~~Bias Correction of Climate Models using a~~ Bayesian Hierarchical Model for Bias Correcting Climate Models

Jeremy Carter[1,2], Erick A. Chacón-Montalván[4], and Amber Leeson[2,3]

[1]Department of Mathematics and Statistics, Lancaster University, Lancaster, UK
[2]Data Science Institute, University of Lancaster, Lancaster, UK
[3]Lancaster Environment Center, University of Lancaster, Lancaster, UK
[4]Escuela Profesional de Ingeniería Estadística, Universidad Nacional de Ingeniería, Lima, Peru

Correspondence: Jeremy Carter (j.carter10@lancaster.ac.uk)

Abstract. Climate models, derived from process understanding, are essential tools in the study of climate change and its wide-ranging impacts~~on the biosphere~~. Hindcast and future simulations provide comprehensive spatiotemporal estimates of climatology that are frequently employed within the environmental sciences community, although the output can be afflicted with bias that impedes direct interpretation. ~~Bias correction approaches using observational data aim~~ Post-processing, bias correction approaches utilise observational data to address this challenge~~. However, approaches~~ , although are typically criticised for not being physically justified and not considering uncertainty in the correction. ~~These aspects are particularly important in cases where observations are sparse, such as for weather station data over Antarctica. This paper attempts to address both of these issues through the development of~~ This paper proposes a novel Bayesian ~~hierarchical model for bias prediction. The model~~ bias correction framework that propagates uncertainty robustly and ~~uses latent Gaussian process distributions to capture~~ models underlying spatial covariance patterns~~,~~. Shared latent Gaussian processes are assumed between the in situ observations and climate model output with the aim of partially preserving the covariance structure from the climate model after bias correction, which is based on well-established physical laws. ~~The Bayesian framework can handle complex modelling structures and provides an approach that is flexible and adaptable to specific areas of application, even increasing the scope of the work to data assimilation tasks more generally. Results in this paper are presented for one-dimensional simulated examples for clarity, although the method implementation has been developed to also work on multidimensional data as found in most real applications. Performance under different simulated scenariosis examined, with the method providing~~ Results demonstrate added value in modelling shared generating processes under several simulated scenarios, with most value added ~~over alternative approaches in~~ for the case of sparse in situ observations and smooth underlying bias. ~~A major benefit of the model is the robust~~ Additionally, the propagation of uncertainty to a simulated final bias corrected time series is illustrated, which is of key importance to a range of stakeholders, from climate scientists engaged in impact studies, decision makers trying to understand the likelihood of particular scenarios and individuals involved in climate change adaption strategies where accurate risk assessment is required for optimal resource allocation. This paper focuses on one-dimensional simulated examples for clarity, although the code implementation is developed to also work on multi-dimensional input data, encouraging follow-on real-world application studies that will further validate performance and remaining limitations. The Bayesian framework

supports uncertainty propagation under model adaptations required for specific applications, providing a flexible approach that increases the scope to data assimilation tasks more generally.

1 Introduction

Climate models are invaluable in the study of climate change and its impacts (Bader et al., 2008; Flato et al., 2013). Formulated from physical laws and with parameterisation and process understanding derived from past observations; climate models provide comprehensive spatiotemporal estimates of our past, current and future climate under different emission scenarios. Global climate models (GCMs) simulate important climatological features and drivers such as storm tracks and the El Niño–Southern Oscillation (ENSO) (Greeves et al., 2007; Guilyardi et al., 2009). In addition, independently developed GCMs agree on the future direction of travel for many important features such as global temperature rise under continued net-positive emission scenarios (Tebaldi et al., 2021). However, GCMs are computationally expensive to run and also exhibit significant systematic errors, particularly on regional scales (Cattiaux et al., 2013; Flato et al., 2013). Regional climate models (RCMs) aim to dynamically downscale GCMs and more accurately represent climatology for specific regions of interest and have parameterisation, tuning and additional physical schemes optimised to the region (Giorgi, 2019; Doblas-Reyes et al., 2021). Despite this, significant systematic errors remain, particularly for regions with complex climatology and with sparse in situ observations available to inform process understanding, such as over Antarctica (Carter et al., 2022). These systematic errors inhibit the direct interpretation of climate model output, particularly important in impact assessments (Ehret et al., 2012; Liu et al., 2014; Sippel et al., 2016).

There are many fundamental causes of systematic errors in climate models, including: the absence or imperfect representation of physical processes; errors in initialisation; influence of boundary conditions and finite resolution (Giorgi, 2019). The inherent complexity and computationally expensive nature of climate models makes direct reduction of systematic errors through climate model development and tuning challenging (Hourdin et al., 2017). ~~End~~ Additionally, end users are typically interested in only a narrow aspect of the output (e.g. possibly only one or two variables), which the climate model is unlikely to be specifically tuned for. Post-processing, bias correction techniques allow improvements to the consistency, quality and value of climate model output, specific to the end user's focus of interest, with manageable computational cost and without requirement of in-depth knowledge behind the climate model itself (Ehret et al., 2012). ~~Different end users are focused on different types of systematic errors, whether that's errors in the mean climatology, the multi-year trends or in other features of the output such as the covariance structure~~Transfer functions are derived between the climate model output and in situ observational data to correct components such as the mean (Das et al., 2022) or probability density functions (PDFs) of the data (Qian and Chang, 2021). This paper ~~follows a common approach to focus on~~ focuses on providing a novel framework for correcting systematic errors in the ~~parameters that describe the probability density function (PDF) at each site. Further, detailed discussion of this is given in Sect. 2 as are approaches to bias correction within this context~~PDF of the climate model output at each grid point.

One of the fundamental issues often attached to post-processing bias correction is the lack of justification based on known physical laws and process understanding (Ehret et al., 2012; Maraun, 2016). ~~Transfer functions are derived that are applied to the climate data to improve some aspect of consistency with observations, such as the mean in for example the delta method (Das et al., 2022) or the overall PDF in the case of quantile mapping (Qian and Chang, 2021).~~ The spatiotemporal field

and associated covariance structure from the climate model, which is consistent with accepted physical laws, is typically not considered and so not preserved. Resulting corrected fields may exhibit too smooth or sharply varying behaviour over the region and discontinuities near observations. In addition, many approaches of bias correction fail to adequately handle uncertainties or estimate them at all. Reliable uncertainty estimation is valuable for inclusion in impact studies to inform resulting decision making. This is especially true for regimes with tipping points, such as ice shelf collapse over Antarctica, where uncertainties in the climatology can cause a regime shift (DeConto and Pollard, 2016).

In this paper these issues are partially addressed through the development of a fully Bayesian ~~approach using a hierarchical structure and~~ hierarchical approach to bias correction. Parameter uncertainties are propagated through the hierarchical model and underlying spatial covariance structures are captured with latent Gaussian processes (~~GP) is proposed for bias correction, discussed in detail in Sect. 3. Parameter uncertainties are propagated through the model and the underlying covariance structure is derived both from observations and the~~ GPs) for both in situ observations and the climate model output.

The approach presented builds on that of Lima et al. (2021), which models the in situ observational data as generated from a GP and uses quantile mapping (Qian and Chang, 2021) to apply the correction to the climate model output. In Lima et al. (2021) the spatial covariance structure of the climate model output is not considered and uncertainty is not propagated to the final bias corrected time series. The novelty of the approach proposed here is that shared latent GPs are modelled between the climate model output and the in situ observational data, which aims to incorporate information from the physically realistic spatial patterns of the climate model output in predictions of the unbiased field. Additionally, uncertainty is propagated through the quantile mapping step, which results in uncertainty bands on the bias corrected output. The approach is developed with the focus of applying bias correction to regions with sparse in situ observations, such as over Antarctica, where capturing uncertainty in the correction is of key importance and where including data from all sources during inference is particularly valuable. ~~In the method, climate model output is assumed to be generated from two underlying and independent stochastic processes, one relating to the true underlying field of interest (that also generates the in situ observations) and one that generates the bias present in the climate data. The aim is to separate these two processes and to infer their covariance structures. Posterior predictive estimates of the true underlying field across the region can then be made, which in turn can be used for bias correction. The ability of the model in doing this depends on factors such as the density of observations and the relative smoothness of the truth and bias components. Simulated data is used to test the performance under~~ Performance under simulated scenarios with differing data density and ~~latent~~ underlying covariance length scales ~~, with results and discussion presented in Sect. 4.~~ is evaluated in this paper and the potential added value assessed when compared with the approach in Lima et al. (2021).

~~The model is developed~~ Developing the bias correction approach in a flexible Bayesian framework ~~, where adjustments~~means further adjustments/advancements that are necessary for real-world scenarios can easily be incorporated while maintaining robust uncertainty propagation. For example, extra predictors, such as elevation and latitude, can be included either in the mean function or covariance matrix of the latent GPs. Alternatively, the model could be expanded to incorporate a temporal component of the bias accounting for variability across different seasons. This flexibility is important and increases the scope of the work, allowing the ~~model~~ methodology to be applied to a wide range of scenarios, including for example application

to many different meteorological fields and also combining observation data from different instruments rather than necessarily with respect to climate model output. Additionally, the Bayesian framework allows incorporation of domain specific, expert knowledge of different data sources and their uncertainties through the choice of prior distributions.

2 Methodology

The goal of the methodology developed and presented in this paper is to evaluate the bias in the climate model output across the domain in a framework that captures uncertainty robustly and that preserves information available from both the in situ observations and climate model output on underlying spatial structures. The resulting predictive bias can be coupled with known bias correction methods, such as quantile mapping, with the benefits of uncertainty quantification and inherited spatial structure. The overarching approach is summarised in Sect. 3.1 with a specific example given in Sect. 3.2. The properties of GPs are discussed in Sect. 3.3

2.1 Model Overview

In a probabilistic framework, the in situ observations and climate model output are treated as realisations from latent spatiotemporal stochastic processes, denoted as $\{Y(s,t) : s \in \mathcal{S}, t \in \mathcal{T}\}$ and $\{Z(s,t) : s \in \mathcal{S}, t \in \mathcal{T}\}$, respectively. Stochastic processes are sequences of random variables indexed by a set, which in this case are the spatial and temporal coordinates in the domain $(\mathcal{S},\mathcal{T})$. A random variable is attributed to each spatiotemporal coordinate $(Y(s,t), Z(s,t))$. The data observed The observed data is then considered a realisation of the joint distribution over a finite set of random variables across the domain.

For the purpose of bias predictionevaluating the time-independent component of the climate model bias, the random variables are treated as independent and identically distributed across time, such that . That is the collection of temporal data for a given location s, $Y(s,t) \mid \boldsymbol{\phi}_Y(s) \overset{i.i.d.}{\sim} \mathcal{F}_Y(\boldsymbol{\phi}_Y(s))$ and $Z(s,t) \mid \boldsymbol{\phi}_Z(s) \overset{i.i.d.}{\sim} \mathcal{F}_Z(\boldsymbol{\phi}_Z(s))$, where $\mathcal{F}_Y(\cdot)$ and $\mathcal{F}_Z(\cdot)$ represent some generic site-level distributions with spatially varying vector parameters spatial location can be considered as multiple realisations from the same random variable. The random variables for each location are distributed respectively as $Y(s) \sim f_Y(\boldsymbol{\phi}_Y(s))$ and $Z(s) \sim f_Y(\boldsymbol{\phi}_Z(s))$, where $\boldsymbol{\phi}_Y(s)$ and $\boldsymbol{\phi}_Z(s)$ represent the collection of parameters that describe the PDF. For example, if the PDF is approximated as normal then $\boldsymbol{\phi}(s) = [\mu(s), \sigma(s)]$.The disparity between each of the PDF parameters for the in situ observations and climate model at each site then gives a measure of bias. The goal is to estimate the parameters $\boldsymbol{\phi}_Y(s)$ and $\boldsymbol{\phi}_Z(s)$. This follows from evaluating the time-independent component of the climate model bias. Consider evaluating bias in the values of January midday near-surface temperature over a region. While the values of nearby days are clearly dependent on each other, since focus is on evaluating time-independent bias, the time component of the data is dropped and only the marginal distribution considered. The marginal distribution in this case gives the probability of a certain value of January midday temperature just based on location and could for example be approximated as normal with mean and variance parameters, as mentioned in Sect. 2.1. In the case of other climatological fields such as rainfall a more appropriate distribution might be that of a Bernoulli-Gamma with its own collection of parameters, as used in Lima et al. (2021). Caution in this treatment should be applied in cases where, for example, the observational site only has a limited number of days of data and

these are bunched around the same relatively short time period, since this period is unlikely to be representative of the time series as a whole.

