reslr: An R Package for Relative Sea Level Modelling

by Maeve Upton, Andrew Parnell, and Niamh Cahill

Abstract We present reslr, an R package to perform Bayesian modelling of relative sea level data. We include a variety of different statistical models previously proposed in the literature, with a unifying framework for loading data, fitting models, and summarising the results. Relative sea-level data often contain measurement error in multiple dimensions, and so our package allows for these to be included in the statistical models. When plotting the output sea level curves, the focus is often on comparing rates of change, and so our package allows for computation of the derivatives of sea level curves with appropriate consideration of the uncertainty. We provide a large example dataset from the Atlantic coast of North America and show some of the results that might be obtained from our package.

1 Introduction

Understanding the rates and spatial patterns of Relative Sea-Level (RSL) change across various timescales, spanning from decades to millennia, poses a significant challenge. The task involves analysing sparse and noisy proxy and/or instrumental data sources that often have large measurement uncertainties. To address these complexities and provide robust assessments, statistical models play a pivotal role and have become indispensable in the task of quantifying RSL changes (as examined by Cahill et al. (2015a) and Khan et al. (2015)) and in the evaluation of temporal and spatial variability (e.g., Kopp, 2013; Kopp et al., 2016; Kemp et al., 2018; Walker et al., 2021). To that end, the paleo sea-level community would benefit from a comprehensive toolset capable of analysing the historical evolution of sea-level changes across different times and locations. This motivated us to create the **reslr** package, which is available on the Comprehensive R Archive Network at `https://cran.r-project.org/web/packages/reslr` or on GitHub at `https://github.com/maeveupton/reslr`. Our package includes a suite of statistical models appropriate for modelling the complexity of sea-level over time and space, while accounting for sea-level data uncertainties and remaining computationally tractable. The output of our package provides insight into temporal and spatial sea-level variability and rates of sea-level change.

The **reslr** package includes a comprehensive dataset of proxy RSL reconstructions for 21 locations along the Atlantic coast of North America (Kemp et al., 2013). These reconstructions rely heavily on dated geological archives obtained from coastal sediments (e.g., Gehrels, 1994) or corals (e.g., Meltzner et al., 2017). Moreover, users of the package have the option to incorporate instrumental sea-level data sourced from the Permanent Service Mean Sea Level (PSMSL) online database, which provides annual RSL measurements for approximately 1,500 tide-gauge stations worldwide (Holgate et al., 2013). By offering this diverse range of data sources, the **reslr** package caters to the needs of researchers seeking to explore and analyse sea-level variations across different locations and time periods.

The **reslr** package offers a range of statistical models which include: linear regression (e.g., Ashe et al. (2019)), change point models (e.g., Cahill et al. (2015b)), integrated Gaussian process (IGP) models (e.g., Cahill et al. (2015a)), temporal splines (e.g., de Boor (1978)), spatio-temporal splines (e.g., Simpson (2018)) and generalised additive models (GAM) (e.g., Upton et al. (2023)). In all cases, a Bayesian framework is employed, facilitating the estimation of unknown parameters based on the RSL data while fully accounting for the associated uncertainties. The **reslr** package enables researchers to gain comprehensive insights into sea-level variations, leveraging the flexibility and robustness of these statistical models.

When it comes to addressing measurement uncertainty in proxy records, the **reslr** package offers two distinct approaches. The first approach involves employing the Errors-in-Variables (EIV) method, which takes into account the inherent uncertainties in the input variables (Dey et al., 2000). This method acknowledges that the input variables are not error-free and incorporates this knowledge into the analysis. The second approach offered by the package is the Noisy Input (NI) uncertainty method. This method tackles uncertainty by inflating the output noise variance with a corrective term that is directly linked to the input noise variance (McHutchon and Rasmussen, 2011). Both the EIV and NI uncertainty methods have their respective advantages, and the **reslr** package recommends the most suitable uncertainty method based on the statistical model being employed. This ensures that researchers can select the appropriate approach to effectively address measurement uncertainties within their specific analysis context.

For each model, the **reslr** package generates informative plots illustrating the model-based estimates of RSL. In the case of more complex models, like the IGP, splines, and GAMs, the resulting

plots not only provide RSL estimates but also offer insights into the rates of RSL change. Of particular significance to the paleo-sea level community, the GAM model provides estimates for separate components that represent potential drivers of RSL change. This feature enables comparisons between different components and contributes to a more comprehensive understanding of the factors influencing RSL fluctuations (Upton et al., 2023). These visual representations enable researchers to gain a clearer understanding of when and where RSL changes occurred, including the magnitude of their temporal variations. Moreover, the package grants users access to the posterior samples used to generate the plots, providing the option to delve deeper into the underlying statistical distributions and uncertainties associated with the estimated RSL changes. The combination of these outputs serves as a valuable resource for researchers, aiding in the investigation and interpretation of RSL dynamics across various spatial and temporal contexts.

The **reslr** package is uniquely tailored to address the challenges inherent in analysing historic RSL changes using proxy records. Its design, characterised by minimal functions and a user-friendly interface, draws inspiration from other packages such as **mgcv**, renowned for its diverse statistical modelling options (Wood, 2015). Crucially, there are currently no competing R packages that match the breadth of capabilities offered by **reslr**. This uniqueness stems from its capability to provide users with a selection of Bayesian statistical models and account for bi-variate uncertainty, crucial in sea-level research.

Our paper has the following structure. First, we introduce the example dataset provided within the package, which serves as the foundation for the examples presented throughout the paper. We also provide insight into additional data sources. Second, we offer an overview of the statistical models available in the package, providing necessary background information. Next, we explore the uncertainty methods employed within these statistical models. Following this, we provide a detailed description of the functionality of the **reslr** package, outlining the diverse outputs and plots accessible to users. Finally, we conclude with important remarks and discuss potential future extensions for the package's advancement. Whilst this paper is just a summary of the features of **reslr**, a more complete vignette containing examples of the full functionality of the package is available at `https://maeveupton.github.io/reslr/`.

2 Data and models

2.1 Data sources

Proxy sea-level data are vital sources of information for examining historic changes in RSL prior to the instrumental data period. A proxy refers to a characteristic that can be observed and used to estimate a variable of interest, which cannot be measured directly, and can be of physical, biological, or chemical nature (e.g., Gornitz, 2009). In sea-level studies, the proxy data can be sourced from microorganisms such as foraminifera (e.g., Edwards and Wright, 2015), geochemical measurements (e.g., Marshall, 2015), or vegetation that has accumulated in the tidal realm (e.g., Kemp and Telford, 2015). The datasets we use have had their proxy measurements transformed into sea level using various techniques which are beyond the scope of our paper (e.g., Gehrels (1994), Shennan et al. (2015), Kemp et al. (2018)). In the **reslr** package, we provide an example proxy dataset which contains 21 proxy sea-level records (See Appendix) from the Atlantic coast of North America as used in Upton et al. (2023).

Within the context of sea-level analysis, instrumental data plays an important role by providing direct measurements obtained from tide gauges and satellites (although the latter is currently not incorporated into the **reslr** package). To enhance the versatility of the package, we have implemented a feature that allows users to download annual tide-gauge data from the PSMSL Level online database and store it in a temporary file, making it readily available when needed (PSMSL, 2023; Holgate et al. (2013); Woodworth and Player (2003)).

To ensure the comparability of the tide-gauge data with proxy records, we apply two processing steps. First, the tide-gauge data in the PSMSL database is given in millimetres relative to a revised local reference datum (a coordinate system that defines the zero level for sea level measurements Pugh and Woodworth (2014)). Within **reslr**, we transform the data by removing 7,000 mm to revert the tide-gauge data into the observed reference frame and convert the RSL to metres following the guidance from the PSMSL website as described in Aarup et al. (2006). The second processing step involves averaging the tide-gauge data to align with the temporal resolution of the more recent proxy data. However, we provide flexibility for users to adjust this averaging period according to the specific characteristics of their data.

2.2 Statistical Models

Within the **reslr** package, a Bayesian hierarchical framework is employed for each statistical modelling technique. Markov Chain Monte Carlo (MCMC) simulations are carried out using the Just Another Gibbs Sampler (JAGS) tool (Plummer, 2003) and implemented using the `rjags` package in R (Plummer et al., 2016). Other tools, such as Stan (Carpenter et al., 2017), are used for MCMC simulations. However, our preference for JAGS is based on its user-friendly interface, computational efficiency, and flexible model design capabilities.

Mathematically, the data level for each statistical model is described as:

$$y = f(\mathbf{x}, t) + \epsilon_y \tag{1}$$

where y is the response data (RSL in metres). $f(\mathbf{x}, t)$ is the process mean that depends on location $\mathbf{x}$ and time t. ϵ_y is the error term given by $\epsilon_y \sim \mathbb{N}(0, \sigma_y^2 + s_y^2)$, where σ_y^2 is the residual variance and s_y the known measurement error associated with RSL, which occurs during the collection of proxy records (e.g., Kemp and Telford (2015)). In Table 1, we provide a list of all the possible options for f within the **reslr** package. Since some of the models we fit do not vary over space (they apply to a single site or treat a set of sites as identical), we use $f(t)$ rather than $f(\mathbf{x}, t)$ to denote the process model. In Figure 1, we provide a simplified graphic to describe the Bayesian hierarchical framework used for each statistical model in Table 1.

When using proxy RSL data, measurement error is also present in the input variable (time) due to the dating technique used. For the input measurements, $\tilde{t}$ is assumed to be a noisy estimate of the true time value t:

$$\tilde{t} = t + \epsilon_t \tag{2}$$

with the error term given by $\epsilon_t \sim \mathbb{N}(0, s_t^2)$ where s_t is the known measurement error associated with time.

