysical Research Letters*, *35*(8), 1–4. https://doi.org/10.1029/2008GL033611

Kemp A. C., Bernhardt C. E., Horton B. P., Kopp R. E., Vane C. H., Peltier W. R., Hawkes A. D., Donnelly J. P., Parnell A. C., & Cahill N. (2014). Late Holocene sea- and land-level change on the US southeastern Atlantic coast. *Marine Geology*, *357*, 90–100. https://doi.org/10.1016/j.margeo.2014.07.010

Kemp A. C., Hawkes A. D., Donnelly J. P., Vane C. H., Horton B. P., Hill T. D., Anisfeld S. C., Parnell A. C., & Cahill N. (2015). Relative sea-level change in Connecticut (USA) during the last 2200 yrs. *Earth and Planetary Science Letters*, *428*, 217–229. https://doi.org/10.1016/j.epsl.2015.07.034

Kemp A. C., Hill T. D., Vane C. H., Cahill N., Orton P. M., Talke S. A., Parnell A. C., Sanborn K., & Hartig E. K. (2017). Relative sea-level trends in New York City during the past 1500 years. *The Holocene*, *27*(8), 1169–1186. https://doi.org/10.1177/0959683616683263

Kemp A. C., Horton B. P., Donnelly J. P., Mann M. E., Vermeer M., & Rahmstorf S. (2011). Climate related sea-level variations over the past two millennia. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(27), 11017–11022. https://doi.org/10.1073/pnas.1015619108

Kemp A. C., Horton B. P., Vane C. H., Bernhardt C. E., Corbett D. R., Engelhart S. E., Anisfeld S. C., Parnell A. C., & Cahill N. (2013). Sea-level change during the last 2500 years in New Jersey, USA. *Quaternary Science Reviews*, *81*, 90–104. https://doi.org/10.1016/j.quascirev.2013.09.024

Kemp A. C., & Telford R. J. (2015). Chapter 31: Transfer functions. In *Handbook of sea-level research* (pp. 470–-499). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118452547.ch31

Kemp A. C., Wright A. J., Edwards R. J., Barnett R. L., Brain M. J., Kopp R. E., Cahill N., Horton B. P., Charman D. J., Hawkes A. D., Hill T. D., & van de Plassche O. O. (2018). Relative sea-level change in Newfoundland, Canada during the past 3,000 years. *Quaternary Science Reviews*, *201*, 89–110. https://doi.org/10.1016/j.quascirev.2018.10.012

Khan N. S., Ashe E., Horton B. P., Dutton A., Kopp R. E., Brocard G., Engelhart S. E., Hill D. F., Peltier W., Vane C. H., & Scatena F. N. (2017). Drivers of Holocene sea-level change in the Caribbean. *Quaternary Science Reviews*, *155*, 13–36. https://doi.org/10.1016/j.quascirev.2016.08.032

Khan N. S., Ashe E., Moyer R. P., Kemp A. C., Engelhart S. E., Brain M. J., Toth L. T., Chappel A., Christie M., Kopp R. E., & Horton B. P. (2022). Relative sea-level change in South Florida during the past 5000 years. *Global and Planetary Change*, *216*, 1–16. https://doi.org/10.1016/j.gloplacha.2022.103902

Kopp R., Horton B., Kemp A., & Tebaldi C. (2015, October). Past and future sea-level rise along the coast of North Carolina, USA. *Climatic Change*, *132*(4), 693–707. https://doi.org/10.1007/s10584-015-1451-x

Kopp R. E., Kemp A. C., Bittermann K., Horton B. P., Donnelly J. P., Gehrels W. R., Hay C. C., Mitrovica J. X., Morrow E. D., & Rahmstorf S. (2016). Temperature-driven global sea-level variability in the Common Era. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(11), E1434–E1441. https://doi.org/10.1073/pnas.1517056113

Marshall W. (2015). Chapter 25: Chronohorizons. In *Handbook of sea-level research* (pp. 373–385). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118452547.ch25

Masson-Delmotte V., Zhai P., Pirani A., Connors S. L., Péan C., Berger S., Caud N., Chen Y., Goldfarb L., Gomis M. I., Huang M., Leitzell K., Lonnoy E., Matthews J. R., Maycock T. K., Waterfield T., Yelekçi O., Yu R., & Zhou B. (2021). Summary for policymakers. In *Climate change 2021: The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

McHutchon A., & Rasmussen C. E. (2011). Gaussian process training with input noise. In *Advances in neural information processing systems* (Vol. *24*, pp. 1–9). Curran Associates, Inc.