The disparity between the spatially varying parameters $\phi_Z(s)$ at the climate model grid points to quantify the bias and to apply quantile mapping to bias correct the climate model output. Gaussian processes are used to model the underlying spatial covariance structure of the parameters, which is required to estimate $\phi_Y(s)$ and $\phi_Z(s)$ in the site-level marginal distributions serves as a measure of bias. Specifically, as in Sect. 2.2, the bias for each parameter ϕ_i can be defined by some discrepancy function $\phi_{B,i}(s) = b_i(\phi_{Y,i}(s), \phi_{Z,i}(s))$. Alternatively, the parameters associated with the climate model output $\phi_{Z,i}(s)$ can be defined as a function of away from the location of the in situ observations. Further discussion around the definition of bias in climate models is provided in appendix **??**.

Consider a collection of n_Y in situ observational sites, where for each site i there exists m_i measurements of some property. In addition, consider gridded output from a climate model at n_z locations, where at each location there exists m_z measurements of the unbiased parameters $\phi_{Y,i}(s)$ and a latent biasfunction $\phi_{B,i}(s)$. In this paper an additive relationship is usedsame property. The data can then be represented through the following:

$$\boldsymbol{y} = [\boldsymbol{y}_{s_1}, \ldots, \boldsymbol{y}_{s_{n_y}}] \tag{1}$$

$$\boldsymbol{y}_{s_i} = [y_{s_i,1}, \ldots, y_{s_i,m_i}] \tag{2}$$

$$\boldsymbol{z} = [\boldsymbol{z}_{s'_1}, \ldots, \boldsymbol{z}_{s'_{n_z}}] \tag{3}$$

$$\boldsymbol{z}_{s'_i} = [z_{s'_i,1}, \ldots, z_{s'_i,m_z}] \tag{4}$$

Defining the collection of in situ observation sites as $\boldsymbol{s}_y = [s_1, \ldots, s_{n_y}]$ and the collection of climate model output locations as $\boldsymbol{s}_z = [s'_1, \ldots, s'_{n_z}]$, then the collection of PDF parameter values for each set of locations is written as:

$$\boldsymbol{\phi}_Y(\boldsymbol{s}_y) = [\boldsymbol{\phi}_Y(s_1), \ldots, \boldsymbol{\phi}_Y(s_{n_y})] \tag{5}$$

$$\boldsymbol{\phi}_Z(\boldsymbol{s}_z) = [\boldsymbol{\phi}_Z(s'_1), \ldots, \boldsymbol{\phi}_Z(s'_{n_z})] \tag{6}$$

The PDF parameters are each modelled as being generated from latent stochastic processes $\{\phi_Y(s)\}$ and $\{\phi_Z(s)\}$. The latent processes that generate the parameters for climate model are considered composed of two independent processes, one that also generates the equivalent parameters for the in situ observations and another that generates some bias, such that $\phi_{Z,i}(s) = \phi_{Y,i}(s) + \phi_{B,i}(s)$. Additionally, the bias $\phi_{B,i}(s)$ is considered independent of the value of $\phi_{Y,i}(s)$. To estimate the parameters across the domain and quantify the bias, these spatially varying parameters are modelled as spatial stochastic processes with $\{\phi_Z(s)\} = \{\phi_Y(s)\} + \{\phi_B(s)\}$. The family of GPs are chosen for the latent processes. A link function is used for the case where the parameter space is not the same as the sample space for GPs. Considering the case of no link function, the following can then be written:

$$\phi_Y(s) \sim \mathcal{GP}(\cdot,\cdot|\boldsymbol{\theta}_{\phi_Y}) \tag{7}$$

$$\phi_B(s) \sim \mathcal{GP}(\cdot,\cdot|\boldsymbol{\theta}_{\phi_B}) \tag{8}$$

$$\phi_Z(s) \sim \mathcal{GP}(\cdot,\cdot|\boldsymbol{\theta}_{\phi_Y},\boldsymbol{\theta}_{\phi_B}) \tag{9}$$

The collection of hyper-parameters ~~$\boldsymbol{\theta}$. It's important to note that since the collection of parameters may not necessarily all belong to the same parameter space, their representation can be standardized by applying a link function transformation to some of the parameters $\tilde{\phi}_i = h_i(\phi_i)$ so that all parameters can~~ for the generating processes are given by $\boldsymbol{\theta}_{\phi_Y}$ and $\boldsymbol{\theta}_{\phi_B}$ respectively. The hyper-parameters used in this paper consist of a mean constant, kernel variance and kernel lengthscale. Note the additive property of GPs allows $\phi_Z(s)$ to also be represented by ~~the same family of stochastic processes. In the methodology presented in this paper the family of Gaussian processes is used to model spatial dependencies. The full~~ a GP, where the mean and covariances are computed from the sum of the relative values from the independent processes. Further discussion around the properties of GPs is provided in appendix **??**. The hierarchical model is ~~then the following, with dependencies~~ illustrated through the plate diagram shown in Fig. 1.

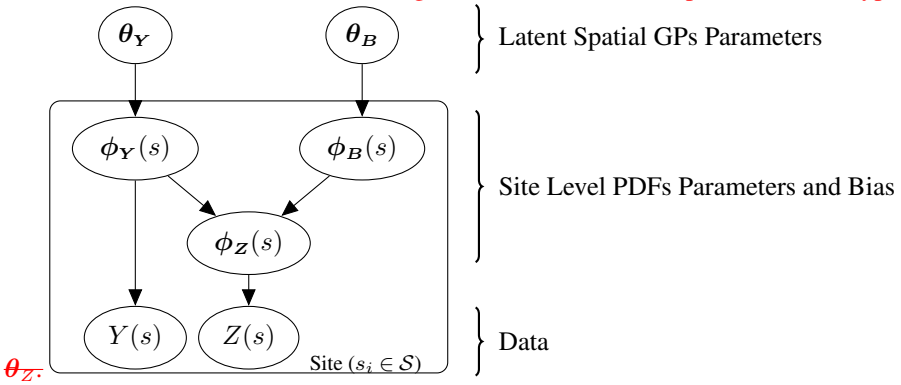
$$Y(s,t) \mid \boldsymbol{\phi}_Y(s) \overset{i.i.d.}{\sim} \mathcal{F}_Y(\boldsymbol{\phi}_Y(s))$$

$$Z(s,t) \mid \boldsymbol{\phi}_Z(s) \overset{i.i.d.}{\sim} \mathcal{F}_Z(\boldsymbol{\phi}_Z(s))$$

$$\begin{cases} \phi_{Z,i}(s) = \phi_{Y,i}(s) + \phi_{B,i}(s) & \text{if correct support,} \\ \phi_{Y,i}(s) \perp\!\!\!\perp \phi_{B,i}(s) \\ \phi_{Y,i}(s) \sim \mathcal{GP}(\cdot,\cdot|\boldsymbol{\theta}_{\phi_{Y,i}}) \\ \phi_{B,i}(s) \sim \mathcal{GP}(\cdot,\cdot|\boldsymbol{\theta}_{\phi_{B,i}}) \\ \\ \tilde{\phi}_{Z,i}(s) = \tilde{\phi}_{Y,i}(s) + \tilde{\phi}_{B,i}(s) & \text{if link function required for correct support.} \\ \tilde{\phi}_{Y,i}(s) \perp\!\!\!\perp \tilde{\phi}_{B,i}(s) \\ \tilde{\phi}_{Y,i}(s) \sim \mathcal{GP}(\cdot,\cdot|\boldsymbol{\theta}_{\tilde{\phi}_{Y,i}}) \\ \tilde{\phi}_{B,i}(s) \sim \mathcal{GP}(\cdot,\cdot|\boldsymbol{\theta}_{\tilde{\phi}_{B,i}}) \end{cases}$$

In addition, a specific example where the PDFs are approximated as normal is presented in appendix **??**.

~~Plate diagram showing a generic version of the full hierarchical model. The in-situ observations Y and climate model output Z are generated from distributions with the collection of parameters ϕ_Y and ϕ_Z respectively. The parameters ϕ_Z are modelled as some function of the parameters ϕ_Y and some independent bias ϕ_B. The parameters ϕ_Y and the corresponding bias ϕ_B are each themselves modelled over the domain as generated from Gaussian processes with hyper-parameters θ_Y and~~



~~θ_Z.~~

Figure 1. Plate diagram illustrating the full hierarchical model. The random variables for the in-situ observations $Y(s)$ and climate model output $Z(s)$ have PDFs with the collection of parameters $\phi_Y(s)$ and $\phi_Z(s)$ respectively, where $\phi_Z(s)$ is modelled as the sum of $\phi_Y(s)$ and some independent bias $\phi_B(s)$. The parameters $\phi_Y(s)$ and the corresponding bias $\phi_B(s)$ are each themselves modelled over the domain as generated from Gaussian processes with hyper-parameters $\boldsymbol{\theta}_Y$ and $\boldsymbol{\theta}_B$.

~~Gaussian processes naturally introduce spatial structure into the parameters and enable inference with misaligned data. Predictive estimates of the PDF parameters for each data source can be made for any set of locations across the domain. Estimates at the climate model output locations are needed for bias correction, while there's also the possibility to compute estimates at higher resolution and combine with a downscaling approach, as in Lima et al. (2021). Additional added benefits of GPs include properties that facilitate inference, for example the additive property where the sum of two independent GPs is itself also represented as a GP. More details following on from this and the application of GPs in the methodology is provided in Sect. 3.3.~~

Inference on the parameters of ~~site-level and spatial distributions of the~~ the hierarchical model given the data is applied in a Bayesian ~~hierarchical~~ framework, where parameters of the model are themselves treated as random variables with distributions. The distribution prior to conditioning on any data is known as the prior distribution and allows the incorporation of a domain specific expert's knowledge in the estimates of the parameters. The updated distribution after conditioning on the observed data is known as the posterior and is approximated using Markov chain Monte Carlo (MCMC) methods, which provide samples ~~from the posterior. An important advantage of this framework is it allows flexible extensions of the model while automatically maintaining robust uncertainty estimation. This results in the model being applicable to a wide range of problems and domains, especially important for correcting~~ of the parameters from the distribution $P(\phi_Y(\boldsymbol{s}_y), \phi_Z(\boldsymbol{s}_z), \boldsymbol{\theta}_{\phi_Y}, \boldsymbol{\theta}_{\phi_B} | \boldsymbol{y}, \boldsymbol{z})$. Estimates of the parameters ϕ_Y and ϕ_Z at any set of new locations $\hat{\boldsymbol{s}}$ can then be made by constructing the posterior predictive distribution, in particular for the purpose of bias correction estimates of ϕ_Y at the climate model locations can be made by sampling from the posterior predictive distribution of $P(\phi_Y(\boldsymbol{s}_z) | \boldsymbol{y}, \boldsymbol{z})$.

After obtaining multiple realisations of $\phi_Y(s_z)$ and $\phi_Z(s_z)$ quantile mapping is then used to bias correct the climate model time series at every grid cell location. Specifically, for each value of the time series from the climate model output ~~since there's a broad range of users interested in different variables and domains with varying levels of complexity~~at a given point $(z_{s'_i,j})$, this involves finding the percentile of that value using the parameters $\phi_Z(s'_i)$ and then mapping the value onto the corresponding value of the equivalent percentile of the PDF estimated for the unbiased process, defined through the parameters $\phi_Y(s'_i)$. The cumulative density function (CDF) returns the percentile of a given value and the inverse CDF returns the value corresponding to a given percentile, which results in the following correction function $\hat{z}_{s_i,j} = F^{-1}_{Y_{s_i}}(F_{Z_{s_i}}(z_{s_i,j}))$, where F represents the CDF at a specific site. The CDF can be estimated as an integral over the parametric form assumed for the PDF. The Bayesian hierarchical model presented provides a collection of realisations for $\phi_Y(s_z)$ and $\phi_Z(s_z)$ from an underlying latent distribution. Applying quantile mapping with each set of realisations then results in a collection of bias corrected time series, with an expectation and uncertainty. The full framework for bias correction proposed in this paper is then illustrated in Fig. 2. The formulation for the posterior and posterior predictive is given in appendix **??**.