We use two methods to account for the time measurement uncertainty. The first is the Errors-in-variables (EIV) method, which assumes that the input variable, e.g., time, is measured as an error-prone substitute and models it directly (Dey et al., 2000). The second uncertainty method is the Noisy Input (NI) method. This method fits an initial model and uses the derivative of the mean of f to calculate a corrective variance term. Then, the model is re-run with this additional corrective variance term, allowing for the input noise variation to be learned from the complete outputs of the model (McHutchon and Rasmussen, 2011). The **reslr** package employs the EIV method for linear regression, change point, and IGP models, while employing the NI uncertainty method for temporal spline, spatio-temporal spline, and GAM models. In general, the EIV method tends to be slower but models the uncertain input process directly, whilst the NI method is faster but requires the model to be fitted twice.

In Table 1, we present a range of statistical modelling techniques for $f(\mathbf{x}, t)$, the component of our approach, available in the **reslr** package. Below we discuss each technique and provide insight into the potential uses of these techniques for the paleo-environmental community.

Statistical Model	**Model Information**	**`model_type` code**
Errors in variables simple linear regression	A straight line of best fit taking into account any age and measurement errors in the RSL values using the method of Cahill et al. (2015a)	`"eiv_slr_t"`
Errors in variables change point model	An extension of the linear regression modelling process. It uses piece-wise linear sections and estimates where/when trend changes occur in the data (Cahill et al., 2015a)	`"eiv_cp_t"`
Errors in variables integrated Gaussian process	A non-linear fit that uses a Gaussian process prior on the rate of sea-level change that is then integrated (Cahill et al., 2015a).	`"eiv_igp_t"`
Noisy Input spline in time	A non-linear fit using regression splines (Upton et al., 2023).	`"ni_spline_t"`
Noisy Input spline in space and time	A non-linear fit for a set of sites across a region using the method of Upton et al. (2023).	`"ni_spline_st"`
Noisy Input Generalised Additive model for the decomposition of the RSL signal	A non-linear fit for a set of sites across a region and provides a decomposition of the signal into regional, local linear, and non-linear local components. This full model is as described in Upton et al. (2023).	`"ni_gam_decomp"`

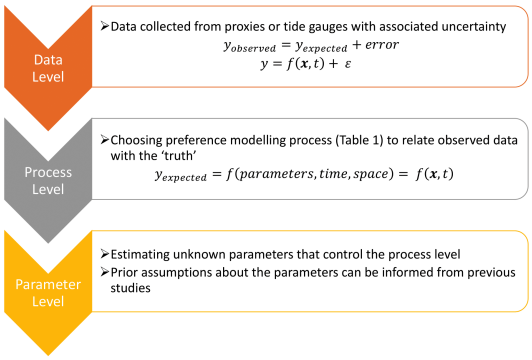

Figure 1: A visual representation of the Bayesian hierarchical model structure used for each statistical model in the reslr package.

Table 1: List of all statistical models available in the **reslr** package. We provide a short description and the relevant literature for each model. The `model_type` code column represents the text input the user should use when implementing their preferred modelling technique.

2.3 EIV Linear Regression

The EIV linear regression model is given by:

$$f(t) = \alpha + \beta t \tag{3}$$

where α is the intercept, β is the slope, and t is time. Earlier studies, for example, Shennan and Horton (2002) and Engelhart et al. (2009), employed linear regression when evaluating the rate of RSL change over the past 4,000 years. The **reslr** package implements a temporal linear regression as its simplicity is popular for approximate estimates of linear rates of RSL change. However, linearity assumptions for RSL change are often unrealistic when examining long-term historical trends.

2.4 EIV Change Point Model

The EIV change point (CP) model, an extension of the linear regression model, assumes the RSL process is piecewise linear and estimates when trend changes occur in the data (Cahill et al., 2015b). Mathematically, the multiple CP model, $f(t)$ is described as:

$$f(t) = \begin{cases} \alpha_1 + \beta_j(t - \lambda_1) \text{ when } j = 1, 2, \\ \alpha_{j-1} + \beta_j(t - \lambda_{j-1}), \text{ when } j = 3, ..., m+1 \end{cases} \tag{4}$$

where α_j is the expected value of the response at the jth CP. λ_j is the time at which the CP occurs with the prior restriction that $\lambda_1 < \lambda_2 < ... < \lambda_m$ and m is the number of CPs (Cahill et al., 2015b). In the `reslr` package, the user can select m to be 1, 2, or 3 CPs. β_1 and β_{m+1} are the slopes before and after the first and last CP, respectively. β_j for $j = 2 \ldots, m$ are the slopes between the $(j-1)$th and jth CP.

This technique has been used in different aspects of the sea-level literature. For example, Kemp et al. (2009) determined the magnitude and the timing of recent accelerated sea-level rise using change point models in North Carolina, USA. Brain et al. (2012) used the CP method to examine the impacts of sediment compaction on reconstructing recent sea-level rise in the United Kingdom. Hogarth et al. (2020) used CP models to obtain more consistent estimates of sea-level rise since 1958 for the British Isles. The main advantage of the CP model is its ability to identify sudden changes in RSL. However, the number of change points must be specified by the user.

2.5 Integrated Gaussian Process

An Integrated Gaussian process (IGP) is a modelling strategy that is extensively used by the sea-level community when examining the temporal evolution of sea level change (e.g. Cahill et al. (2015a); Hawkes et al. (2016); Kemp et al. (2017); Shaw et al. (2018); Dean et al. (2019); Stearns et al. (2023); Kirby et al. (2023)).

The IGP uses Gaussian Process (GP) to directly estimate the rate of change of the response

(Holsclaw et al., 2013). In order to extract the original $f(t)$ we integrate $p(t)$:

$$f(t) = \alpha + \int_0^t p(u)du \tag{5}$$

where α is the intercept, and is the rate of change, $p(t) = \frac{df}{dt}$, described as:

$$p(t) \sim GP(\mu(t), k(t, t')) \tag{6}$$

with t time and $\mu(t)$ the mean function, and $k(t, t')$ is the covariance function. The covariance function provides insight into the relationship between the outcome variables, i.e., if the input variables, t and t', are in close proximity, the corresponding outcomes will be more correlated, and vice versa (Rasmussen and Williams, 2006). It is written as (Cahill et al., 2015a):

$$k(t, t') = \nu^2 \rho^{(t-t')^2} \tag{7}$$

where ρ is the correlation parameter and ν^2 is the variance of the rate process.

The technique, described by Cahill et al. (2015a), offers insights into examining rates of Relative Sea Level (RSL) change using proxy records from a single location. Apart from the IGP model, the **reslr** package does not rely on GP methods. We acknowledge that the use of GP modelling has gained considerable traction within the sea-level research community, particularly for investigating the spatio-temporal evolution of sea-level changes, as evidenced by notable studies (Kopp et al., 2009; Kopp, 2013; Kopp et al., 2016; Kemp et al., 2018; Walker et al., 2021). Nevertheless, in the context of the **reslr** package, we have intentionally opted for computationally efficient alternatives—splines and GAMs—as detailed in our prior work (Upton et al., 2023). These methods offer practical and effective approaches to analysing sea-level data, accommodating the complexities of spatio-temporal dynamics while ensuring computational tractability.

2.6 Temporal Spline

Splines are mathematical tools used in a wide range of settings from interpolation to data smoothing. There are a variety of different splines available, yet in this research we focus on B-splines (de Boor, 1978; Dierckx, 1995) and P-splines (Eilers and Marx, 1996). Mathematically, B-splines are described in the following way:

$$f(t) = \sum_{k=1}^{K} b_k(t)\beta_k \tag{8}$$

where $b_k(t)$ is the spline basis function and β_k is the spline coefficient.

Following on from B-splines, Eilers and Marx (1996) describe a method to overcome the difficulty of choosing the correct number of knots by developing penalised spline or P-splines. Penalised differences in the spline coefficients control the smoothness of the spline based on differences (of order d) of the spline coefficients. The first order differences are written as:

$$\Delta\beta_k = \beta_k - \beta_{k-1} \tag{9}$$

The spline coefficient will be centered on the previous value with an inverse smoothness parameter σ_β^2:

$$\Delta\beta_k \sim \mathcal{N}(0, \sigma_\beta^2) \tag{10}$$

In our package, P-splines are used for the NI spline in time and the extendable nature of these splines allows for different components to be examined within the GAM, which is described below.

2.7 Spatio-Temporal Spline

We use a spatio-temporal spline to examine RSL evolving over time at multiple locations. We include a tensor product to capture the variability over time and space (represented with longitude and latitude). For each individual covariate, time (t) and longitude (x_1) and latitude (x_2), we construct a B-spline basis (Wood, 2017a). These basis functions are combined product-wise in the following way (Wood, 2006):

$$f(t, x_1, x_2) = \sum_{h=1}^{H}\sum_{i=1}^{I}\sum_{j=1}^{J} b_h(t)b_i(x_1)b_j(x_2)\beta_{hij} \tag{11}$$

where β_{hij} is the spline coefficient. H is the number of knots for $b_h(t)$ the spline basis function in time t. I is the number of knots for $b_i(x_1)$ the spline basis function for longitude. J is the number of knots for $b_j(x_2)$ the spline basis functions for latitude values. The prior for the spline coefficient is given as:

$$\beta_{hij} \sim \mathcal{N}(0, \sigma_\beta^2) \tag{12}$$

where σ_β^2 is the smoothness parameter for the spatio-temporal spline. The **reslr** package uses B-splines for the NI spline in space and time, allowing for multiple sites to be examined. The advantage of the tensor B-spline approach is that the basis functions are simple to construct, each depending on only one input variable. However, the number of parameters to estimate does increase considerably.