Meltzner A. J., Switzer A. D., Horton B. P., Ashe E., Qiu Q., Hill D. F., Bradley S. L., Kopp R. E., Hill E. M., Majewski J. M., Natawidjaja D. H., & Suwargadi B. W. (2017). Half-metre sea-level fluctuations on centennial timescales from mid-Holocene corals of Southeast Asia. *Nature Communications*, *8*(1), 14387. https://doi.org/10.1038/ncomms14387

Neukom R., Barboza L. A., Erb M. P., Shi F., Emile-Geay J., Evans M. N., Franke J., Kaufman D. S., Lücke L., Rehfeld K., Schurer A., Zhu F., Brönnimann S., Hakim G. J., Henley B. J., Ljungqvist F. C., McKay N., Valler V., von Gunten L., & PAGES 2k Consortium (2019). Consistent multidecadal variability in global temperature reconstructions and simulations over the common era. *Nature Geoscience*, *12*(8), 643–649. https://doi.org/10.1038/s41561-019-0400-0

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Parnell A. C., & Gehrels W. R. (2015). Chapter 32: Using chronological models in late Holocene sea-level reconstructions from saltmarsh sediments. In *Handbook of sea-level research* (Chapter 32, pp. 500–513). John Wiley & Sons, Ltd.

Parnell A. C., Haslett J., Allen J. R., Buck C. E., & Huntley B. (2008). A flexible approach to assessing synchroneity of past events using Bayesian reconstructions of sedimentation history. *Quaternary Science Reviews*, *27*(19–20), 1872–1885. https://doi.org/10.1016/j.quascirev.2008.07.009

Peltier W. (2004). Global glacial isostasy and the surface of the ice-age Earth: The ICE-5G (VM2) model and GRACE. *Annual Review of Earth and Planetary Sciences*, *32*, 111–149. https://doi.org/10.1146/earth.2004.32.issue-1

Peltier W. R., Argus D. F., & Drummond R. (2015). Space geodesy constrains ice age terminal deglaciation: The global ICE-6G_C (VM5a) model. *Journal of Geophysical Research: Solid Earth*, *120*(1), 450–487. https://doi.org/10.1002/jgrb.v120.1

Piecuch C. G., Huybers P., & Tingley M. P. (2017). Comparison of full and empirical Bayes approaches for inferring sea-level changes from tide-gauge data. *Journal of Geophysical Research: Oceans*, *122*(3), 2243–2258. https://doi.org/10.1002/jgrc.v122.3

Plummer M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing, TUWien*, (pp. 1–10). ISSN 1609-395X.

Plummer M. (2015). Cuts in Bayesian graphical models. *Statistics and Computing*, *25*(1), 37–43. https://doi.org/10.1007/s11222-014-9503-z

Plummer M., Best N., Cowles K., & Vines K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, *6*(1), 7–11. ISSN 1609-3631.

Plummer M., Stukalov A., & Denwood M. (2016). *Rjags: Bayesian graphical models using MCMC*. R Package Version 4(6).

Porcu E., Furrer R., & Nychka D. (2021). 30 Years of space-time covariance functions. *WIREs Computational Statistics*, *13*(2), 1512. https://doi.org/10.1002/wics.v13.2

Pugh D., & Woodworth P. (2014a). Mean sea-level changes in time. In *Sea-level science: Understanding tides, surges, tsunamis and mean sea-level changes* (pp. 252–295). Cambridge University Press. https://doi.org/10.1017/CBO9781139235778.013

Pugh D., & Woodworth P. (2014b). Tidal forces. In *Sea-level science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes* (pp. 36–59). Cambridge University Press. https://doi.org/10.1017/CBO9781139235778.006

Redfield A. C. (1972). Development of a New England salt marsh. *Ecological Monographs*, *42*(2), 201–237. https://doi.org/10.2307/1942263

Roy K., & Peltier W. (2015). Glacial isostatic adjustment, relative sea level history and mantle viscosity: Reconciling relative sea level model predictions for the US East coast with geological constraints. *Geophysical Journal International*, *201*(2), 1156–1181. https://doi.org/10.1093/gji/ggv066

Rubin D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics*, *12*, 1151–1172. https://doi.org/10.1214/aos/1176346785

Sachs H. M., Webb III T., & Clark D. (1977). Paleoecological transfer functions. *Annual Review of Earth and Planetary Sciences*, *5*(1), 159–178. https://doi.org/10.1146/earth.1977.5.issue-1

Scholl D. W., & Stuiver M. (1967). Recent submergence of southern Florida: A comparison with adjacent coasts and other eustatic data. *Geological Society of America Bulletin*, *78*(4), 437–454. https://doi.org/10.1130/0016-7606(1967)78[437:RSOSFA]2.0.CO;2

Shennan I., Bradley S. L., & Edwards R. (2018). Relative sea-level changes and crustal movements in Britain and Ireland since the Last Glacial Maximum. *Quaternary Science Reviews*, *188*, 143–159. https://doi.org/10.0277–3791

Shennan I., Long A. J., & Horton B. P. (2015). *Handbook of sea-level research*. John Wiley & Sons, Ltd.