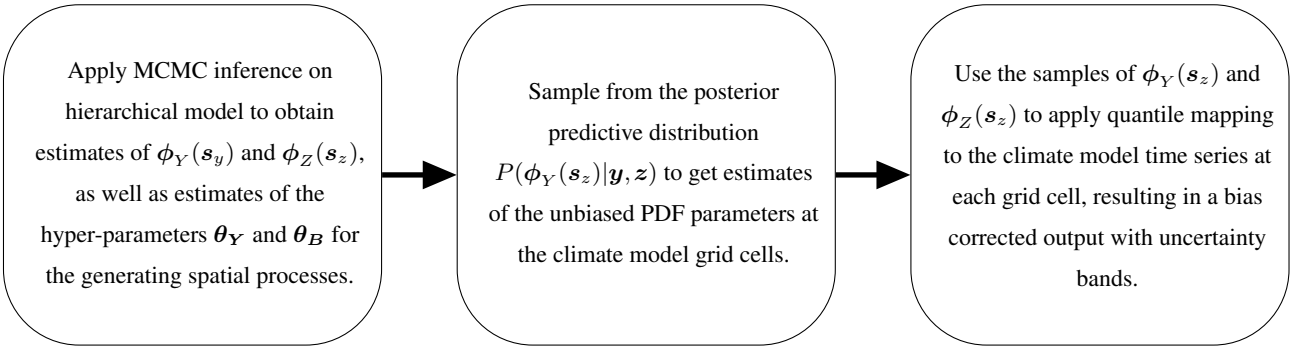


Figure 2. The full bias correction framework proposed in this paper broken down into the key steps.

3 ~~Simulated Examples~~Data Generation

The goal of the ~~model~~ hierarchical model in Fig. 1 is primarily to estimate, with reliable uncertainties, the true unbiased values of the PDF parameters at each location of the climate model output so bias correction can be applied. ~~The model additionally infers the spatial structure of these parameters and their bias. Results are presented to~~ Simulated examples are generated that highlight the advantage of two key features of the methodology over other approaches in the literature: modelling shared spatial covariance between the in situ data and climate model output through the inclusion of a shared generating latent process (Sect. ~~3.1~~3.1) and the Bayesian hierarchical ~~nature and~~ framework with uncertainty propagation (Sect. ~~3.2~~3.2). One dimensional simulated examples are chosen for clarity in illustrating these features, although it is noted the implementation works for higher dimensional domains as is useful in real-world scenarios. ~~The steps for generating the data and the results are presented separately for each example, while the discussion of results is done together in Sect. 5.~~

3.1 ~~Shared Latent Generating Processes: Non-Hierarchical Example~~

3.1 Non-Hierarchical Examples: Data Generation

~~A~~ To illustrate the potential advantage of modelling shared generating spatial processes, non-hierarchical ~~example is presented where direct~~ examples are generated for simplicity. Direct measurements are assumed for one parameter of the PDFs for the in situ observations ~~$\phi_Y(s)$~~ $\phi_Y(\boldsymbol{s}_y)$ and for the climate model output ~~$\phi_Z(s)$~~$\phi_Z(\boldsymbol{s}_z)$. The goal is to predict the ~~parameter $\phi_Y(s)$ across the spatial domain~~ unbiased parameter at the climate model locations $\phi_Y(\boldsymbol{s}_z)$ using information from both ~~the simulated in situ observations and climate model output~~sets of input data $\phi_Y(\boldsymbol{s}_y)$ and $\phi_Z(\boldsymbol{s}_z)$, which are related through ~~$\phi_Z = \phi_Y + \phi_B$.~~ ~~The parameters $\phi_Y(s)$ and $\phi_B(s)$ are considered independent and generated from Gaussian processes~~$\phi_Z(\boldsymbol{s}) = \phi_Y(\boldsymbol{s}) + \phi_B(\boldsymbol{s})$. Comparison is made to the approach of inferring ~~$\phi_Y(s)$ just~~ $\phi_Y(\boldsymbol{s}_z)$ from the in situ data alone $\phi_Y(\boldsymbol{s}_y)$, as in Lima et al. (2021). ~~The purpose of this example is to illustrate the advantage of modelling shared latent generating processes between the observational data and the climate model output, as in Fig. ??.~~ Relative performance is evaluated for three alternative simulated scenarios that correspond to different possible real-world situations.

~~The simulated data in this example~~ The data is generated assuming the ~~dependency~~ model in Fig. ~~4 and the relationship $\phi_Z = \phi_B + \phi_Y$, where ϕ_Y and ϕ_B are assumed independent. The latent Gaussian process distributions that generate ϕ_Y and ϕ_B across the domain~~ 1, where the GPs are taken with constant mean and an ~~RBF kernel. The hyper-parameters of these latent distributions~~ radial basis function (RBF) kernel with constant kernel length scale and kernel variance. The specific values of the hyper-parameters used to generate the data and the number of ~~simulated observations are set for three scenarios , as~~ observations under the different scenarios is given in Table ~~. The prior distributions~~ 1.

For each scenario, a sample of the parameters ~~are taken as the same for each scenario. Specifics of the data generation are given in Sect. of the appendix.~~

~~The three scenarios~~ $\phi_Y(\boldsymbol{s}^\star)$ and $\phi_B(\boldsymbol{s}^\star)$ is taken from the distributions $\mathcal{GP}_{\phi_Y}$ and $\mathcal{GP}_{\phi_B}$ at regularly spaced, high-resolution intervals. These samples are referred to here as complete realisations and represent underlying fields for each parameter across the domain. Direct 'observations' of the parameter $\phi_Y(\boldsymbol{s}_y)$ from the underlying field are simulated at lower-resolution, randomised locations after conditioning the distribution $\mathcal{GP}_{\phi_Y}$ on the complete realisation and introducing some noise. In order to simulate input data for the parameter $\phi_Z(\boldsymbol{s}_z)$ of the climate model output, samples are first generated for $\phi_Y(\boldsymbol{s}_z)$ and $\phi_B(\boldsymbol{s}_z)$ at regularly spaced intervals after conditioning the distributions $\mathcal{GP}_{\phi_Y}$ and $\mathcal{GP}_{\phi_B}$ on the complete realisations, then the sum of these samples at each location is taken to give $\phi_Z(\boldsymbol{s}_z)$. The input data $\phi_Y(\boldsymbol{s}_y)$ and $\phi_Z(\boldsymbol{s}_z)$ can be considered as the training set, while the underlying realisations generated for $\phi_Y(\boldsymbol{s}_z)$ are the test set used for validating the model performance.

Data is generated for three scenarios chosen to represent different potential real-world situations~~and the data generated for each is shown~~ , illustrated in Fig. 3. The first scenario (Fig. 3a) represents an example case where it is expected that there is ample data provided in the form of in situ observations to capture the features of the underlying complete realisation of ϕ_Y without significant added value provided from inclusion of the climate model output during inference. The second scenario (Fig. 3b) is an adjustment where the in situ observations are relatively sparse and the underlying bias is relatively smooth. In this situation the climate model output should provide significant added value in estimating ϕ_Y across the domain since it is

only afflicted by a comparatively simple bias that is easy to estimate. The final scenario (Fig. 3c) also involves sparse in situ observational data but with a reduced smoothness of the bias compared to the other scenarios. In this scenario the climate model output should provide added value in estimating ϕ_Y across the domain but this will be limited compared to scenario two due to the difficulty of disaggregating the components and estimating the comparatively more complex bias.

Dependent Variable	Model Parameters	Scenario 1	Scenario 2	Scenario 3
~~In-Situ~~	Kernel Variance (v_{ϕ_Y})	1.0	1.0	1.0
~~In-Situ~~ Unbiased PDF Parameter ϕ_Y	Kernel Lengthscale (l_{ϕ_Y})	3.0	3.0	3.0
~~In-Situ~~	Mean Constant (m_{ϕ_Y})	1.0	1.0	1.0
~~In-Situ Observation~~	Noise (σ_{ϕ_Y})	0.1	0.1	0.1
~~Bias~~	# Observations	80.0	20.0	20.0
	Kernel Variance (v_{ϕ_B})	1.0	1.0	1.0
~~Bias~~ Bias PDF Parameter ϕ_B	Kernel Lengthscale (l_{ϕ_B})	10.0	20.0	5.0
~~Bias~~	Mean Constant (m_{ϕ_B})	-1.0	-1.0	-1.0
~~# In-Situ Observations 80.0 20.0~~	~~20.0~~ # ~~Climate Model Predictions~~ Observations	100.0	80.0	80.0
Climate Model PDF Parameter ϕ_Z				

Table 1. A table showing the hyper-parameters of the two latent Gaussian processes used to generate the complete underlying realisations of ~~ϕ_Y~~ $\phi_Y(\boldsymbol{s}^\star)$, ~~$\phi_B$~~ $\phi_B(\boldsymbol{s}^\star)$ and hence ~~ϕ_Z~~$\phi_Z(\boldsymbol{s}^\star)$, as well as observations of ~~ϕ_Y~~ $\phi_Y(\boldsymbol{s}_y)$ and ~~ϕ_Z~~$\phi_Z(\boldsymbol{s}_z)$, on which inference is done for three scenarios. The number of observations representing in-situ data and climate model output is also given.

3.2 ~~Bayesian Framework: Hierarchical Example~~

~~A hierarchical example is presented in this section where the in situ data and climate model output are simulated at each site as generated from normal distributions, as in the specific example given in Sect. . The goal of the model is the same as in Sect. , that is to predict the parameters of the PDFs for the climate model output and in situ observations at the locations of the climate model output. An example of how uncertainty in these predictions can be propagated through bias correction techniques such as quantile mapping is then presented. The purpose of this section is to demonstrate the model working in the intended hierarchical structure and to illustrate the benefit of having a fully Bayesian hierarchical model for uncertainty estimation.~~

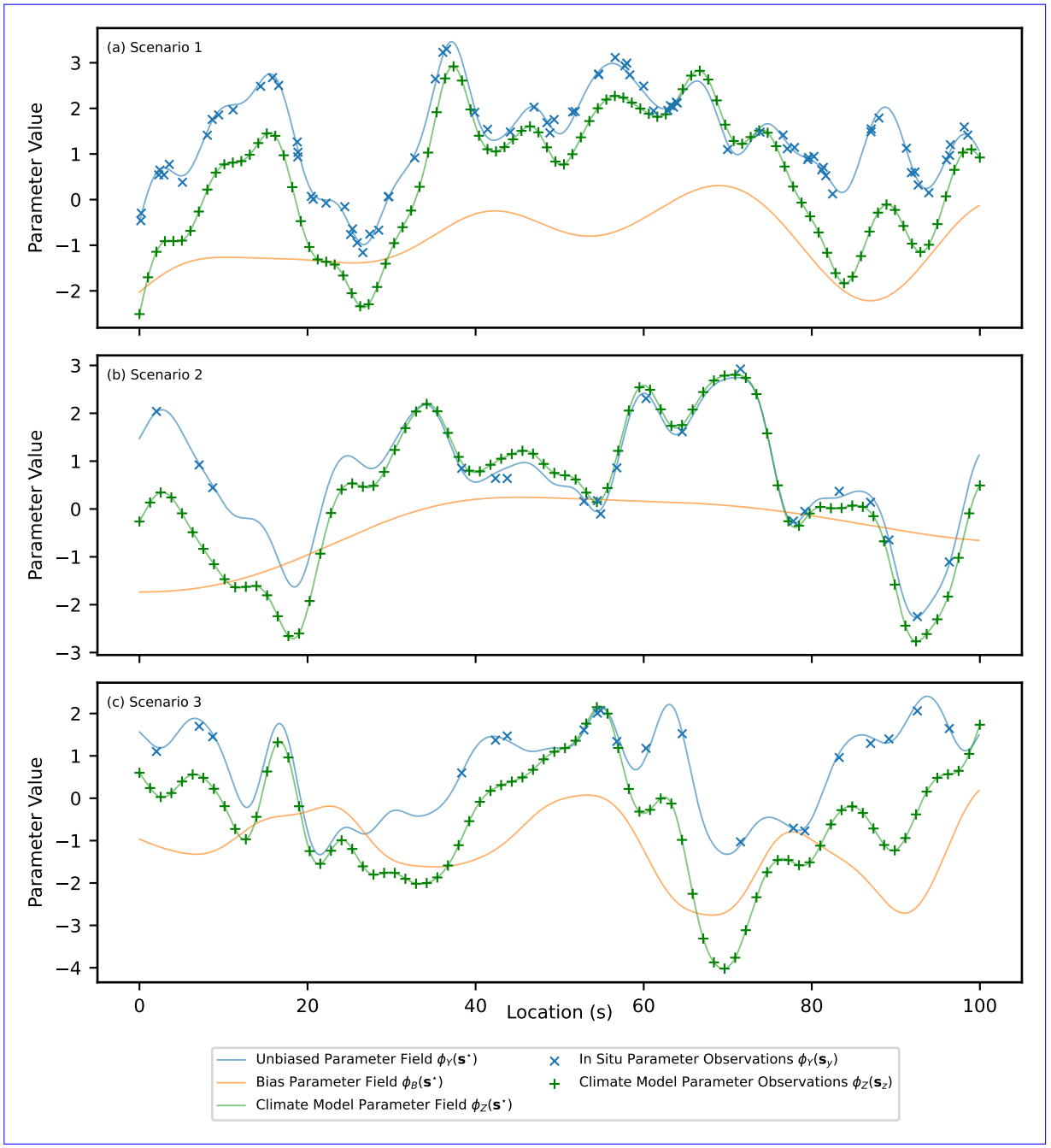


Figure 3. A figure showing simulated observed data for the PDF parameters $\phi_Y(\boldsymbol{s}_y)$ and $\phi_Z(\boldsymbol{s}_z)$, as well as the underlying complete realisations for each parameter and the underlying bias ($\phi_Y(\boldsymbol{s}^\star)$, $\phi_Z(\boldsymbol{s}^\star)$ and $\phi_B(\boldsymbol{s}^\star)$). Three scenarios are shown and correspond to data generated from parameters in Table 1.