2.8 Generalised Additive Models

Generalised additive models are an extension of generalised linear models that use a basis expansion and a smoothing penalty to create linear predictors that are dependent on the sum of smooth functions of the predictor variable (GAMs; (Wood, 2017b)). The model developed by Upton et al. (2023) uses splines and random effects to create a spatio-temporal relative sea level surface. It identifies variations of sea-level at different spatial and temporal scales, encompassing multiple underlying processes and avoiding a focus on specific physical processes. The decomposition of this mean relative sea level surface can be written as:

$$f(\mathbf{x}, t) = r(t) + g(z_\mathbf{x}) + h(z_\mathbf{x}) + l(\mathbf{x}, t) \tag{13}$$

where $r(t)$ is the regional component at time t represented with a spline in time. $g(z_\mathbf{x})$ is the linear local component at location x represented by a random effect with $z_\mathbf{x}$ representing each data site. $h(z_\mathbf{x})$ is the spatial vertical offset for each data site. $l(\mathbf{x}, t)$ is the non-linear local component represented with a spline in space-time.

The regional component ($r(t)$) represents temporal processes that are common to all locations, including barystatic and thermosteric contributions, where the former is caused by the transfer of mass between land-based ice and oceans (Gregory et al., 2019) and the latter is influenced by changes in global temperature creating density variations within the oceans (Grinsted, 2015). It is described using a spline in time:

$$r(t) = \sum_{s=1}^{k_r} b_{r_s}(t)\beta_s^r \tag{14}$$

where β_s^r is the s^{th} spline coefficient, k_r is the number of knots, and $b_{r_s}(t)$ is the s^{th} spline basis function at time t. The prior for the spline coefficients of the regional component β_s^r are:

$$\beta_s^r \sim \mathcal{N}(0, \sigma_r^2) \tag{15}$$

where the smoothness of the model fit is controlled by σ_r, the standard deviation of the spline coefficient.

The linear local component ($g(z_x)$) of the sea level model aims to capture linear trends present in the relative sea level signal. One such cause is glacial isostatic adjustment (GIA), which is a response of the Earth, the gravitational field, and the ocean to changes in the size of ice sheets (Whitehouse, 2018). On relatively short timescales, it is approximated to be linear through time with spatial variability along the Atlantic coastline of North America (Engelhart et al., 2009). Mathematically, we define our linear local component as an unstructured random effect for each site which is formulated as:

$$g(z_{\mathbf{x}_j}) = \beta_j^g t \tag{16}$$

where β_j^g is a slope parameter specific for each site j. The prior for the linear local component is given by:

$$\beta_j^g \sim \mathcal{N}(m_{g_j}, s_{g_j}^2) \tag{17}$$

where m_{g_j} and $s_{g_j}^2$ are the empirically estimated rate and associated variance for the dataset (refer to Upton et al., 2023, for a detailed description).

The site-specific vertical offset h is a random effect used to capture vertical shifts associated with measurement variability between sites and is formulated as:

$$h(z_{\mathbf{x}_j}) = \beta_j^h \tag{18}$$

where β_j^h contains the random effect coefficients for site j. The prior for the site-specific vertical offset β_j^h is given as:

$$\beta_j^h \sim \mathcal{N}(0, \sigma_h^2) \tag{19}$$

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

The non-linear local component ($l(\mathbf{x}, t)$) captures structured and unstructured RSL variability on century timescales, including dynamic sea-level changes (atmospheric and oceanic circulation patterns (Gregory et al., 2019)) and site-specific processes (e.g., sediment compaction affecting solid Earth's surface (Horton et al., 2018)). It is described using a spatio-temporal spline function formed using a tensor product and is formulated as:

$$l(\mathbf{x}, t) = \sum_{s=1}^{k_l} b_{l_s}(\mathbf{x}, t)\beta_s^l \tag{20}$$

where β_s^l is the s^{th} spline coefficient, k_l is the number of knots, and $b_{l_s}(\mathbf{x}, t)$ is the s^{th} spline basis function at time t and location $\mathbf{x}$. The prior for the spline coefficient β_s^l is given as:

$$\beta_s^l \sim \mathcal{N}(0, \sigma_l^2) \tag{21}$$

where σ_l^2 is the variance of the spline coefficients over space and time.

As described in Upton et al. (2023), B-splines are used for both the regional and local terms as this model structure balances both model usability and computational efficiency for examining proxy-based sea level reconstructions on a regional to local scale. B-splines also allow for easier prior elicitation of the smoothness parameters since they directly control the variability of the spline weights in the model.

3 Implementation

Within the package, we keep the number of functions to a minimum to ensure accessibility for users with varying experience in R. We run the statistical models using an MCMC algorithm and include a summary function to obtain a high-level insight into the outputs. We use S3 classes to access the summary, print, and plot commands. The package has functions to plot the input data and resulting model fits using `ggplot2` (Wickham, 2016). The user has access to all the underlying information used to create these plots, allowing these visualisations to be re-created. In addition, the functions within the package are extendable, allowing advanced users access to more complex outputs. A detailed description for each function and associated commands for the **reslr** package can be found in the vignettes.

In this section, we provide insight into the example dataset and additional data sources using tide-gauge data within the **reslr** package. A discussion is provided into each function using two separate case studies: a single location and multiple locations. In the first case study, we demonstrate the Noisy Input temporal spline (`model_type = "ni_spline_t"`) which is an example modelling strategy for a single location. In the second case study, we examine multiple locations using the Noisy Input GAM decomposition (`model_type = "ni_gam_decomp"`).

3.1 Example proxy dataset

We include an example dataset called `NAACproxydata`. The full dataset with the names of the locations and associated literature is in the Appendix. The `NAACproxydata` is a data frame with 1715 rows and 8 columns which include:

- *Region*: Region name
- *Site*: Site name
- *Latitude*: Latitude of the site
- *Longitude*: Longitude of the site
- *RSL*: Relative Sea level in metres
- *RSL_err*: 1 standard deviation error associated with relative sea level measured in metres
- *Age*: Age in years Common Era (CE)
- *Age_err*: 1 standard deviation error associated with the age in years CE

3.2 Including tide-gauge data

The tide-gauge data available to users can be obtained from the PSMSL online database ((Holgate et al., 2013; Woodworth and Player, 2003)) through the **reslr** package. To ensure compatibility with the proxy records, several processing steps are performed within the package, as discussed earlier.

When incorporating tide-gauge data, users have three methods to select their preferred tide gauge(s). The first option is to provide a list of tide-gauge names from the PSMSL database, allowing users the freedom to select any tide gauge available. The second option is to automatically identify the nearest tide gauge to the proxy location that has more than 20 years of observations. This option proves particularly useful when examining proxy records and extending the temporal range to capture recent changes in RSL.

The final option enables the selection of all tide gauges within a 1-degree radius (latitude and longitude) of the proxy location, provided they have more than 20 years of observations. This option grants users access to a wide array of tide gauges within a larger geographic area. Moreover, users can combine the first option with either the second or the third option, allowing for the freedom to choose specific tide gauges while incorporating the nearest tide gauge or multiple tide gauges.

All the values mentioned in this paragraph are arguments that can be adjusted within the function, giving users flexibility in customizing their data selection process according to their specific requirements.

3.3 Case Study for 1 location

In the following sections, we use one site, Cedar Island, North Carolina, USA, from the example dataset, `NAACproxydata`:

```
CedarIslandNC <- reslr::NAACproxydata %>%
  dplyr::filter(Site == "Cedar Island")

glimpse(CedarIslandNC)
```

```
#> Rows: 104
#> Columns: 8
#> $ Region    <chr> "North Carolina", "North Carolina", "North Carolina", "North~
#> $ Site      <chr> "Cedar Island", "Cedar Island", "Cedar Island", "Cedar Islan~
#> $ Latitude  <dbl> 34.971, 34.971, 34.971, 34.971, 34.971, 34.971, 34.971, 34.9~
#> $ Longitude <dbl> -76.38, -76.38, -76.38, -76.38, -76.38, -76.38, -76.38, -76.~
#> $ RSL       <dbl> -0.12, -0.14, -0.16, -0.18, -0.19, -0.21, -0.22, -0.23, -0.2~
#> $ Age       <dbl> 2005, 1996, 1988, 1979, 1974, 1963, 1957, 1951, 1941, 1937, ~
#> $ Age_err   <dbl> 2.25, 2.00, 5.00, 5.75, 5.50, 5.50, 7.00, 7.75, 7.75, 8.00, ~
#> $ RSL_err   <dbl> 0.06, 0.06, 0.06, 0.06, 0.06, 0.06, 0.06, 0.06, 0.06, 0.06, ~
```

For a single location such as this case study, we recommend using an EIV IGP or a NI spline in time as they estimate RSL changes over time. In this example, tide-gauge data is not included but it is an option available to the user if they require. The next example will demonstrate a more complex analysis with the inclusion of tide gauges.