Spiegelhalter D. J., Best N. G., Carlin B. P., & Van Der Linde A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583–639. https://doi.org/10.1111/1467-9868.00353

Stammer D., Cazenave A., Ponte R. M., & Tamisiea M. E. (2013). Causes for contemporary regional sea level changes. *Annual Review of Marine Science*, *5*(1), 21–46. https://doi.org/10.1146/marine.2013.5.issue-1

Stearns R. B., & Engelhart S. E. (2017). A high-resolution reconstruction of late-holocene relative sea level in Rhode Island, USA.

Törnqvist T. E., Rosenheim B. E., Hu P., & Fernandez A. B. (2015). Chapter 23: Radiocarbon dating and calibration. In *Handbook of sea-level research* (pp. 347–360). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118452547.ch23

Upton M., Parnell A., & Cahill N. (2023). Reslr: An R package for relative sea level modelling. *Rjournal*. https://doi.org/10.32614/CRAN.package.reslr

Downloaded from https://academic.oup.com/jrssc/advance-article/doi/10.1093/jrssc/qlae044/7756353 by Maynooth University user on 23 September 2024

Vacchi M., Engelhart S. E., Nikitina D., Ashe E. L., Peltier W. R., Roy K., Kopp R. E., & Horton B. P. (2018). Postglacial relative sea-level histories along the eastern Canadian coastline. *Quaternary Science Reviews*, *201*, 124–146. https://doi.org/10.1016/j.quascirev.2018.09.043

Walker J. S., Kopp R. E., Little C. M., & Horton B. P. (2022). Timing of emergence of modern rates of sea-level rise by 1863. *Nature Communications*, *13*(1), 1–8. https://doi.org/10.1038/s41467-022-28564-6

Walker J. S., Kopp R. E., Shaw T. A., Cahill N., Khan N. S., Barber D. C., Ashe E. L., Brain M. J., Clear J. L., Corbett D. R., & Horton B. P. (2021). Common Era sea-level budgets along the US Atlantic coast. *Nature Communications*, *12*(1), 1841. https://doi.org/10.1038/s41467-021-22079-2

Wenzel M., & Schröter J. (2010). Reconstruction of regional mean sea level anomalies from tide gauges using neural networks. *Journal of Geophysical Research: Oceans*, *115*(C8), 1–15. https://doi.org/10.1029/2009JC005630

Whitehouse P. L. (2018). Glacial isostatic adjustment modelling: Historical perspectives, recent advances, and future directions. *Earth Surface Dynamics*, *6*, 401–429. https://doi.org/10.5194/esurf-6-401-2018

Winter S. D., & Depaoli S. (2023). Illustrating the value of prior predictive checking for Bayesian structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *30*(6), 1000–1021. https://doi.org/10.1080/10705511.2022.2164286

Wood S. (2015). *Package 'mgcv'*. R Package Version 1(29), 729.

Wood S. N. (2006). Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics*, *62*(4), 1025–1036. https://doi.org/10.1111/biom.2006.62.issue-4

Wood S. N. (2016). Just another Gibbs additive modeller: Interfacing JAGS and mgcv. *Journal of Statistical Software*, *75*(7), 1-15. https://doi.org/10.18637/jss.v075.i07

Wood S. N. (2017). *Generalized additive models: An introduction with R*. CRC Press.

Woodworth P., & Player R. (2003). The permanent service for mean sea level: An update to the 21st century. *Journal of Coastal Research*, *19*, 287–295. https://www.jstor.org/stable/4299170

Wöppelmann G., Zerbini S., & Marcos M. (2006). Tide gauges and Geodesy: A secular synergy illustrated by three present-day case studies. *Comptes Rendus Geoscience*, *338*(14), 980–991. La Terre observée depuis l'espace. https://doi.org/10.1016/j.crte.2006.07.006

Wright A. J., Edwards R. J., van de Plassche O., Blaauw M., Parnell A. C., van der Borg K., de Jong A. F., Roe H. M., Selby K., & Black S. (2017). Reconstructing the accumulation history of a saltmarsh sediment core: Which age-depth model is best? *Quaternary Geochronology*, *39*, 35–67. https://doi.org/10.1016/j.quageo.2017.02.004

Downloaded from https://academic.oup.com/jrsssc/advance-article/doi/10.1093/jrsssc/qlae044/7756353 by Maynooth University user on 23 September 2024