~~The simulated data in this~~

3.2 Hierarchical Example: Data Generation

Following on from the non-hierarchical examples, to illustrate the advantage of uncertainty propagation in the Bayesian framework a hierarchical example is generated~~assuming the dependencies in Sect. and Fig. . Defining $Y(s,t)$ and $Z(s,t)$ as the in-situ~~ . In situ data and climate model output ~~respectively, then the time-independent PDF~~ are simulated at each site ~~is taken as normal~~ as generated from normal distributions, such that $Y(s) \sim \mathcal{N}(\mu_Y(s), \sigma_Y(s))$ and $Z(s) \sim \mathcal{N}(\mu_Y(s), \sigma_Y(s))$ as in appendix **??**. The following relationship is assumed for the mean parameters $\mu_Z(s) = \mu_Y(s) + \mu_B(s)$, where $\mu_B(s)$ is the bias in the mean for the climate data. For the standard deviation, the parameters are first transformed using a logarithmic link function and then the relationship $\tilde{\sigma}_Z(s) = \tilde{\sigma}_Y(s) + \tilde{\sigma}_B(s)$ is assumed, where $\tilde{\sigma}_B(s)$ is the bias in the transformed parameter. The latent distributions that generate $\mu_Y(s)$, $\mu_B(s)$, $\tilde{\sigma}_Y(s)$ and $\tilde{\sigma}_B(s)$ across the domain are assumed as independent GPs with constant mean and an RBF kernel. The hyper-parameters for these latent generating processes are set for a single scenario, as given in Table ~~. Further specifics of the data generation is provided in Sect. of the appendix~~2 along with the number of simulated observation locations and the number of samples per location.

A sample of the parameters $\mu_Y(\boldsymbol{s}^\star)$, $\mu_B(\boldsymbol{s}^\star)$, $\tilde{\sigma}_Y(\boldsymbol{s}^\star)$ and $\tilde{\sigma}_B(\boldsymbol{s}^\star)$ is taken from the distributions $\mathcal{GP}_{\mu_Y}$, $\mathcal{GP}_{\mu_B}$, $\mathcal{GP}_{\tilde{\sigma}_Y}$ and $\mathcal{GP}_{\tilde{\sigma}_B}$ at regularly spaced, high resolution intervals. These samples are referred to as complete realisations and represent the underlying fields for each PDF parameter across the domain.

After conditioning the latent GPs on the complete realisations a sample is generated for $\mu_Y(\boldsymbol{s}_y)$ and $\tilde{\sigma}_Y(\boldsymbol{s}_y)$ at a selection of lower-resolution, randomised locations that represent simulated in situ observation sites. Multiple observations of $Y(s_i)$ are then generated at each in situ observation site by sampling from the corresponding normal distribution $\mathcal{N}(\mu_Y(s_i), \tilde{\sigma}_Y(s_i))$. In the case of the simulated climate model output, samples are first generated for $\mu_Y(\boldsymbol{s}_z)$, $\mu_B(\boldsymbol{s}_z)$, $\tilde{\sigma}_Y(\boldsymbol{s}_z)$ and $\tilde{\sigma}_B(\boldsymbol{s}_z)$ at regularly spaced intervals after conditioning the latent distributions on the complete realisations, then the sum of these samples at each location is taken to give $\mu_Z(\boldsymbol{s}_z)$ and $\tilde{\sigma}_Z(\boldsymbol{s}_z)$. The climate model output is then generated at each of these locations from the corresponding normal distribution $Z(s_i) \sim \mathcal{N}(\mu_Z(s_i), \tilde{\sigma}_Z(s_i))$.

~~There~~ In the generated example there are 40 locations corresponding to simulated in situ observation sites, where for each site 20 measurements are generated. ~~Likewise, there~~ There are 80 locations corresponding to simulated climate model output grid points and at each location 100 samples are generated. This reflects the typical scenario where the climate model output has greater spatiotemporal coverage than in situ observations but is ~~also afflicted with greater~~ afflicted with bias. In Fig. 4 examples of the generated samples of $Y(s_i)$ and $Z(s_i)$ are shown corresponding to the nearest sites for three locations. It is clear that, due to limited observations, there will be significant uncertainty in estimates of the mean and standard deviation parameters at each site and it's important this uncertainty is propagated ~~when estimating the parameters across the domain~~to inference of the hyper-parameters for the latent GPs and also to estimates of the unbiased PDF parameters at the climate model locations ($\mu_Y(\boldsymbol{s}_z)$ and $\tilde{\sigma}_Y(\boldsymbol{s}_z)$). The underlying, complete realisations of the parameters ~~$\mu_Y(s)$, $\mu_Z(s)$, $\sigma_Y(s)$ and $\sigma_Z(s)$~~$\mu_Y(\boldsymbol{s}^\star)$, $\mu_Z(\boldsymbol{s}^\star)$, $\sigma_Y(\boldsymbol{s}^\star)$ and $\sigma_Z(\boldsymbol{s}^\star)$, as well as the bias ~~$\mu_B(s)$ and $\sigma_B(s)$~~$\mu_B(\boldsymbol{s}^\star)$ and $\sigma_B(\boldsymbol{s}^\star)$, are shown in Fig. 5. In addition, the empirical mean value and standard deviation of the generated data is ~~given~~ illustrated at the simulated in situ observation and climate model sites. The goal of the hierarchical model is then to predict the unbiased values for the parameters of the PDFs at the

locations of the climate model output ($\mu_Y(\boldsymbol{s}_z)$ and $\tilde{\sigma}_Y(\boldsymbol{s}_z)$), while propagating uncertainty. An example of how the uncertainty in predictions of $\mu_Y(\boldsymbol{s}_z)$ and $\tilde{\sigma}_Y(\boldsymbol{s}_z)$ is propagated through quantile mapping is then provided.

Dependent Variable	Model Parameters	Hierarchical Scenario
~~In-Situ Mean,~~ Unbiased PDF Mean μ_Y	Kernel Variance (v_{μ_Y})	1.0
	Kernel Lengthscale (l_{μ_Y})	3.0
~~In-Situ Mean,~~	Mean Constant (m_{μ_Y})	1.0
~~In-Situ Transformed Variance,~~ Unbiased PDF Transformed Variance $\tilde{\sigma}^2_Y$	Kernel Variance ($v_{\tilde{\sigma}^2_Y}$)	1.0
	Kernel Lengthscale ($l_{\tilde{\sigma}^2_Y}$)	3.0
~~In-Situ Transformed Variance,~~	Mean Constant ($m_{\tilde{\sigma}^2_Y}$)	1.0
~~Bias Mean,~~ Bias PDF Mean μ_B	Kernel Variance (v_{μ_B})	1.0
	Kernel Lengthscale (l_{μ_B})	10.0
~~Bias Mean,~~	Mean Constant (m_{μ_B})	-1.0
~~Bias Transformed Variance,~~ Bias PDF Transformed Variance $\tilde{\sigma}^2_B$	Kernel Variance ($v_{\tilde{\sigma}^2_B}$)	1.0
	Kernel Lengthscale ($l_{\tilde{\sigma}^2_B}$)	10.0
~~Bias Transformed Variance,~~	Mean Constant ($m_{\tilde{\sigma}^2_B}$)	-1.0
Unbiased Output Y	# ~~Spatial Locations of In-Situ Observations~~ Observation Sites	40.0
	# ~~Spatial Locations of Climate Model Predictions~~ Observations per Site	~~80.0~~ 20.0
Climate Model Output Z	# ~~Samples per Location of In-Situ Observations~~ Observation Sites	~~20.0~~ 80.0
	# ~~Samples per Location of Climate Model Predictions~~ Observations per Site	100.0

Table 2. A table showing the hyper-parameters used to generate the complete underlying realisations and the measurement data on which inference is done for the hierarchical scenario. The number of sites where data is generated along with the number of samples for each site is also given.

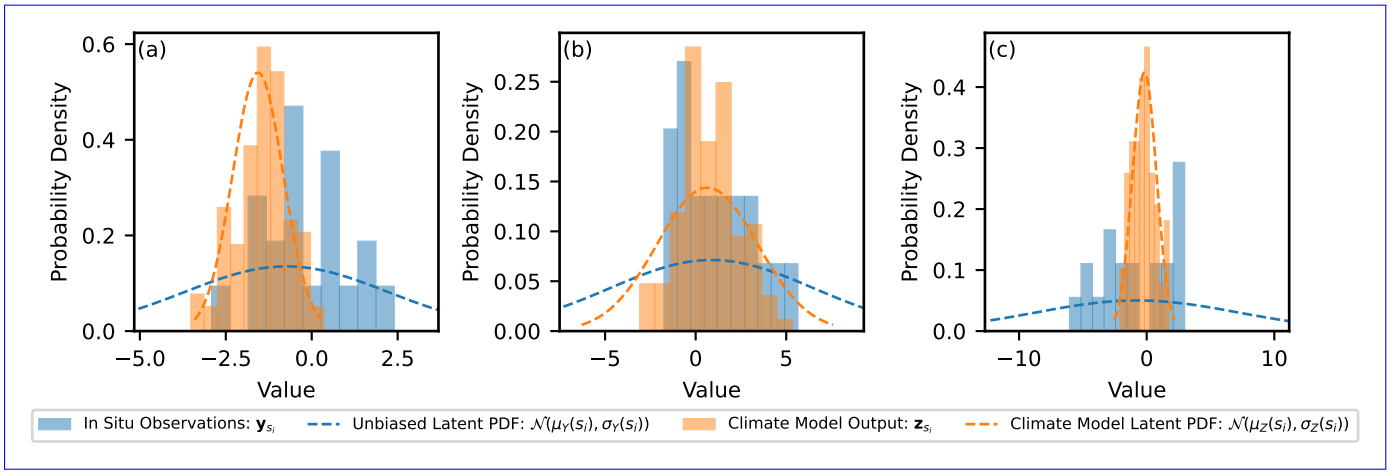


Figure 4. Histograms for the climate model output at three locations and the corresponding data from the nearest in situ observation site. The locations are a) s=11.4, b) s=46.8 and c) s=79.7. The latent normal distribution the data was generated from is illustrated as a dotted line.