After selecting the data site from the example dataset, we use the `reslr_load` function to process the data prior to running the statistical model, and it has a number of different settings that the user can alter depending on the model choice. One such setting is the `prediction_grid_res` option. This provides the resolution at which predictions of RSL and RSL rates are made and subsequently plotted. We set the default at 50 years, and if a finer grid is required, the user can alter the setting for `prediction_grid_res`. The `reslr_load` function includes additional settings to include tide-gauge data and linear rates, which will be discussed in the next case study. For the single site case study, we demonstrate the `reslr_load` function:

```
CedarIslandNC_input <- reslr_load(data = CedarIslandNC)
```

The output of this function is a list of two data frames called `data` and `data_grid`. The `data` dataframe is the inputted data with an additional column called `data_type_id` which distinguishes proxy records from tide-gauge data. The `data_grid` is a data frame that is evenly spaced in time based on the `prediction_grid_res` value chosen by the user and is used to create the plots. A brief insight into the outputs of the `reslr_input` function can be obtained using the print function, which provides the number of observations and the sources of the data, as shown below:

```
print(CedarIslandNC_input)
```

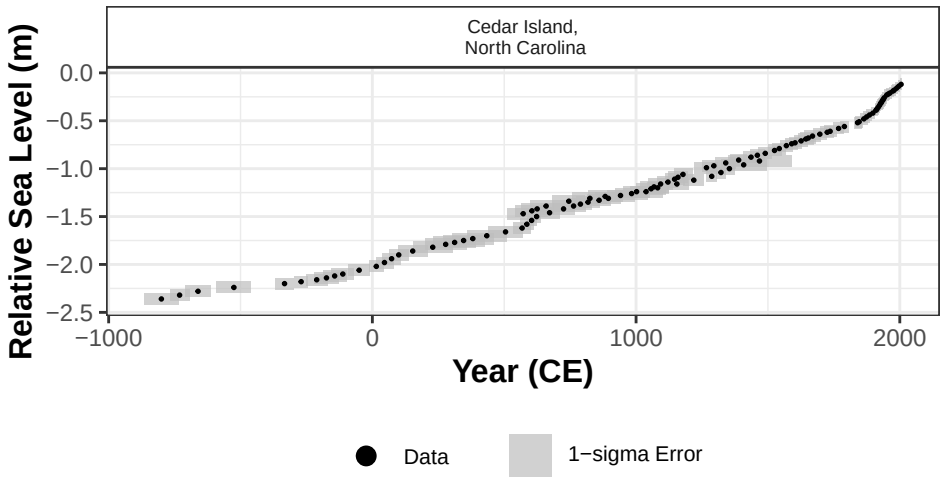

Figure 2: A plot of the raw data for our example site Cedar Island North Carolina. The x-axis is time in years in the Common Era (CE) and the y-axis is relative sea level in metres. The grey boxes are 1 standard deviation vertical and horizontal (temporal) uncertainty. The black dots are the midpoints of the uncertainty boxes.

```
#> This is a valid reslr input object with 104 observations and  1 site(s).
#> There are  1  proxy site(s) and  0  tide gauge site(s).
#> The age units are; Common Era.
#> Decadally averaged tide gauge data was not included. It is recommended for the ni_gam_decomp model
#> The linear_rate or linear_rate_err was not included. It is required for the ni_gam_decomp model
```

The next step is using the plot function to plot the raw data, shown in Figure 2, using the following:

```
plot(CedarIslandNC_input, plot_caption = FALSE)
```

Noisy Input Spline in time

The NI spline in time ("ni_spline_t") examines how the response variable, RSL, varies in time. While the EIV-IGP method is commonly used in the sea-level community, we demonstrate that the NI spline in time is a faster alternative. Unlike the Gaussian process, which has a computational complexity that grows exponentially with the number of data points, the spline equivalent uses pre-computed basis functions, resulting in a more efficient computation (Wood, 2017b).

For this model type, we use the `reslr_mcmc` function to implement the MCMC simulation using JAGS, and the model type setting is selected to be `model_type = "ni_spline_t"`.

```
res_ni_spline_t <- reslr_mcmc(input_data = CedarIslandNC_input,
                              model_type = "ni_spline_t")
```

The output of the `reslr_mcmc` function is a list that stores the JAGS model run, the input dataframe, and the dataframes for plotting the results. The user can set the size of the credible intervals by changing the `CI` setting in this function; the current default is `CI = 0.95`. In addition, the user can alter the number of iterations which will be required if the model is not converging.

To obtain a brief insight into the outputs of the `reslr_mcmc` function, the user can use the print function which provides the number of iterations and the model type:

```
print(res_ni_spline_t)
```

```
#> This is a valid reslr output object with 104 observations and  1 site(s).
#> There are  1  proxy site(s) and  0  tide gauge site(s).
#> The age units are; Common Era.
#> The model used was the Noisy Input Spline in time model.
#> The input data has been run via reslr_mcmc and has produced 3000 iterations over 3 MCMC chains.
```

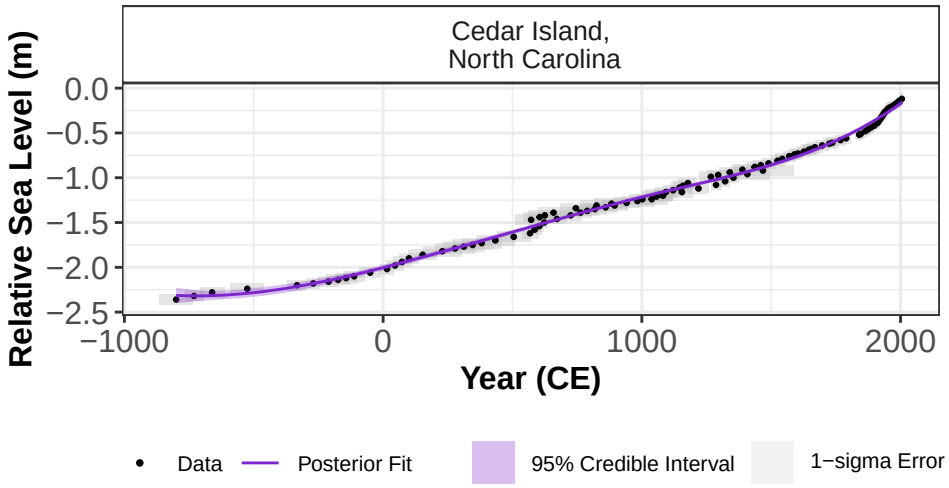

Figure 3: The plot of the Noisy Input spline in time model fit for our example site, Cedar Island, North Carolina. The x-axis is time in years in the Common Era (CE) and the y-axis is relative sea level in metres. The grey boxes are 1 standard deviation vertical and horizontal (temporal) uncertainty. The black dots are the midpoints of the uncertainty boxes. The solid purple line represents the posterior model fit with a 95% credible interval denoted by shading.

The convergence of the MCMC algorithm can be examined for the "ni_spline_t" model using the `summary` function and ensures the scale reduction factor (R-hat) is close to 1 (Gelman and Rubin, 1992; Gelman et al., 2013). If the model run has converged, the package will print: "No convergence issues detected". If the package prints: "Convergence issues detected, a longer run is necessary". The user is recommended to update the `reslr_mcmc` function with additional iterations as described above. The `summary` function provides insight into the parameter estimates from the model using the following:

```
summary(res_ni_spline_t)
```

```
#> No convergence issues detected.
```

```
#> # A tibble: 2 x 7
#>   variable      mean      sd     mad       q5    q95  rhat
#>   <chr>        <dbl>   <dbl>   <dbl>    <dbl>  <dbl> <dbl>
#> 1 sigma_beta 2.04    0.617   0.520   1.29     3.24    1.00
#> 2 sigma_y    0.00633 0.00478 0.00456 0.000513 0.0153  1.00
```

For the parameter estimates, "sigma_beta" acts as a smoothness parameter controlling the penalisation of the spline coefficients for the spline in time model, and "sigma_y" represents the data model variation. These are σ_y and σ_β as described in Section 2.

The final results from the "ni_spline_t" model can be illustrated using the `plot` function, and the corresponding dataframes are stored in the `res_ni_spline_t` object called `output_dataframes` as a named list element. Figure 3 demonstrates the posterior model fit for our example site using:

```
plot(res_ni_spline_t,
     plot_type = "model_fit_plot",
     plot_caption = FALSE)
```

In Figure 4, the rate of change of this posterior model fit is presented and can be viewed using:

```
plot(res_ni_spline_t,
     plot_type = "rate_plot",
     plot_caption = FALSE)
```

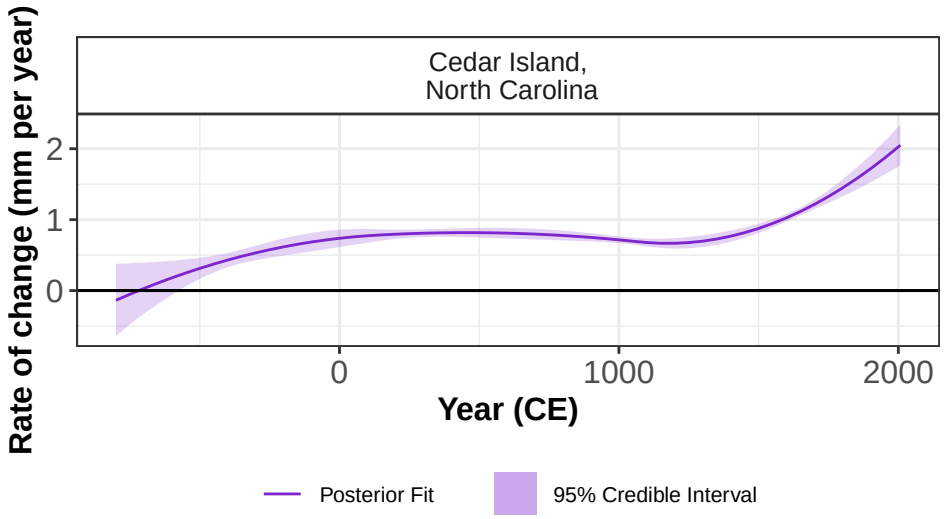

Figure 4: The rate of change model fit using the Noisy Input spline in time model for our example site, Cedar Island, North Carolina. The rate is calculated by taking the first derivative of the total model fit. The x-axis is time in years in the Common Era (CE) and the y-axis is the instantaneous rate of change of sea level in mm per year. The solid purple line represents the posterior model fit with a 95% credible interval denoted by shading. There is a black horizontal line which is the zero rate of change for this site.