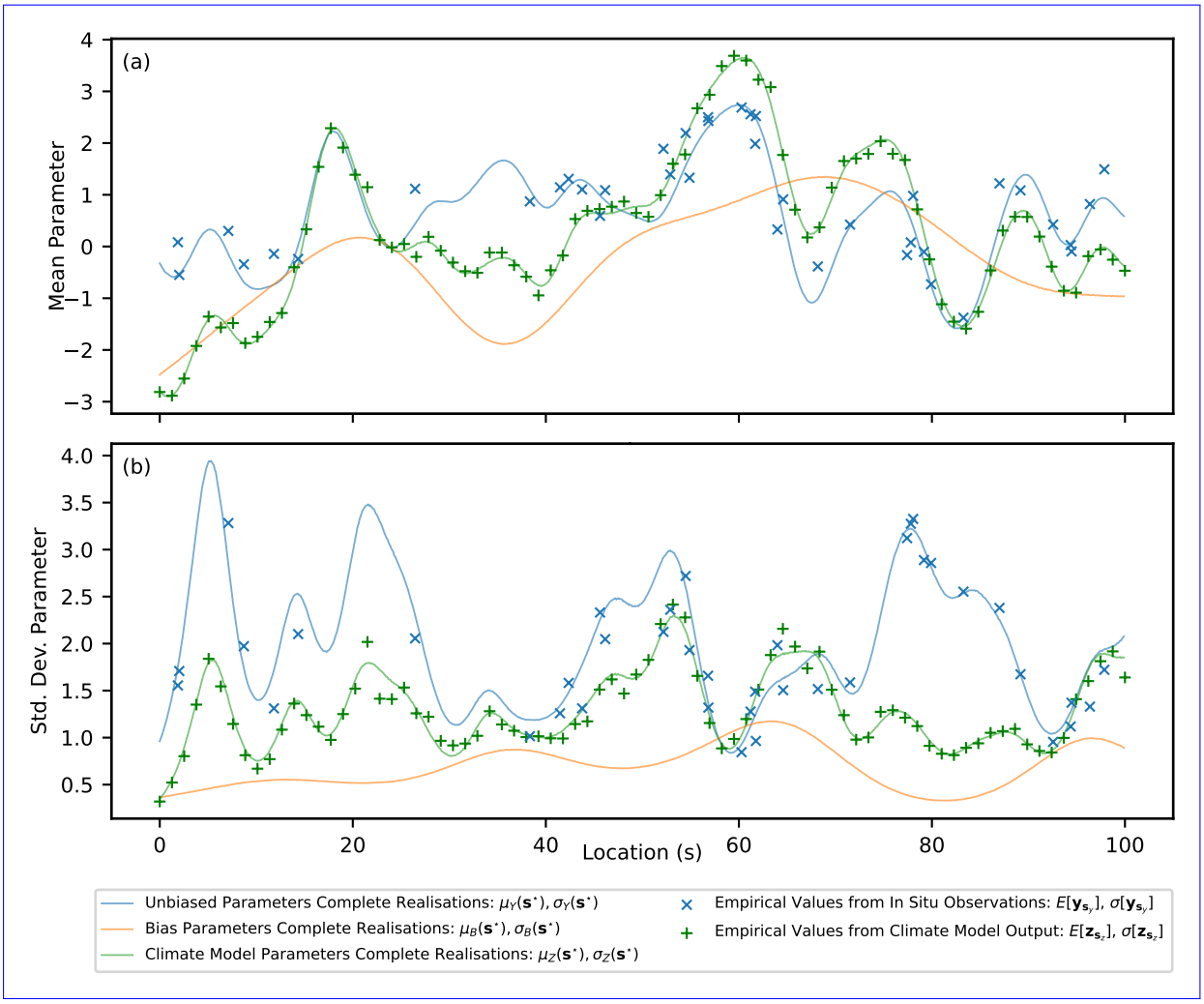


Figure 5. Simulated complete realisations for the parameters $\mu_Y(\boldsymbol{s}^\star)$, $\mu_B(\boldsymbol{s}^\star)$, $\mu_Z(\boldsymbol{s}^\star)$, $\tilde{\sigma}_Y(\boldsymbol{s}^\star)$, $\tilde{\sigma}_B(\boldsymbol{s}^\star)$ and $\tilde{\sigma}_Z(\boldsymbol{s}^\star)$ as well as the empirical values at the observation locations for the in situ and climate model data.

4 Results

~~Inference of the parameters of the models~~ Inference is done in a Bayesian framework using MCMC and the No-U-Turn Sampler (NUTS) algorithm (Hoffman and Gelman, 2014) implemented in Numpyro (Phan et al., 2019). The parameters/hyper-parameters are treated as random variables with associated probability distributions. A prior distribution is set for each ~~parameter~~ hyper-parameter and represents the belief on the distribution before observing any data, which typically incorporates knowledge from application specific experts. In the examples presented, relatively non-informative priors are chosen since the data is simulated and represents generic examples. The posterior distribution of each parameter/hyper-parameter is the updated distribution after

observing and conditioning on the data. Estimates of the ~~parameters $\phi_Y(s)$, $\phi_Z(s)$~~ PDF parameters $\phi_Y(\hat{s})$, $\phi_Z(\hat{s})$ and the corresponding bias ~~$\phi_B(s)$ across the domain given the posterior and the observed data is~~ $\phi_B(\hat{s})$ at new locations away from the observation sites, are then referred to as samples from the posterior predictive.

4.1 Non-Hierarchical Examples: Results

~~The expectation, standard deviation and 95% credible intervals for the prior distribution and posterior distribution after inference of each parameter under the three different scenariosis given in Table . Comparisons are shown in the statistics between the posterior distributions of the full model presented in this paper, referred to as the shared process model,~~ The shared latent process model presented in this paper is fit to the three non-hierarchical example scenarios, as discussed in Sect. 3.1. Input data for $\phi_Y(\boldsymbol{s}_y)$ and ~~the case where only the parameters for~~ $\phi_Z(\boldsymbol{s}_z)$ are provided and the hyper-parameters for the latent GPs that generate the unbiased and biased components inferred. Comparisons in estimates of the hyper-parameters for the unbiased process (m_{ϕ_X}, v_{ϕ_X} and l_{ϕ_X}) are made to the approach of only fitting to the in situ data~~are modelled as generated from a latent Gaussian processes~~, referred to here as the single process ~~model. In Sect. of the appendix, an illustration is given of~~ approach since the latent process generating the bias is not modelled. The difference between the shared and single process approaches is detailed further in appendix **??**. The expectation, standard deviation and 95% credible intervals for the prior and posterior distributions of ~~each parameter after inference with the shared process model for scenario one~~the hyper-parameters under the three different scenarios is given in Table 3.

Under all scenarios ~~and for both the shared process and single process models~~ the 95% credible interval of the posterior for every hyper-parameter bounds the value specified in generating the data. The expectation for the posterior distribution of the ~~shared process model~~ unbiased hyper-parameters is in general closer to the specified value ~~than~~ in the shared process model compared with the single process model and the range of the credible interval is smaller. In scenario one and three the differences between the shared and single process models posteriors are relatively insignificant for the mean constant (m_{ϕ_X}) and kernel variance (v_{ϕ_X}), although the shared process model ~~does show a~~ shows a noticeable reduction in the uncertainty of the ~~length scale for the latent process generating ϕ_Y~~kernel length scale (l_{ϕ_X}). In scenario two the difference is more significant and clear improvement is shown ~~in both the expectation and uncertainty of latent parameter estimates~~ for the shared process model~~. Improvement is also clear in estimates from the 3rd scenario, although the relative difference in performance between models is less significant~~, both in the expectation and uncertainty of hyper-parameter estimates.

~~Predictions for the underlying fields of the parametersϕ_Y(s), $\phi_Z(s)$ and the corresponding bias $\phi_B(s)$ across the domain given~~ After applying MCMC inference on the parameters/hyper-parameters that generate the data, ~~referred to as the posterior predictive , are shown~~ posterior predictive estimates are made for the unbiased PDF parameter values at the simulated locations of the climate model ($\phi_Y(\boldsymbol{s}_z)$). These estimates are presented in Fig. 6 for each scenario and for both the shared and single process models. ~~The true underlying fields that the simulated observations were sampled from is also shown. The single process model is only concerned with estimating the underlying field of $\phi_Y(s)$ across the domain given observations of the parameter for the in situ data, so in Figs. a, c and e the climate model output and bias fields are excluded. To perform bias correction of the climate modeloutput through quantile mapping, posterior predictive estimates of $\phi_Y(s)$ at the climate model output~~

~~locations are required~~Additionally, estimates of the underlying bias $\phi_B(s)$ are shown for the shared process model, since the bias is explicitly modelled. The relative ~~ability~~ performance of the shared and single process models ~~to estimate this is further assessed through~~ is quantified by computing R^2 scores ~~,~~ between the predictions of $\phi_Y(\boldsymbol{s}_z)$ and the actual values used in generating the data, with results presented in Table 4.

In Fig. 6 it can be seen that the predictions of ~~$\phi_Y(s)$ across the domain~~ $\phi_Y(\boldsymbol{s}_z)$ in the shared process case (Fig. ~~b, dand~~ 6d, 6e and 6f) are closer to the true underlying field and with smaller but still realistic uncertainty compared to the single process model. In scenario one, the difference between the posterior predictive distributions for ~~$\phi_Y(s)$ across the domain~~ $\phi_Y(\boldsymbol{s}_z)$ between the two approaches is not substantial, with both models performing adequately, having R^2 scores of 0.99 and 0.97 respectively. In scenario two, the difference between estimates of ~~$\phi_Y(s)$~~ $\phi_Y(\boldsymbol{s}_z)$ between the models is significant with R^2 scores of 0.99 and 0.68 for the shared and single process models respectively. Finally, in scenario three ~~there is again a significant difference in the estimates of $\phi_Y(s)$ between the models,~~ the difference is again significant with R^2 scores of 0.74 and 0.52 respectively, although ~~the difference is reduced~~ less significant compared with scenario two.

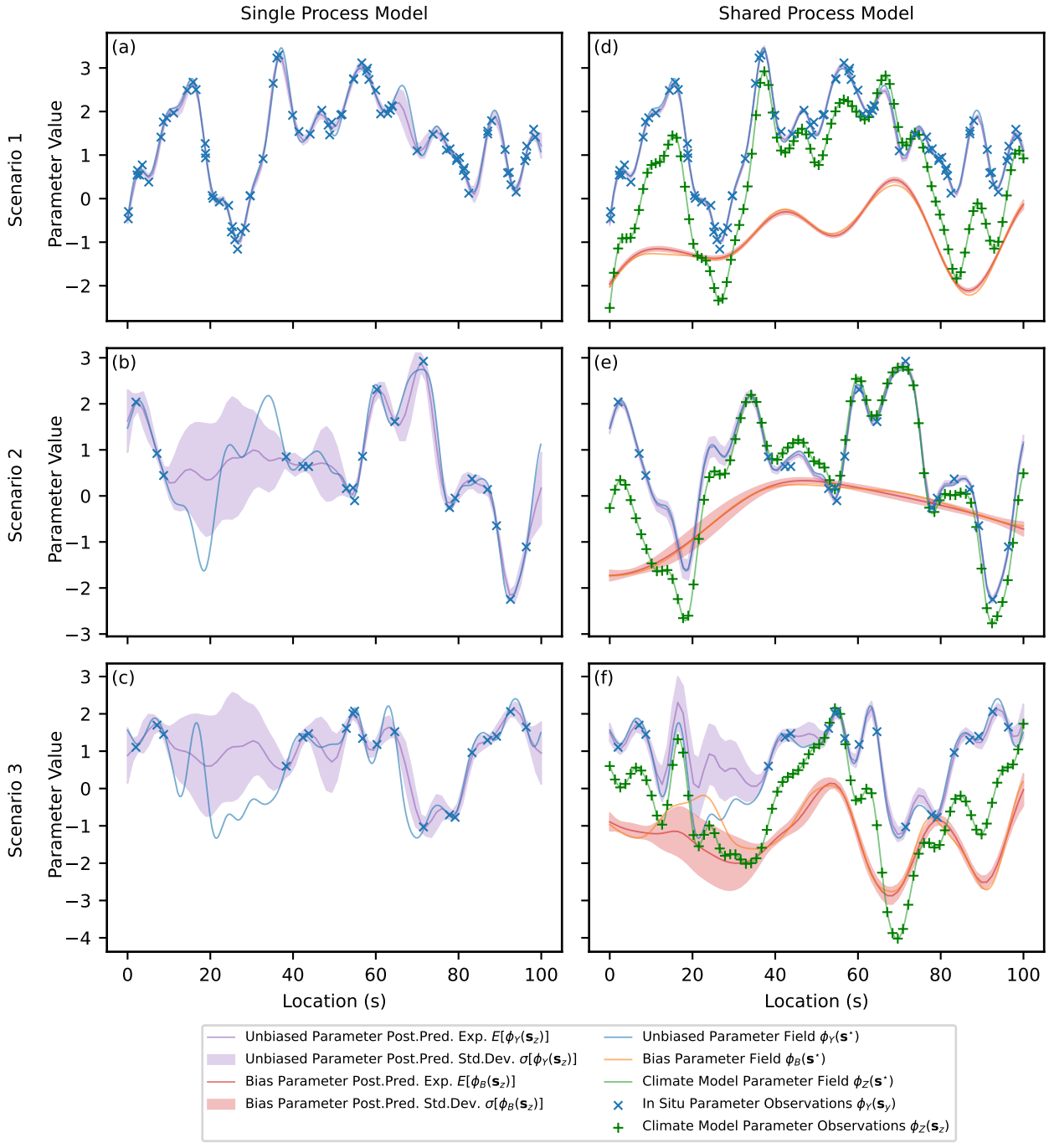


Figure 6. Expectation and 1σ uncertainty of the posterior predictive distributions of the parameter $\phi_Y(\boldsymbol{s}_z)$ and the corresponding bias $\phi_B(\boldsymbol{s}_z)$ for three scenarios. The underlying functions (complete realisations) as well as the simulated input data are also shown.

4.2 Hierarchical Example

The model presented in this paper is fit to the hierarchical example from Sect. 3.2. The expectation, standard deviation and 95% credible intervals for the prior and posterior distributions of each ~~parameter~~ hyper-parameter of the latent generating processes are given in Table 5. The 95% credible interval of the posterior for every hyper-parameter bounds the value specified in generating the data. As expected the posterior distribution for each ~~parameter~~ hyper-parameter is concentrated closer to the value specified when generating the data than the relatively non-informative prior distributions. The prior and posterior distributions for each ~~parameter~~ hyper-parameter are plot in Fig. **??** of the appendix.

PDF Parameter	Model Hyper-Parameter	Specified Value	Prior Distribution Exp.	Prior Distribution Std. Dev.	Prior Distribution C.I. L.	Prior Distribution C.I. U.	Posterior Dist. Exp.	Posterior Dist. Std. Dev.	Posterior Dist. C.I. L.	Posterior Dist. C.I. U.
Unbiased Mean μ_Y	Kernel Variance v_{μ_Y}	1.0	0.67	0.67	0.02	2.46	1.00	0.32	0.49	1.63
	Kernel Lengthscale l_{μ_Y}	3.0	15.00	8.66	3.09	36.12	3.00	0.22	2.56	3.43
	Mean Constant m_{μ_Y}	1.0	0.00	2.00	-3.92	3.92	0.73	0.28	0.17	1.26
Unbiased Transformed Variance $\tilde{\sigma}_Y$	Kernel Variance $v_{\tilde{\sigma}^2_Y}$	1.0	0.67	0.67	0.02	2.46	0.70	0.25	0.30	1.17
	Kernel Lengthscale $l_{\tilde{\sigma}^2_Y}$	3.0	15.00	8.66	3.09	36.12	2.94	0.24	2.47	3.40
	Mean Constant $m_{\tilde{\sigma}^2_Y}$	1.0	0.00	2.00	-3.92	3.92	1.12	0.24	0.66	1.61
Bias Mean μ_B	Kernel Variance v_{μ_B}	1.0	0.67	0.67	0.02	2.46	1.38	0.63	0.42	2.58
	Kernel Lengthscale l_{μ_B}	10.0	15.00	8.66	3.09	36.12	12.02	3.59	5.08	18.50
	Mean Constant m_{μ_B}	-1.0	0.00	2.00	-3.92	3.92	-0.78	0.56	-1.89	0.29
Bias Transformed Variance $\tilde{\sigma}_B$	Kernel Variance $v_{\tilde{\sigma}^2_B}$	1.0	0.67	0.67	0.02	2.46	0.92	0.48	0.24	1.86
	Kernel Lengthscale $l_{\tilde{\sigma}^2_B}$	10.0	15.00	8.66	3.09	36.12	8.97	1.96	5.07	12.58
	Mean Constant $m_{\tilde{\sigma}^2_B}$	-1.0	0.00	2.00	-3.92	3.92	-0.86	0.42	-1.73	-0.06

Table 5. A table showing summary statistics for the prior and posterior distributions including the expectation (Exp.), standard deviation (Std. Dev.) and lower and upper bounds for the 95% credible interval (C.I. L. and C.I. U.). The specified value for each hyper-parameter used to generate the data is also shown.

~~The posterior predictive estimate for the underlying fields of $\mu_Y(s)$, $\mu_B(s)$, $\sigma_Y(s)$ and $\sigma_B(s)$ across the domain given the data is~~ After fitting the model, posterior predictive estimates are made of the unbiased mean and standard deviation parameters at the simulated locations of the climate model ($\mu_Y(\boldsymbol{s}_z)$ and $\sigma_Y(\boldsymbol{s}_z)$). Additionally, estimates of the bias in the parameters are made ($\mu_B(\boldsymbol{s}_z)$ and $\sigma_B(\boldsymbol{s}_z)$). These estimates along with the true underlying values are shown in Fig. ~~. The true underlying fields of the parameters are also shown, as are the~~ 7. The empirical mean and standard deviation ~~values of the samples of simulated in situ observations and climate model outputs~~ of the input data is also given at the locations where they are sampled. The ~~posterior predictive appears~~ estimates visually appear to perform well at capturing the spatial features of the underlying fields ~~while also exhibiting a reasonable~~ and at estimating a one sigma uncertainty range~~that bounds the majority of the underlying function~~. For example, in the range of ~~$s \in [15, 25]$~~ $s \in [28, 38]$, where the main data source is the biased climate model output,

the prediction accurately captures the spatial features of the ~~true, unobserved latent mean~~ unbiased parameters ($\mu_Y(s)$ and ~~standard deviation~~ $\sigma_Y(s)$~~. Uncertainty in the parameters of $\mu_Y(s)$~~) with an uncertainty that bounds the true underlying value over most of the region. Additionally, uncertainty in the unbiased parameters at the in situ observation sites ($\mu_Y(\boldsymbol{s}_y)$ and ~~$\sigma_Y(s)$ at the observation sites,~~ $\sigma_Y(\boldsymbol{s}_y)$) due to limited samples ~~, is propagated through the model. This is~~ is clearly reflected in the ~~uncertainty shown in estimatesof the posterior predictive at the observation sites~~estimates.

~~Bias correction of samples from~~

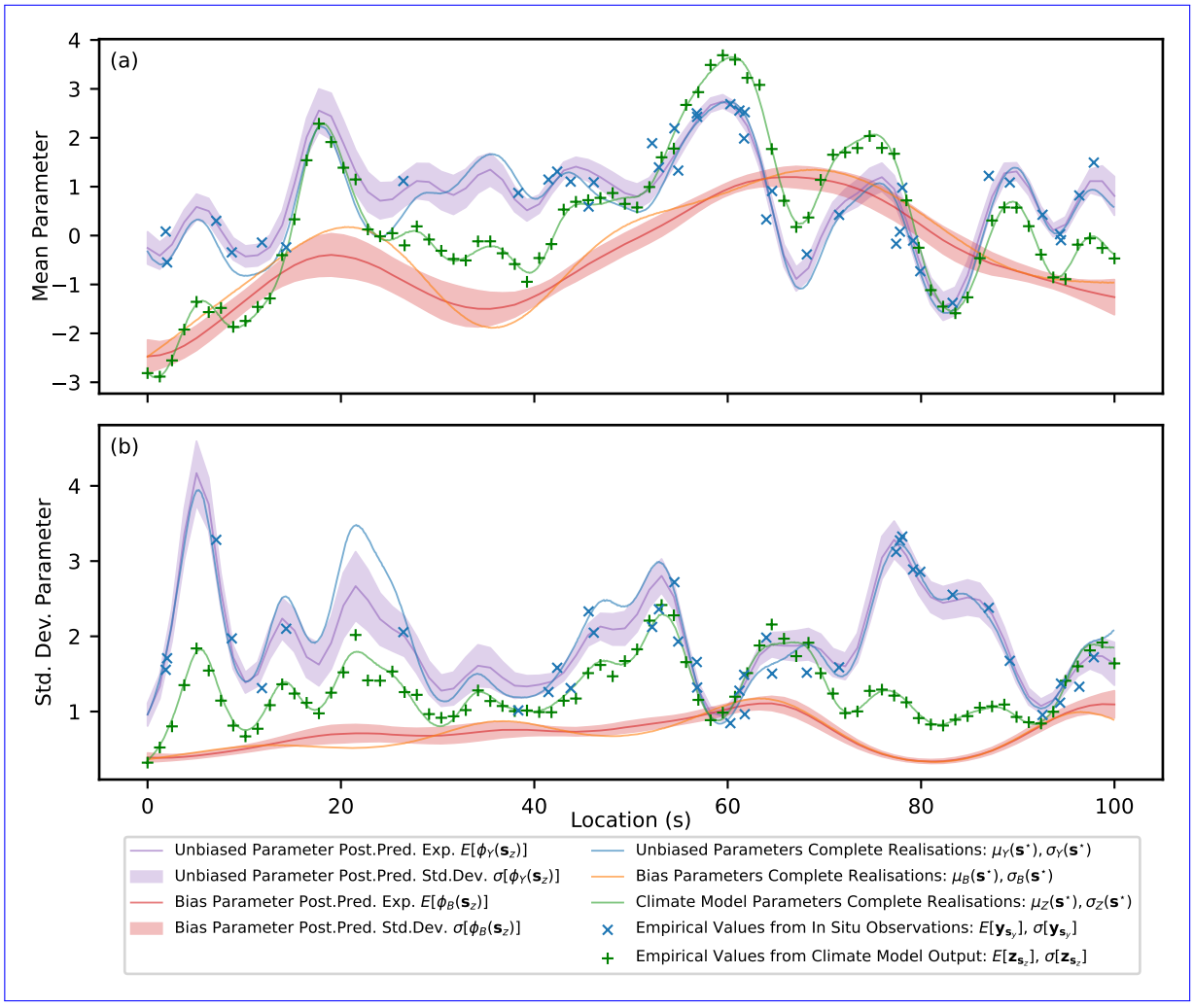


Figure 7. A figure showing the expectation and one sigma uncertainty of the posterior predictive distribution across the domain for the parameters $\mu_Y(\boldsymbol{s}_z)$, $\mu_B(\boldsymbol{s}_z)$, $\sigma_Y(\boldsymbol{s}_z)$ and $\sigma_B(\boldsymbol{s}_z)$ as well as the true underlying functions.

Quantile mapping is applied to the climate model output for a single site ($\boldsymbol{z}_{s_z}$) and the bias corrected time series ($\hat{\boldsymbol{z}}_{s_z}$) is shown in Fig. 8. The site chosen is at $s = 11.4$ and is the same as in Fig. 4a. A generic time series for the climate model

output and nearest in situ observations is generated from the correct mean and standard deviations of the samples. Quantile mapping ~~of the climate model time series~~ is performed for each posterior predictive realisation of ~~$\mu_Y(s)$, $\mu_Z(s)$, $\sigma_Y(s)$ and $\sigma_Z(s)$~~$\mu_Y(s_i)$, $\mu_Z(s_i)$, $\sigma_Y(s_i)$ and $\sigma_Z(s_i)$. This results in multiple realisations of bias corrected time series with an expectation and uncertainty.

~~A figure showing the expectation and one sigma uncertainty of the posterior predictive distribution across the domain for the parameters $\mu_Y(s_z)$, $\mu_B(s_z)$, $\sigma_Y(s_z)$ and $\sigma_B(s_z)$ as well as the true underlying functions.~~

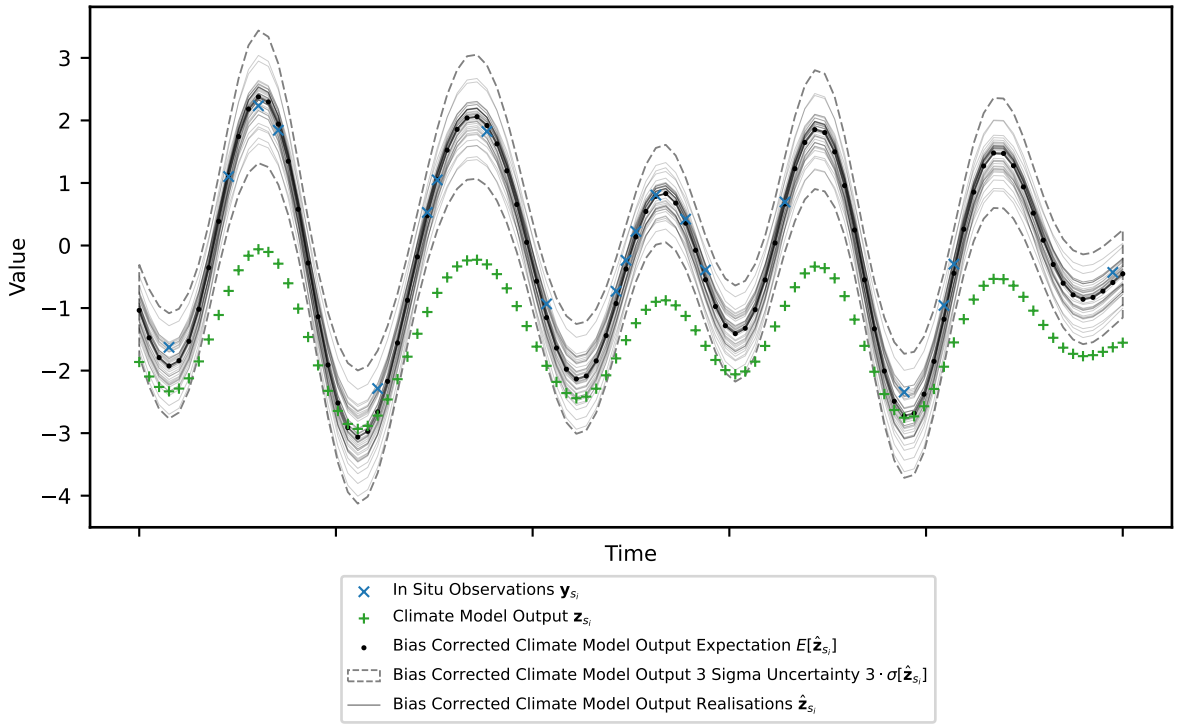


Figure 8. Simulated time series for the climate model output at location $s = 11.4$ and for the nearest in situ observation site. Realisations of the climate model bias corrected time series are shown along with the expectation and three sigma uncertainty range.