3.4 Case Study for multiple sites

The sea-level research community is commonly interested in temporal and spatial variations in RSL. To cater to this interest, the **reslr** package offers two models for spatio-temporal modeling. The first model is a noisy-input spline that accounts for noise in both time and space, providing a robust representation of RSL dynamics. The second model, a more intricate option, is the Noisy Input GAM. Mathematical details concerning the Noisy Input GAM can be found in Upton et al. (2023). In our upcoming example, we will focus on this model as it empowers users to explore the decomposition of the RSL signal over time and space, unraveling valuable insights into the underlying dynamics.

Noisy Input Generalised Additive Model for decomposition of response signal

We demonstrate the functions settings required for the NI GAM. This model requires an adequate number of sites to perform the decomposition and the minimum sites required will depend on the signal in the data. In this example, we use nine sites from the example dataset, `NAACproxydata`, which are selected in the following manner:

```
multi_site <- reslr::NAACproxydata %>%
  dplyr::filter(Site %in% c("Cedar Island","Nassau",
                            "East River Marsh", "Swan Key",
                            "Placentia",
                            "Pelham Bay","Fox Hill Marsh",
                            "Snipe Key","Big River Marsh"))
```

Next, the `reslr_load` function is required for the preparation of input data for the NI GAM, which necessitates additional information not required by earlier models. Firstly, the statistical model relies on an estimate of the "linear local rate" and its associated uncertainty. By setting `include_linear_rate = TRUE`, the package incorporates this rate, which is assumed to stem from physical processes like Glacial Isostatic Adjustment (GIA). Users have the flexibility to include their preferred linear rate values as additional columns (`linear_rate` and `linear_rate_err`) in the input dataframe. If these values are not provided, the package automatically calculates them using the available data.

Secondly, users are encouraged to include tide-gauge data by setting `include_tide_gauge = TRUE`. As discussed previously, users need to make a decision regarding the inclusion

of the closest tide gauge (`TG_minimum_dist_proxy = TRUE`), selecting specific tide gauges by providing a list of names (`list_preferred_TGs = c("ARGENTIA")`), or including all tide gauges within a one-degree proximity of the proxy site (`all_TG_1deg = TRUE`). Additionally, the tide-gauge data requires values for the `linear_rate` and `linear_rate_err` columns, which are calculated using the ICE-5G (VM2) Model (Peltier, 2004) with an uncertainty value of 0.3 mm/year (Engelhart et al., 2009), both provided within the **reslr** package.

Thirdly, the tide-gauge data is averaged over a decade to align with the resolution of proxy records. If necessary, users can adjust the size of the averaging window to accommodate varying sediment accumulation rates. For example, a longer sediment accumulation rate would result in a larger average, such as 20 years. The default setting for `sediment_average_TG` is 10 years, which we will use in our example.

The final setting of the `reslr_load` function is `prediction_grid_res`, allowing users to modify the resolution of the output plots. The default setting of 50 years serves as a starting point, but users have the flexibility to explore alternative options. For our example, we will utilize nine proxy sites and select all tide-gauges within a one-degree range of our proxy site, maximizing the number of data points to demonstrate the capabilities of our package. The specific settings employed are described below:

```
multi_site_input <- reslr_load(
  data = multi_site,
  include_tide_gauge = TRUE,
  include_linear_rate = TRUE,
  TG_minimum_dist_proxy = TRUE,
  all_TG_1deg = TRUE)
```

Similar to the previous example, the output of this function is a list of two dataframes called `data` and `data_grid`. The `data` dataframe is the inputted data with additional columns for the data_type_id which will contain "ProxyRecord" and "TideGaugeData". The `data_grid` is a dataframe that is evenly spaced in time based on the `prediction_grid_res` value chosen by the user and is used to create the plots. In this example, we have 9 proxy sites and 26 tide gauges sites and the `print` provides insights into the outputs of the `reslr_input` function yields insights such as the number of observations (1,130).

```
print(multi_site_input)
```

```
#> This is a valid reslr input object with 1130 observations and  35 site(s).
#> There are  9  proxy site(s) and  26  tide gauge site(s).
#> The age units are; Common Era.
#> Decadally averaged tide gauge data included by the package.
#> The linear_rate and linear_rate_err has been included.
```

A plot of the raw data can be created using the `plot` function, with an option to plot the tide gauges and the proxy records together or have separate plots for each data source. Figure 5 demonstrates the resulting plot for the proxy records only using the following function:

```
plot(x = multi_site_input,
     plot_proxy_records = TRUE,
     plot_tide_gauges = FALSE,
     plot_caption = FALSE)
```

For this model type, the `reslr_mcmc` function should specify the `model_type = "ni_gam_decomp"`, and the MCMC simulation settings can be altered to ensure convergence.

```
res_ni_gam_decomp <- reslr_mcmc(
  input_data =  multi_site_input,
  model_type = "ni_gam_decomp"
)
```

The output of the `reslr_mcmc` function is a list that stores the JAGS model run, the input dataframe, and the dataframes for plotting the results. Identical to the other model processes, the convergence of the MCMC algorithm is examined and the parameter estimates from the model can be investigated using the following:

```
summary(res_ni_gam_decomp)
```

```
#> No convergence issues detected.
#> # A tibble: 4 x 7
#>   variable       mean      sd     mad     q5    q95  rhat
```

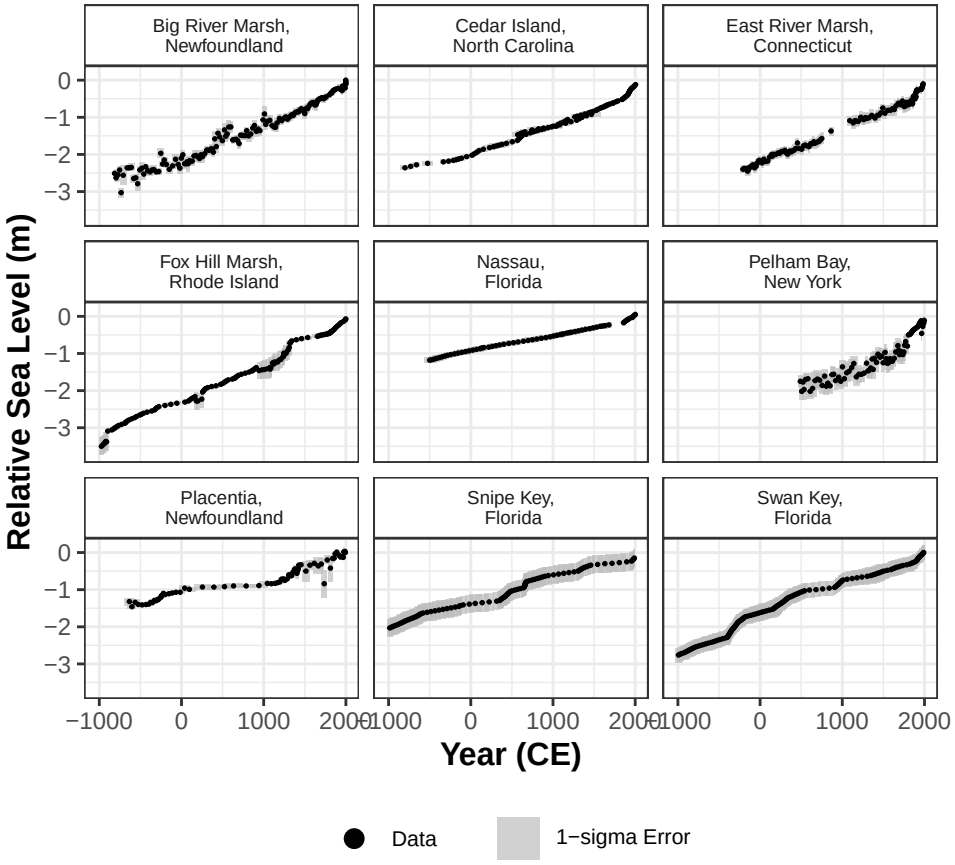

Figure 5: A plot of the raw data for our nine example sites along the Atlantic coast of North America. The x-axis is time in years in the Common Era (CE) and the y-axis is relative sea level in metres. The grey boxes are 1 standard deviation vertical and horizontal (temporal) uncertainty. The black dots are the midpoints of the uncertainty boxes.

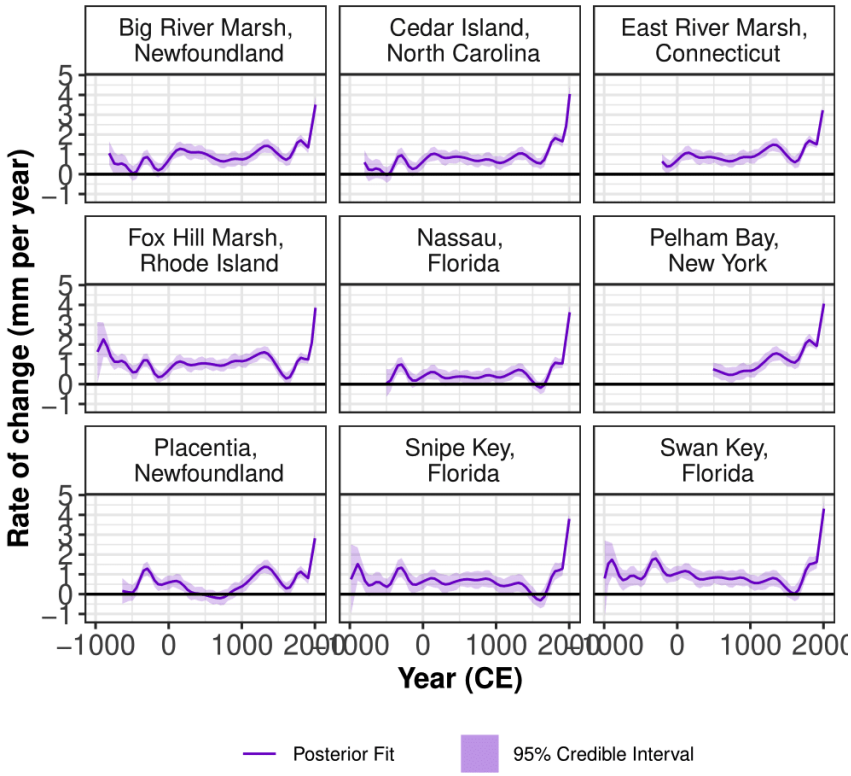

Figure 6: The rate of change for the total model fit for the Noisy Input generalised additive model for sites along the Atlantic coast of North America. It is calculated by finding the derivative of the total model fit. The solid purple line is the mean rate of change fit and the shading denotes a 95% credible interval for each site along the Atlantic coast of North America. The x-axis is time in years in the Common Era (CE), and the y-axis is the rate of change in mm per year.

```
#>   <chr>         <num>   <num>   <num>  <num>  <num> <num>
#> 1 sigma_beta_h 1.82   0.257   0.247   1.45   2.29    1.00
#> 2 sigma_beta_r 0.288  0.0547  0.0505  0.213  0.391   1.00
#> 3 sigma_beta_l 0.939  0.148   0.145   0.717  1.21    1.00
#> 4 sigma_y      0.0155 0.00107 0.00108 0.0138 0.0173  1.00
```

For the parameter estimates, we provide the standard deviation associated with each component of the NI GAM decomposition. Specifically, "sigma_beta_r" represents the standard deviation of the spline coefficient for the regional component, "sigma_beta_l" represents the standard deviation of the spline coefficient for the non-linear local component, "sigma_beta_h" denotes the standard deviation of the site-specific vertical offset component, and "sigma_y" indicates the data model variation. These names correspond to the algebraic components described in Section 2 above.

One of the key advantages of the NI GAM approach is its ability to decompose regional RSL change into separate components. The results from the `ni_gam_decomp` model can be visualized using the `plot` function, which generates individual plots for each component. Additionally, all components, except for the linear local component, have corresponding rate plots. Users can access the data used to create each plot in the `res_ni_gam_decomp` object as separate dataframes for each component.

In our example, we demonstrate the rate of change for the total model fit in Figure 6. This figure illustrates the rate of change at each site, which is useful for understanding the variations of the relative sea-level signal, i.e., $f(\mathbf{x}, t)$. To plot the rate of change, users can employ the following method:

```
plot(res_ni_gam_decomp,
     plot_type = "rate_plot",
     plot_caption = FALSE)
```

The regional component ($r(t)$) captures the mean of RSL change along the Atlantic coast of North America. The associated rate of change of the regional component, as seen in

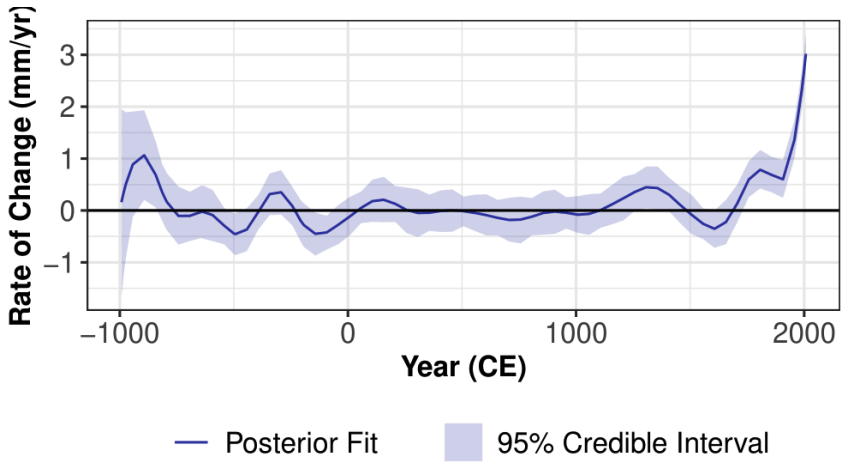

Figure 7: The rate of change for the regional component of the Noisy Input generalised additive model for the nine proxy sites and the eleven tide gauges along the Atlantic coast of North America. It is calculated by finding the derivative of the regional component fit. The solid blue line is the mean rate of change fit, and the shading denotes a 95% credible interval. The x-axis is time in years in the Common Era (CE), and the y-axis is the rate of change in mm per year.

Figure 7, provides an important visual insight into the rate at which this trend varied over the past 3,000 years. It is accessed by:

```
plot(res_ni_gam_decomp,
     plot_type = "regional_rate_plot",
     plot_caption = FALSE)
```

4 Summary

In this paper, we have presented an overview of the **reslr** package and discussed its various features and design decisions. Our goal was to address the specific needs of the paleo sea-level community and provide an efficient and flexible R package that caters to different types of source data, whilst maintaining a simple workflow that does not require the user to learn too many different functions.

Through two case studies, we demonstrated the simplicity and accessibility of the package. The first case study examined a single site using the NI spline in time. Our results showed that the **reslr** package can provide RSL estimates and associated rate of change values over time for a single location. In the second case study, we showcased the capabilities of the **reslr** package when analysing data from multiple locations. We highlighted its flexibility, allowing for the decomposition of the relative sea-level signal into different components. Additionally, we presented a comprehensive method for incorporating tide-gauge data, which can help to provide valuable insights into recent changes in RSL not captured by proxy records.

There are several potential extensions for the **reslr** package. A potential enhancement involves replacing our current MCMC algorithm, employed via JAGS software (Plummer, 2003), with more efficient alternatives like Integrated Nested Laplace Approximations (INLA: Rue et al., 2009). Another improvement could be the integration of other instrumental data sources, such as satellite data, enabling the examination of other variables related to climate change. Overall, the **reslr** package offers a powerful toolkit for the paleo sea-level community, and we anticipate that it will continue to evolve and expand its capability to meet the evolving needs of researchers in this field.

5 Acknowledgements

Upton's work is supported by A4 (Aigéin, Aeráid, agus athrú Atlantaigh); the project is funded by the Marine Institute (grant: PBA/CC/18/01). Parnell's work is supported by the

SFI awards 17/CDA/4695; 16/IA/4520; 12/RC/2289P2. Cahill'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.

6 Appendix

6.1 Example dataset

The **reslr** package contains a dataset used as an example called `NAACproxydata`. This dataset contains proxy records from the Atlantic coast of North America as used in Upton et al. (2023). The 21 different proxy data sites and the references for each data source can be found in Table 3. The **reslr** package contains a dataset used as an example called `NAACproxydata`. This dataset contains proxy records from the Atlantic coast of North America as used in Upton et al. (2023). The 21 different proxy data sites and the references for each data source can be found in Table 3.

Site Name	Reference
Barn Island, Connecticut	(Donnelly et al., 2004; Gehrels et al., 2020)
Big River Marsh, Newfoundland	(Kemp et al., 2018)
Cape May Courthouse, New Jersey	(Kemp et al., 2013; Cahill et al., 2016)
Cedar Island, North Carolina	(Kemp et al., 2011, 2017)
Cheesequake, New Jersey	(Walker et al., 2021)
Chezzetcook Inlet, Nova Scotia	(Gehrels et al., 2020)
East River Marsh, Connecticut	(Kemp et al., 2015; Stearns et al., 2023)
Fox Hill Marsh, Rhode Island	(Stearns et al., 2023)
Leeds Point, New Jersey	(Kemp et al., 2013; Cahill et al., 2016)
Les Sillons, Magdelen Islands	(Barnett et al., 2017)
Little Manatee River, Florida	(Gerlach et al., 2017)
Nassau, Florida	(Kemp et al., 2014)
Pelham Bay, New York	(Kemp et al., 2017; Stearns et al., 2023)
Placentia, Newfoundland	(Kemp et al., 2018)
Revere, Massachusetts	(Donnelly, 2006)
Saint Simeon, Quebec	(Barnett et al., 2017)
Sanborn Cove, Maine	(Gehrels et al., 2020)
Sand Point, North Carolina	(Kemp et al., 2011, 2017)
Snipe Key, Florida	(Khan et al., 2022)
Swan Key, Florida	(Khan et al., 2022)
Wood Island, Massachusetts	(Kemp et al., 2011)

Table 3: Presents the names of all the sites available in the example dataset within the **reslr** package. For each site, we include the reference in the literature to the source of the data.