5 Discussion

The bias correction framework proposed in this paper models the parameters of the PDFs for the in situ observations and climate model output across the domain using a Bayesian hierarchical model. This allows estimates to be made of the unbiased PDF parameters at the climate model locations and quantile mapping can then be applied to bias correct the climate model time series. The hierarchical model uses GPs to model the spatial covariance structure of the ~~data~~ PDF parameters and assumes that each parameter of the ~~PDF for the~~ climate model output is generated from two independent ~~, latent GPs. One GP~~ GPs: One that generates an unbiased component and ~~so also the equivalent parameter for the PDF of~~ another that generates a bias. The

GP that generates the unbiased component is also modelled as generating the equivalent PDF parameters for the in situ data~~, while the other GP generates a bias. This~~ . This approach reflects the belief that the climate model provides skillful estimates of ~~these~~ the PDF parameters across the domain and that the spatial covariance structure, generated from equations based on established physical laws, has ~~similar features~~ spatial features similar to the true ~~underlying structure~~unbiased PDF parameter values. The climate model output, while afflicted with bias, ~~has~~ provides comprehensive spatiotemporal coverage and ~~provides~~ useful information in the inference of the unbiased PDF ~~values~~ parameters across the domain~~,~~. This is assuming the bias signal can be adequately deconstructed from the climate model output with the use of in situ observations. ~~Modelling shared latent processes provides added value over the approach where~~ In Sect. 4.1 of the results, the added value of modelling shared latent GPs between the in situ observations and climate model output is demonstrated. This is compared with the approach of modelling a latent GP for the in situ observations alone and inferring the unbiased PDF parameters ~~are inferred from the in situ observations alone~~without incorporating information from the climate model output, as in Lima et al. (2021) ~~. This is demonstrated in Sect. of the results, where~~ and here referred to as the single process approach.

The added value is assessed for three scenarios with differing density of observations and spatial complexity of the bias signal.

~~Added value is assessed with respect to :~~ The methods used to assess the added value include: comparisons of summary statistics for the posterior distributions of the GP hyper-parameters~~of the latent GPs; visual examination~~ ; visual comparisons of the expectation and standard deviation for posterior predictive estimates of the unbiased PDF parameters across the domain; and comparison of R^2 scores for the unbiased PDF parameter at the locations of the climate model output. It is shown that ~~most added value is provided,~~ across all these measures, ~~in the case of scenario two,~~ the most added value is provided where the in situ observations are ~~sparse compared to~~ relatively sparse compared with the climate model output and the underlying bias is relatively ~~smooth compared to~~ spatially smooth compared with the the unbiased signal~~. The bias can be estimated with high accuracy and precision~~, as in scenario two. In this scenario, despite sparse in situ observations, since ~~it~~ the bias signal varies smoothly across the domain ~~, which also means~~ the climate model output can be ~~disaggregated and the unbiased component estimated across the domain with high accuracy and precision.~~ accurately and precisely disaggregated into its unbiased and biased components. This leads to improved estimates of the unbiased PDF parameters at the climate model locations when considering shared GPs compared with the single process approach that uses in situ observations alone, see Fig. 6.

As the density of in situ observations is increased to similar levels as the climate model output itself, then the value added from the climate model output in inference of the unbiased parameters is reduced, illustrated through results for scenario one. The number of in situ observations is sufficient to adequately capture the spatial features of the underlying process (Fig. 6a) as well as the latent spatial covariance structure, encoded through the hyper-parameter estimates of the latent GP (Table 3). Additionally, ~~if~~ as the complexity of the bias signal is increased, through for example reducing the length scale of the latent generating process, as in scenario three, then again added value is reduced. The relatively more complex bias structure compared with scenario two makes it more difficult to disaggregate the climate model output into its biased and unbiased components. ~~While~~ Despite this, while added value is reduced for scenarios one and three relative to scenario two, incorporating the climate model output in inference is still shown to improve overall performance. Modelling the generating

process for the bias explicitly ~~is also informative and~~ also provides informative information that is potentially useful for future climate model development.

In addition to modelling shared latent processes, another important feature of the methodology presented in this paper is the Bayesian framework~~, where~~ . In this framework the parameters/hyper-parameters of the hierarchical model are treated as random variables with associated ~~distributions. This framework is flexible and allows for robust uncertainty propagation, which is important for making the model~~ probability distributions. Uncertainty is inherently propagated through the framework, making the code implementation flexible to further development and so applicable to a wide range of real-world ~~applications where bias prediction is required~~scenarios. Additionally, ~~expert knowledge can~~ a Bayesian framework allows expert knowledge to be incorporated in the inference through the choice of prior distributions, which ~~is~~ can be especially important where the data is sparse. In Sect. 4.2, results for a simulated ~~one dimensional~~ hierarchical example illustrate uncertainty propagation between ~~parameter values of the PDF at each sample site and the values of~~ the PDF parameter values and the hyper-parameters of the latent generating processes. Uncertainty present in the different levels of the hierarchical model are incorporated in the final posterior predictive estimates of the unbiased PDF parameters at the climate model locations, see Fig. 7. Multiple realisations from the posterior predictive can then be used in quantile mapping ~~, which is illustrated in Fig.. This results in~~ to produce multiple realisations of the final bias corrected time series, with an expectation and uncertainty range~~. Robust uncertainty computation that incorporates the spatial relationships between points~~ , illustrated in Fig. 8. Reliable uncertainty bands on the final bias corrected time series is important for impact assessments and resulting decision making. ~~Having~~ Additionally, having multiple realisations for the final bias corrected time series ~~is also useful for~~ allows further propagation of uncertainty in process models driven by climate model output, such as land surface models (Liu et al., 2014).

The simulated examples presented provide an initial proof of concept, although future studies validating the methodology against real-world applications are important for understanding the remaining limitations and areas for further development. The current primary limitation is expected to be that the underlying spatial covariance structures are assumed stationary. That is that the covariance length scale is assumed constant across the domain, whereas for real-world applications over large and complex topographic domains the length scale will be expected to change depending on the specific topography of the region. Further development of the methodology to incorporate non-stationary kernels would therefore be valuable, although is beyond the scope of this paper. Another important limitation to consider is the assumption that the bias is time independent. In situations where the bias varies gradually through time and uniformly across the domain, the methodology can be further developed such that the mean function of the GPs is modelled with a time dependency. If the bias varies in time non-uniformly across the domain, spatiotemporal GPs will need to be considered, which is again beyond the scope of this paper. Secondary limitations, include the assumption that the unbiased and biased components of the PDF parameter values are independent. In situations where there is a dependence between these components, the methodology presented is still expected to perform adequately, although information is lost by not modelling the dependency explicitly. Additionally, many real-world applications will necessitate specific model adjustments, such as incorporating a mean function dependent on factors like elevation and latitude. Finally, the computational complexity of the model is an important remaining consideration, with inference time of GPs scaling as the cube of the number of data points. Incorporating techniques such as using sparse variational

GPs or upscaling the climate model output, while outside the scope of this paper, will aid computational performance under demanding real-world scenarios and will facilitate further model development.

6 Conclusion

Current approaches for bias prediction and correction do not aim to preserve the spatial covariance structure of the climate model output ~~(Ehret et al., 2012). Climate models are fundamentally based on established physical laws and so the covariance structures are desirable since it is reasonable to assume that they are physically realistic. In addition, current approaches~~ and typically either neglect uncertainty or inadequately model uncertainty propagation ~~through the model. In this paper a~~ (Ehret et al., 2012). This paper presents a novel fully Bayesian hierarchical ~~model~~ framework for bias correction ~~is presented where~~ with uncertainty propagation and latent GP distributions ~~are~~ used to capture and preserve underlying covariance structures. ~~The Bayesian nature allows robust uncertainty propagation under a flexible modelling framework where the model is easily expanded for specific real-world scenarios, increasing the scope of the work~~In this framework bias is considered in the parameters of the time-independent PDF at each site. Estimates of the unbiased PDF parameters are made at the climate model locations and then quantile mapping is applied to produce the final bias corrected time series. The novelty of the approach lies in the fully Bayesian implementation, assuming shared latent GPs between the in situ data and climate model output and in propagating uncertainty through the quantile mapping step.

Simple simulated examples are chosen to illustrate ~~the~~ key features of the ~~model~~framework. In Sect. ~~??~~4.1, results are displayed for ~~a~~ non-hierarchical ~~example~~ examples where the focus is on illustrating the ~~nature of GPs and how assuming a shared latent GP between~~ advantage of modelling spatial covariance in both the in situ data and climate model output~~allows inference on the unbiased field from both sources of data~~, assuming shared latent GPs. This is shown to be particularly important in the case of sparse ~~data and a simple bias~~ in situ observations and bias that varies smoothly across the domain, where the climate model output ~~provides significant value added in predictions~~itself provides significant added value in predictions of the unbiased PDF parameters. In Sect. ~~??~~4.2, results are presented for a hierarchical case and focus is on illustrating how the model propagates uncertainty between the different levels and to the final ~~parameter predictions that are used in bias correction. Uncertainty in the parameter estimates is easily propagated in bias correction of the time series from~~ unbiased PDF parameter predictions at the climate model ~~at every location through the existing approach of quantile mapping. This~~ locations. In addition, a simulated example of propagating this uncertainty through quantile mapping is then provided to demonstrate how this results in a bias corrected time series with uncertainty bands, which is desirable for use in impact studies ~~that compute predictions on responses to climate change~~ and for informing ~~decisions based on these. This is especially true in~~ decision making. Adequately modelling uncertainty in the bias corrected time series is expected to be especially important over areas where the climatology is hard to model and in situ observations are sparse, such as ~~Antarctica, meaning the uncertainty is expected to be significant (Carter et al., 2022).~~

~~The model presented is~~ over Antarctica (Carter et al., 2022). The framework presented provides a step towards adequately capturing uncertainty and incorporating underlying spatial covariance structures from the climate model in bias correction. ~~The~~

~~primary limitation is the assumption that the spatial structure of the site-level parameters can be adequately modelled through a stationary GP. Over large and complex topographic regions it is likely that the covariance length scale will vary across the domain and this is something that will need assessing for each specific application. Additionally, many real-world applications will necessitate specific model adjustments, such as incorporating a mean function dependent on factors like elevation and latitude, handling non-Gaussian data, and accounting for other bias structures. Future papers validating the framework proposed in this paper against a range of~~ While initial results are promising, further studies applied to real-world ~~applications, assessing the added value of the approach as well as remaining limitations, is an important next step~~datasets are important to further validate the approach and explore remaining limitations. The Bayesian ~~approach adopted means~~ implementation provides a flexible modelling framework, where adjustments to the methodology needed for specific applications can be made ~~with uncertainty inherently propagated adequately~~while inherently propagating uncertainty.

Code and data availability. The code used to generate the simulated data, fit the model, make predictions and create the figures/tables is available at: https://doi.org/10.5281/zenodo.10053653 (Carter, a).

The data used to create the plots is available at: https://doi.org/10.5281/zenodo.10053531 (Carter, b).

Author contributions. J.Carter: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - Original Draft. E.Chacón-Montalván: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - Review & Editing, Supervision. A.Leeson: Conceptualization, Writing - Review & Editing, Supervision.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. J.Carter and A.Leeson are supported by the Data Science for the Natural Environment project (EPSRC grant number EP/R01860X/1). The code for analysis is written in Python 3.9.13 and makes extensive use the following libraries: Numpyro (Phan et al., 2019); TinyGP (Foreman-Mackey); NumPy (Harris et al., 2020); Matplotlib (Hunter, 2007).

References

Bader, D., Covey, C., Gutowski, W., Held, I., Kunkel, K., Miller, R., Tokmakian, R., and Zhang, M.: Climate Models: An Assessment of Strengths and Limitations, Climate Models: An Assessment of Strengths and Limitations, 2008.