References

T. Aarup, M. Merrifield, B. Pérez Gómez, I. Vassie, and P. Woodworth. Manual on Sea-level Measurements and Interpretation, Volume IV : An update to 2006. *Intergovernmental Oceanographic Commission of UNESCO*, 4, 2006. URL `https://psmsl.org/train_and_info/training/manuals`. [p62]

E. L. Ashe, N. Cahill, C. Hay, N. S. Khan, A. Kemp, S. E. Engelhart, B. P. Horton, A. C. Parnell, and R. E. Kopp. Statistical modeling of rates and trends in Holocene relative sea level. *Quaternary Science Reviews*, 204:58–77, 2019. doi: https://doi.org/10.1016/j.quascirev.2018.10.032. [p61]

R. L. Barnett, P. Bernatchez, M. Garneau, and M.-N. Juneau. 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, 2017. doi: https://doi.org/10.1002/jqs.2931. [p76]

M. J. Brain, A. J. Long, S. A. Woodroffe, D. N. Petley, D. G. Milledge, and A. C. Parnell. Modelling the effects of sediment compaction on salt marsh reconstructions of recent sea-level rise. *Earth and Planetary Science Letters*, 345-348:180–193, 2012. ISSN 0012-821X. doi: https://doi.org/10.1016/j.epsl.2012.06.045. [p64]

N. Cahill, A. C. Kemp, B. P. Horton, and A. C. Parnell. Modeling sea-level change using Errors-in-Variables integrated Gaussian Process 1. *The Annals of Applied Statistics*, 9(2): 547–571, 2015a. doi: https://doi.org/10.1214/15-AOAS824. [p61, 63, 64, 65]

N. Cahill, S. Rahmstorf, and A. C. Parnell. Change points of global temperature. *Environmental research letters*, 10(8):84002, 2015b. doi: https://doi.org/10.1088/1748-9326/10/8/084002. [p61, 64]

N. Cahill, A. C. Kemp, B. P. Horton, and A. C. Parnell. A Bayesian hierarchical model for reconstructing relative sea level: from raw data to rates of change. *Climate of the Past*, 12 (2):525–542, 2016. doi: https://doi.org/10.5194/cp-12-525-2016. [p76]

B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. A. Brubaker, J. Guo, P. Li, and A. Riddell. Stan: A probabilistic programming language. *Journal of statistical software*, 76, 2017. doi: 10.18637/jss.v076.i01. URL https://www.jstatsoft.org/index.php/jss/article/view/v076i01. [p63]

C. de Boor. A Practical Guide to Spline. *Applied Mathematical Sciences, New York: Springer, 1978*, 27, 01 1978. doi: 10.2307/2006241. [p61, 65]

S. Dean, B. P. Horton, N. Evelpidou, N. Cahill, G. Spada, and D. Sivan. Can we detect centennial sea-level variations over the last three thousand years in Israeli archaeological records? *Quaternary Science Reviews*, 210:125–135, 2019. ISSN 0277-3791. doi: https://doi.org/10.1016/j.quascirev.2019.02.021. URL https://www.sciencedirect.com/science/article/pii/S0277379117310831. [p64]

D. K. Dey, S. K. Ghosh, and B. K. Mallick. *Generalized linear models: a Bayesian perspective*. CRC Press, 2000. doi: https://doi.org/10.1201/9781482293456. [p61, 63]

P. Dierckx. *Curve and surface fitting with splines*. Oxford University Press, 1995. URL https://books.google.ie/books?id=-RIQ3SR0sZMC. [p65]

J. P. Donnelly. A Revised Late Holocene Sea-Level Record for Northern Massachusetts, USA. *Journal of Coastal Research*, 22(5):1051–1061, 2006. doi: https://doi.org/l0.2112/04-0207.1. [p76]

J. P. Donnelly, P. Cleary, P. Newby, and R. Ettinger. 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), 2004. doi: https://doi.org/10.1029/2003gl018933. [p76]

R. Edwards and A. Wright. *Foraminifera: Handbook of Sea-Level Research*, chapter 13, pages 191–217. John Wiley & Sons, Ltd, 2015. ISBN 9781118452547. doi: https://doi.org/10.1002/9781118452547.ch13. [p62]

P. Eilers and B. Marx. Flexible Smoothing with B-splines and Penalities. *Statistical Science*, 11, 1996. doi: https://doi.org/10.1214/ss/1038425655. [p65]

S. E. Engelhart, B. P. Horton, B. C. Douglas, W. R. Peltier, and T. E. Törnqvist. Spatial variability of late Holocene and 20th century sea-level rise along the Atlantic coast of the United States. *Geology*, 37(12):1115–1118, 12 2009. ISSN 0091-7613. doi: 10.1130/G30360A.1. [p64, 66, 72]

R. W. Gehrels. Determining Relative Sea-Level Change from Salt-Marsh Foraminifera and Plant Zones on the Coast of Maine, U.S.A. *Journal of Coastal Research*, 10(4):990–1009, 1994. URL https://www.jstor.org/stable/4298291. [p61, 62]

W. Gehrels, S. Dangendorf, N. L. M. Barlow, M. H. Saher, A. J. Long, P. L. Woodworth, C. G. Piecuch, and K. Berk. A Preindustrial Sea Level Rise Hotspot Along the Atlantic Coast of North America. *Geophysical Research Letters*, 47(4), 2020. doi: https://doi.org/10.1029/2019GL085814. [p76]

A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472, 1992. ISSN 08834237. URL http://www.jstor.org/stable/2246093. [p70]

A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian data analysis*. CRC press, 2013. doi: https://doi.org/10.1201/b16018. [p70]

M. J. Gerlach, S. E. Engelhart, A. C. Kemp, R. P. Moyer, J. M. Smoak, C. E. Bernhardt, and N. Cahill. Reconstructing Common Era relative sea-level change on the Gulf Coast of Florida. *Marine Geology*, 390:254–269, 2017. ISSN 0025-3227. doi: https://doi.org/10.1016/j.margeo.2017.07.001. [p76]

V. Gornitz. *Paleoclimate Proxies, An Introduction*, pages 716–721. Springer Netherlands, 2009. doi: https://doi.org/10.1007/978-1-4020-4411-3. [p62]

J. M. Gregory, S. M. Griffies, C. W. Hughes, J. A. Lowe, J. A. Church, I. Fukimori, N. Gomez, R. E. Kopp, F. Landerer, G. L. Cozannet, R. M. Ponte, D. Stammer, M. E. Tamisiea, and R. S. van de Wal. Concepts and Terminology for Sea Level: Mean, Variability and Change, Both Local and Global. *Surveys in Geophysics*, 40(6):1251–1289, 11 2019. ISSN 15730956. URL https://link.springer.com/article/10.1007/s10712-019-09525-z. [p66, 67]

A. Grinsted. Projected Change—Sea Level. *Second assessment of climate change for Baltic Sea basin*, pages 253–263, 2015. URL https://link.springer.com/chapter/10.1007/978-3-319-16006-1_14. [p66]

A. D. Hawkes, A. C. Kemp, J. P. Donnelly, B. P. Horton, W. R. Peltier, N. Cahill, D. F. Hill, E. Ashe, and C. R. Alexander. Relative sea-level change in northeastern Florida (USA) during the last 8.0ka. *Quaternary Science Reviews*, 142:90–101, 2016. ISSN 0277-3791. doi: https://doi.org/10.1016/j.quascirev.2016.04.016. URL https://www.sciencedirect.com/science/article/pii/S0277379116301275. [p64]

P. Hogarth, C. Hughes, S. Williams, and C. Wilson. Improved and extended tide gauge records for the british isles leading to more consistent estimates of sea level rise and acceleration since 1958. *Progress in Oceanography*, 184:102333, 2020. ISSN 0079-6611. doi: https://doi.org/10.1016/j.pocean.2020.102333. URL https://www.sciencedirect.com/science/article/pii/S0079661120300720. [p64]

S. J. Holgate, A. Matthews, P. L. Woodworth, L. J. Rickards, M. E. Tamisiea, E. Bradshaw, P. R. Foden, K. M. Gordon, S. Jevrejeva, and J. Pugh. New Data Systems and Products at the Permanent Service for Mean Sea Level. *Journal of Coastal Research*, 29(3):493 – 504, 2013. doi: 10.2112/JCOASTRES-D-12-00175.1. [p61, 62, 68]

T. Holsclaw, B. Sansó, H. K. H. L. Lee, K. Heitmann, S. Habib, D. Higdon, and U. Alam. Gaussian process modeling of derivative curves. *Technometrics*, 55(1):57–67, 2013. ISSN 00401706, 15372723. URL http://www.jstor.org/stable/24587388. [p65]

B. P. Horton, R. E. Kopp, A. J. Garner, C. C. Hay, N. S. Khan, K. Roy, and T. A. Shaw. Mapping Sea-Level Change in Time, Space, and Probability. *Annual Review of Environment and Resources*, 43(1):481–521, 2018. ISSN 1543-5938. doi: https://doi.org/10.1146/annurev-environ-102017-025826. [p67]

A. C. Kemp and R. J. Telford. *Transfer functions: Handbook of Sea-Level Research*, chapter 31, pages 470–499. John Wiley & Sons, Ltd, 2015. ISBN 9781118452547. doi: https://doi.org/10.1002/9781118452547.ch31. [p62, 63]

A. C. Kemp, B. P. Horton, S. J. Culver, D. R. Corbett, O. van de Plassche, W. R. Gehrels, B. C. Douglas, and A. C. Parnell. Timing and magnitude of recent accelerated sea-level rise (North Carolina, United States). *Geology*, 37(11):1035–1038, 11 2009. ISSN 0091-7613. doi: https://doi.org/10.1130/G30352A.1. [p64]

A. C. Kemp, B. P. Horton, J. P. Donnelly, M. E. Mann, M. Vermeer, and S. Rahmstorf. Climate related sea-level variations over the past two millennia. *Proceedings of the National Academy of Sciences*, 108(27):11017–11022, 2011. doi: 10.1073/pnas.1015619108. [p76]

A. C. Kemp, B. P. Horton, C. H. Vane, C. E. Bernhardt, D. R. Corbett, S. E. Engelhart, S. C. Anisfeld, A. C. Parnell, and N. Cahill. Sea level change during the last 2500 years in New Jersey, USA. *Quaternary Science Reviews*, 81:90–104, 2013. ISSN 02773791. doi: 10.1016/j.quascirev.2013.09.024. [p61, 76]

A. C. Kemp, C. E. Bernhardt, B. P. Horton, R. E. Kopp, C. H. Vane, W. R. Peltier, A. D. Hawkes, J. P. Donnelly, A. C. Parnell, and N. Cahill. Late Holocene sea- and land-level change on the U.S. southeastern Atlantic coast. *Marine Geology*, 357:90–100, 2014. ISSN 0025-3227. doi: https://doi.org/10.1016/j.margeo.2014.07.010. [p76]