Carter, J.: Bias Correction of Climate Models using a Bayesian Hierarchical Model: Code, https://doi.org/10.5281/zenodo.10053653, a.

Carter, J.: Data used in generation of results in 'Bias Correction of Climate Models using a Bayesian Hierarchical Model' J.Carter et. al., https://doi.org/10.5281/zenodo.10053531, b.

Carter, J., Leeson, A., Orr, A., Kittel, C., and van Wessem, M.: Variability in Antarctic Surface Climatology Across Regional Climate Models and Reanalysis Datasets, https://doi.org/10.5281/zenodo.6367850, 2022.

Cattiaux, J., Douville, H., and Peings, Y.: European temperatures in CMIP5: origins of present-day biases and future uncertainties, Climate Dynamics, 41, 2889–2907, https://doi.org/10.1007/s00382-013-1731-y, 2013.

Das, A., Rokaya, P., and Lindenschmidt, K.-E.: The impact of a bias-correction approach (delta change) applied directly to hydrological model output when modelling the severity of ice jam flooding under future climate scenarios, Climatic Change, 172, 19, https://doi.org/10.1007/s10584-022-03364-5, 2022.

DeConto, R. M. and Pollard, D.: Contribution of Antarctica to past and future sea-level rise, Nature, 531, 591–597, https://doi.org/10.1038/nature17145, number: 7596 Publisher: Nature Publishing Group, 2016.

Doblas-Reyes, F. J., Sörensson, A. A., Almazroui, M., Dosio, A., Gutowski, W. J., Haarsma, R., Hamdi, R., Hewitson, B., Kwon, W.-T., Lamptey, B. L., Maraun, D., Stephenson, T. S., Takayabu, I., Terray, L., Turner, A., and Zuo, Z.: Linking global to regional climate change, 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, edited by 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. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., pp. 1363–1512, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896.001, 2021.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrology and Earth System Sciences, 16, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012, publisher: Copernicus GmbH, 2012.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models, in: Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Intergovernmental Panel on Climate Change (IPCC), pp. 741–866, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9781107415324.020, 2013.

Foreman-Mackey, D.: dfm/tinygp: The tiniest of Gaussian Process libraries, https://doi.org/10.5281/zenodo.7646759.

Giorgi, F.: Thirty Years of Regional Climate Modeling: Where Are We and Where Are We Going next?, Journal of Geophysical Research: Atmospheres, 124, 5696–5723, https://doi.org/10.1029/2018JD030094, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018JD030094, 2019.

Greeves, C., Pope, V., Stratton, R., and Martin, G.: Representation of Northern Hemisphere winter storm tracks in climate models, Clim. Dynam., 28, 683–702, https://doi.org/10.1007/s00382-006-0205-x, 2007.

Guilyardi, E., Wittenberg, A., Fedorov, A., Collins, C., Capotondi, A., Van Oldenborgh, G. J., and Stockdale, T.: Understanding El Niño in Ocean-Atmosphere General Circulation Models: Progress and Challenges, Bulletin of the American Meteorological Society, 90, 325–340, https://doi.org/10.1175/2008BAMS2387.1, 2009.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357–362, https://doi.org/10.1038/s41586-020-2649-2, number: 7825 Publisher: Nature Publishing Group, 2020.

Hoffman, M. D. and Gelman, A.: The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, Journal of Machine Learning Research, 15, 1593–1623, http://jmlr.org/papers/v15/hoffman14a.html, 2014.

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., Folini, D., Ji, D., Klocke, D., Qian, Y., Rauser, F., Rio, C., Tomassini, L., Watanabe, M., and Williamson, D.: The Art and Science of Climate Model Tuning, Bulletin of the American Meteorological Society, 98, 589–602, https://doi.org/10.1175/BAMS-D-15-00135.1, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, 2017.

Hunter, J. D.: Matplotlib: A 2D Graphics Environment, Computing in Science Engineering, 9, 90–95, https://doi.org/10.1109/MCSE.2007.55, conference Name: Computing in Science Engineering, 2007.

Lima, C. H. R., Kwon, H.-H., and Kim, Y.-T.: A Bayesian Kriging model applied for spatial downscaling of daily rainfall from GCMs, Journal of Hydrology, 597, 126 095, https://doi.org/10.1016/j.jhydrol.2021.126095, 2021.

Liu, M., Rajagopalan, K., Chung, S. H., Jiang, X., Harrison, J., Nergui, T., Guenther, A., Miller, C., Reyes, J., Tague, C., Choate, J., Salathé, E. P., Stöckle, C. O., and Adam, J. C.: What is the importance of climate model bias when projecting the impacts of climate change on land surface processes?, Biogeosciences, 11, 2601–2622, https://doi.org/10.5194/bg-11-2601-2014, publisher: Copernicus GmbH, 2014.

Maraun, D.: Bias Correcting Climate Change Simulations - a Critical Review, Current Climate Change Reports, 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, 2016.

Phan, D., Pradhan, N., and Jankowiak, M.: Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro, https://doi.org/10.48550/arXiv.1912.11554, arXiv:1912.11554 [cs, stat], 2019.

Qian, W. and Chang, H. H.: Projecting Health Impacts of Future Temperature: A Comparison of Quantile-Mapping Bias-Correction Methods, International Journal of Environmental Research and Public Health, 18, 1992, https://doi.org/10.3390/ijerph18041992, 2021.

Sippel, S., Otto, F. E. L., Forkel, M., Allen, M. R., Guillod, B. P., Heimann, M., Reichstein, M., Seneviratne, S. I., Thonicke, K., and Mahecha, M. D.: A novel bias correction methodology for climate impact simulations, Earth System Dynamics, 7, 71–88, https://doi.org/10.5194/esd-7-71-2016, publisher: Copernicus GmbH, 2016.

Tebaldi, C., Debeire, K., Eyring, V., Fischer, E., Fyfe, J., Friedlingstein, P., Knutti, R., Lowe, J., O'Neill, B., Sanderson, B., van Vuuren, D., Riahi, K., Meinshausen, M., Nicholls, Z., Tokarska, K. B., Hurtt, G., Kriegler, E., Lamarque, J.-F., Meehl, G., Moss, R., Bauer, S. E., Boucher, O., Brovkin, V., Byun, Y.-H., Dix, M., Gualdi, S., Guo, H., John, J. G., Kharin, S., Kim, Y., Koshiro, T., Ma, L., Olivié, D., Panickal, S., Qiao, F., Rong, X., Rosenbloom, N., Schupfner, M., Séférian, R., Sellar, A., Semmler, T., Shi, X., Song, Z., Steger, C., Stouffer, R., Swart, N., Tachiiri, K., Tang, Q., Tatebe, H., Voldoire, A., Volodin, E., Wyser, K., Xin, X., Yang, S., Yu, Y., and Ziehn, T.: Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6, Earth System Dynamics, 12, 253–293, https://doi.org/10.5194/esd-12-253-2021, publisher: Copernicus GmbH, 2021.

(a) Scenario 1

Dependent Variable	Model Parameter	Specified Value	Prior Distribution Exp.	Std. Dev.	C.I. L.	C.I. U.	Posterior Dist. (Shared Process) Exp.	Std. Dev.	C.I. L.	C.I. U.	Posterior Dist. (Single Process) Exp.	Std. Dev.	C.I. L.	C.I. U.
	Kernel Variance v_{ϕ_Y}	1.0	0.67	0.67	0.02	2.46	1.25	0.30	0.73	1.86	1.04	0.31	0.57	1.69
	Kernel Lengthscale l_{ϕ_Y}	3.0	15.00	8.66	3.09	36.12	2.96	0.06	2.85	3.08	2.73	0.20	2.32	3.10
	Mean Constant m_{ϕ_Y}	1.0	0.00	2.00	-3.92	3.92	1.14	0.28	0.61	1.68	1.23	0.26	0.74	1.76
	Noise σ_{ϕ_Y}	0.1	2.00	2.00	0.05	7.38	0.11	0.01	0.09	0.12	N/A	N/A	N/A	N/A
	Kernel Variance v_{ϕ_B}	1.0	15.00	8.66	3.09	36.12	2.10	1.30	0.48	4.72	N/A	N/A	N/A	N/A
	Kernel Lengthscale l_{ϕ_B}	10.0	0.00	2.00	-3.92	3.92	11.45	1.28	9.07	14.00	N/A	N/A	N/A	N/A
	Mean Constant m_{ϕ_B}	-1.0	0.25	0.14	0.01	0.49	-1.00	0.64	-2.31	0.24	N/A	N/A	N/A	N/A

(b) Scenario 2

Dependent Variable	Model Parameter	Specified Value	Prior Distribution Exp.	Std. Dev.	C.I. L.	C.I. U.	Posterior Dist. (Shared Process) Exp.	Std. Dev.	C.I. L.	C.I. U.	Posterior Dist. (Single Process) Exp.	Std. Dev.	C.I. L.	C.I. U.
	Kernel Variance v_{ϕ_Y}	1.0	0.67	0.67	0.02	2.46	1.13	0.28	0.66	1.66	1.49	0.53	0.65	2.55
	Kernel Lengthscale l_{ϕ_Y}	3.0	15.00	8.66	3.09	36.12	2.97	0.06	2.86	3.09	3.70	0.44	2.83	4.56
	Mean Constant m_{ϕ_Y}	1.0	0.00	2.00	-3.92	3.92	0.70	0.27	0.15	1.22	0.69	0.40	-0.14	1.44
	Noise σ_{ϕ_Y}	0.1	2.00	2.00	0.05	7.38	0.12	0.03	0.08	0.18	N/A	N/A	N/A	N/A
	Kernel Variance v_{ϕ_B}	1.0	15.00	8.66	3.09	36.12	1.24	0.99	0.16	3.23	N/A	N/A	N/A	N/A
	Kernel Lengthscale l_{ϕ_B}	20.0	0.00	2.00	-3.92	3.92	23.69	5.79	12.29	34.90	N/A	N/A	N/A	N/A
	Mean Constant m_{ϕ_B}	-1.0	0.25	0.14	0.01	0.49	-0.66	0.64	-1.87	0.62	N/A	N/A	N/A	N/A

(c) Scenario 3

Dependent Variable	Model Parameter	Specified Value	Prior Distribution Exp.	Std. Dev.	C.I. L.	C.I. U.	Posterior Dist. (Shared Process) Exp.	Std. Dev.	C.I. L.	C.I. U.	Posterior Dist. (Single Process) Exp.	Std. Dev.	C.I. L.	C.I. U.
	Kernel Variance v_{ϕ_Y}	1.0	0.67	0.67	0.02	2.46	1.18	0.33	0.62	1.83	0.85	0.33	0.30	1.50
	Kernel Lengthscale l_{ϕ_Y}	3.0	15.00	8.66	3.09	36.12	3.00	0.07	2.87	3.14	3.08	0.49	2.03	3.96
	Mean Constant m_{ϕ_Y}	1.0	0.00	2.00	-3.92	3.92	0.95	0.30	0.35	1.53	0.90	0.29	0.33	1.48
	Noise σ_{ϕ_Y}	0.1	2.00	2.00	0.05	7.38	0.16	0.06	0.03	0.27	N/A	N/A	N/A	N/A
	Kernel Variance v_{ϕ_B}	1.0	15.00	8.66	3.09	36.12	1.50	1.02	0.28	3.56	N/A	N/A	N/A	N/A
	Kernel Lengthscale l_{ϕ_B}	5.0	0.00	2.00	-3.92	3.92	6.34	1.71	3.23	9.20	N/A	N/A	N/A	N/A
	Mean Constant m_{ϕ_B}	-1.0	0.25	0.14	0.01	0.49	-1.17	0.50	-2.11	-0.10	N/A	N/A	N/A	N/A

Table 3. A table showing summary statistics for the prior and posterior distributions including the expectation (Exp.), standard deviation (Std. Dev.) and lower and upper bounds for the 95% credible interval (C.I. L. and C.I. U.). The posterior distributions for the shared and single process models are given. The specified value for each parameter used to generate the data is also shown.

	R^2 Scores: Posterior Predictive Estimates of $\phi_Y(\boldsymbol{s}_z)$			
	Shared Process Model		Single Process Model	
Scenario	Exp.	Std.Dev.	Exp.	Std.Dev.
1	0.99	0.00	0.97	0.01
2	0.99	0.01	0.68	0.07
3	0.74	0.12	0.52	0.10

Table 4. A table showing the expectation and standard deviation of R^2 scores for the posterior predictive estimates of the unbiased PDF parameter at the climate model output locations $\phi_Y(\boldsymbol{s}_z)$ for the shared and single process models for each scenario.