A. C. Kemp, A. D. Hawkes, J. P. Donnelly, C. H. Vane, B. P. Horton, T. D. Hill, S. C. Anisfeld, A. C. Parnell, and N. Cahill. Relative sea level change in Connecticut (USA) during the last 2200 yrs. *Earth and Planetary Science Letters*, 428:217–229, 2015. ISSN 0012-821X. doi: https://doi.org/10.1016/j.epsl.2015.07.034. [p76]

A. C. Kemp, T. D. Hill, C. H. Vane, N. Cahill, P. M. Orton, S. A. Talke, A. C. Parnell, K. Sanborn, and E. K. Hartig. Relative sea-level trends in New York City during the past 1500 years. *The Holocene*, 27(8):1169–1186, 2017. doi: 10.1177/0959683616683263. [p64, 76]

A. C. Kemp, A. J. Wright, R. J. Edwards, R. L. Barnett, M. J. Brain, R. E. Kopp, N. Cahill, B. P. Horton, D. J. Charman, A. D. Hawkes, T. D. Hill, and O. van de Plassche. Relative sea-level change in Newfoundland, Canada during the past 3000 years. *Quaternary Science Reviews*, 201:89–110, 2018. doi: https://doi.org/10.1016/j.quascirev.2018.10.012. [p61, 62, 65, 76]

N. S. Khan, E. Ashe, T. A. Shaw, M. Vacchi, J. Walker, W. R. Peltier, R. E. Kopp, and B. P. Horton. Holocene relative sea-level changes from near-, intermediate-, and far-field locations. *Current Climate Change Reports*, 1(4):247–262, 2015. doi: https://doi.org/10.1007/s40641-015-0029-z. [p61]

N. S. Khan, E. Ashe, R. P. Moyer, A. C. Kemp, S. E. Engelhart, M. J. Brain, L. T. Toth, A. Chappel, M. Christie, R. E. Kopp, and B. P. Horton. Relative sea-level change in South Florida during the past 5000 years. *Global and Planetary Change*, 216:103902, 2022. ISSN 0921-8181. doi: https://doi.org/10.1016/j.gloplacha.2022.103902. [p76]

J. R. Kirby, E. Garrett, and W. R. Gehrels. Holocene relative sea-level changes in northwest ireland: An empirical test for glacial isostatic adjustment models. *The Holocene*, 2023. doi: https://doi.org/10.1177/09596836231169992. [p64]

R. E. Kopp. Does the mid-Atlantic United States sea level acceleration hot spot reflect ocean dynamic variability? *Geophysical Research Letters*, 40(15):3981–3985, 8 2013. doi: 10.1002/GRL.50781. [p61, 65]

R. E. Kopp, F. J. Simons, J. X. Mitrovica, A. C. Maloof, and M. Oppenheimer. Probabilistic assessment of sea level during the last interglacial stage. *Nature*, 462(7275):863–867, 2009. ISSN 00280836. doi: https://doi.org/10.1038/nature08686. [p65]

R. E. Kopp, A. C. Kemp, K. Bittermann, B. P. Horton, J. P. Donnelly, R. W. Gehrels, C. C. Hay, J. X. Mitrovica, E. D. Morrow, and S. Rahmstorf. 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, 2016. ISSN 10916490. doi: 10.1073/pnas.1517056113. [p61, 65]

W. Marshall. *Chronohorizons: Handbook of Sea-Level Research*, chapter 25, pages 373–385. John Wiley & Sons, Ltd, 2015. ISBN 9781118452547. doi: https://doi.org/10.1002/9781118452547.ch25. [p62]

A. McHutchon and C. E. Rasmussen. Gaussian Process training with input noise. *Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011, NIPS 2011*, pages 1–9, 2011. URL https://proceedings.neurips.cc/paper_files/paper/2011/file/a8e864d04c95572d1aece099af852d0a-Paper.pdf. [p61, 63]

A. J. Meltzner, A. D. Switzer, B. P. Horton, E. Ashe, Q. Qiu, D. F. Hill, S. L. Bradley, R. E. Kopp, E. M. Hill, J. M. Majewski, D. H. Natawidjaja, and B. W. Suwargadi. Half-metre sea-level fluctuations on centennial timescales from mid-Holocene corals of Southeast Asia. *Nature Communications*, 8(1):14387, 2017. doi: 10.1038/ncomms14387. [p61]

W. Peltier. 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, 2004. doi: https://doi.org/10.1146/annurev.earth.32.082503.144359. [p72]

M. Plummer. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing, TUWien*, 124, 2003. URL https://api.semanticscholar.org/CorpusID:265976949. [p63, 75]

M. Plummer, A. Stukalov, and M. Denwood. rjags: Bayesian graphical models using MCMC. *R package version*, 4(6), 2016. URL https://search.r-project.org/CRAN/refmans/rjags/html/rjags-package.html. [p63]

D. Pugh and P. Woodworth. *Tidal forces: Sea-Level Science: Understanding Tides, Surges, Tsunamis and Mean Sea-Level Changes*, page 36–59. Cambridge University Press, 2014. doi: 10.1017/CBO9781139235778.006. [p62]

C. E. Rasmussen and C. K. I. Williams. Gaussian Processes for Machine Learning. Technical report, MIT Press, 2006. [p65]

H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 71(2):319–392, 2009. doi: https://doi.org/10.1111/j.1467-9868.2008.00700.x. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9868.2008.00700.x. [p75]

T. A. Shaw, A. J. Plater, J. R. Kirby, K. Roy, S. Holgate, P. Tutman, N. Cahill, and B. P. Horton. Tectonic influences on late Holocene relative sea levels from the central-eastern Adriatic coast of Croatia. *Quaternary Science Reviews*, 200:262–275, 2018. ISSN 0277-3791. doi: https://doi.org/10.1016/j.quascirev.2018.09.015. URL https://www.sciencedirect.com/science/article/pii/S0277379118300659. [p64]

I. Shennan and B. Horton. Holocene land- and sea-level changes in Great Britain. *Journal of Quaternary Science*, 17(5-6):511–526, 2002. doi: https://doi.org/10.1002/jqs.710. [p64]

I. Shennan, A. J. Long, and B. P. Horton. Handbook of sea-level research. *John Wiley & Sons, Ltd*, 2015. doi: https://doi.org/10.1002/9781118452547. [p62]

G. L. Simpson. Modelling palaeoecological time series using generalised additive models. *Frontiers in Ecology and Evolution*, 6:149, 2018. doi: https://doi.org/10.3389/fevo.2018.00149. [p61]

R. B. Stearns, S. E. Engelhart, A. C. Kemp, T. D. Hill, M. J. Brain, and D. R. Corbett. Within-region replication of late Holocene relative sea-level change: An example from southern New England, United States. *Quaternary Science Reviews*, 300:107868, 2023. ISSN 0277-3791. doi: https://doi.org/10.1016/j.quascirev.2022.107868. URL https://www.sciencedirect.com/science/article/pii/S0277379122004991. [p64, 76]

M. Upton, A. Parnell, A. Kemp, E. Ashe, G. McCarthy, and N. Cahill. A noisy-input generalised additive model for relative sea-level change along the atlantic coast of north america. *arXiv preprint arXiv:2301.09556*, 2023. doi: https://doi.org/10.48550/arXiv.2301.09556. [p61, 62, 63, 65, 66, 67, 71, 76]

J. S. Walker, R. E. Kopp, T. A. Shaw, N. Cahill, N. S. Khan, D. C. Barber, E. L. Ashe, M. J. Brain, J. L. Clear, D. R. Corbett, and B. P. Horton. Common Era sea-level budgets along the U.S. Atlantic coast. *Nature Communications*, 12(1):1841, 2021. doi: 10.1038/s41467-021-22079-2. [p61, 65, 76]

P. L. Whitehouse. Glacial isostatic adjustment modelling: historical perspectives, recent advances, and future directions. *Earth Surf. Dynam*, 6:401–429, 2018. doi: https://doi.org/10.5194/esurf-6-401-2018. [p66]

H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. URL https://ggplot2.tidyverse.org. [p67]

S. Wood. Package 'mgcv'. *R package version*, 1(29):729, 2015. URL https://cran.r-project.org/web/packages/mgcv/index.html. [p62]

S. N. Wood. Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models. *Biometrics*, 62(4):1025–1036, 2006. doi: https://doi.org/10.1111/j.1541-0420.2006.00574.x. [p65]

S. N. Wood. P-splines with derivative based penalties and tensor product smoothing of unevenly distributed data. *Statistics and Computing*, 27(4):985–989, 2017a. doi: https://doi.org/10.1007/s11222-016-9666-x. [p65]

S. N. Wood. *Generalized additive models: an introduction with R*. CRC Press, 2017b. doi: https://doi.org/10.1201/9781315370279. [p66, 69]

P. Woodworth and R. Player. The Permanent Service for Mean Sea Level: An update to the 21st century. *Journal of Coastal Research*, 19:287–295, 03 2003. doi: https://www.jstor.org/stable/4299170. [p62, 68]

Maeve Upton
Maynooth University
Hamilton Institute, ICARUS and Department of Mathematics and Statistics
Maynooth, Ireland
ORCiD: 0000-0002-1151-6657
uptonmaeve010@gmail.com

Andrew Parnell
Maynooth University
Hamilton Institute, ICARUS and Department of Mathematics and Statistics
Maynooth, Ireland
ORCiD: 0000-0001-7956-7939
andrew.parnell@mu.ie

Niamh Cahill
Maynooth University
Department of Mathematics and Statistics and ICARUS
Maynooth, Ireland
ORCiD: 0000-0003-3086-550X
niamh.cahill@mu.ie