Probabilistic downscaling to detect regional present and future climate hazards

Sherman Lo, Ritabrata Dutta, Peter Dueben, Peter Watson

University of Warwick, ECMWF, University of Bristol

28th April 2020

Funding

- University of Warwick
- Alan Turing Institute

Introduction

Introduction

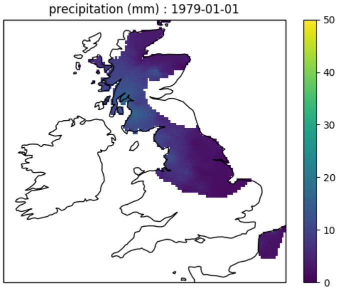


Figure: Precipitation

Introduction

We have computer simulations of the weather/climate, they are called *model fields*.

Introduction

We have computer simulations of the weather/climate, they are called *model fields*.

- Temperature
- Precipitable water content
- Humidity
- Geopotential
- Wind speed and velocity

Introduction

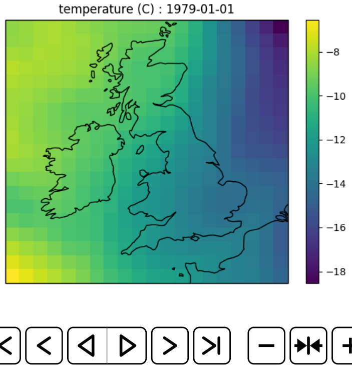
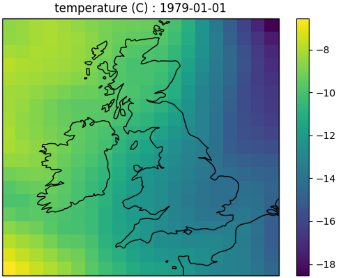
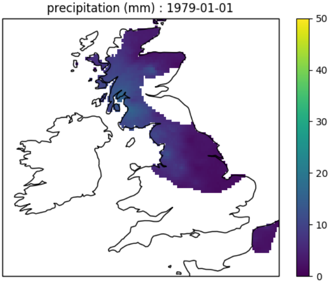


Figure: Air temperature

Introduction



(a) Model fields



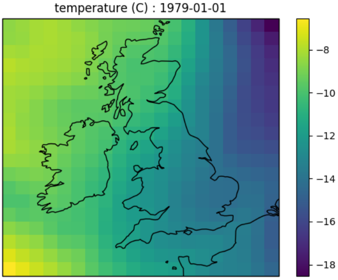
(b) Precipitation

Figure: Sample from data

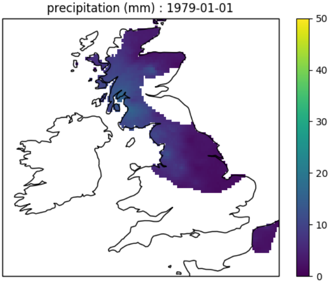
Introduction

- 40 years of data (1979 - 2019)
- Observed precipitation ($\sim 10\,\text{km}$)
- Simulated model fields ($\sim 80\,\text{km}$)

Introduction



(a) Model fields



(b) Precipitation

Figure: Sample from data

Introduction

Can we use the model fields on a *coarse grid* to forecast the precipitation on the *fine grid*? This task is known as *downscaling*.

Introduction

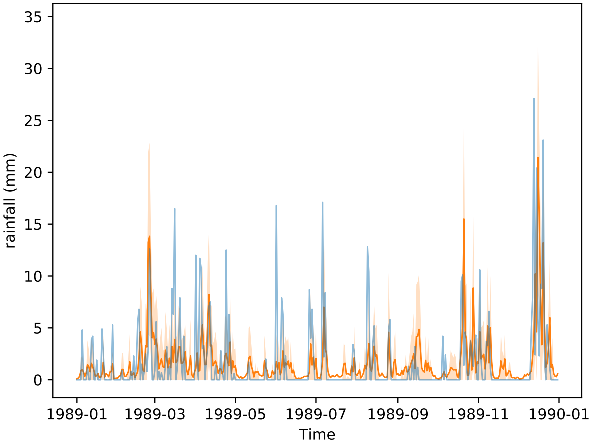


Figure: Forecast (orange) and observed (blue)

Introduction

$1.2 \pm 1.5\,\text{mm}$

Statistical model

Statistical model

- Occurrence of precipitation

Statistical model

- Occurrence of precipitation
- Quantity of precipitation

Statistical model

- Occurrence of precipitation
- Quantity of precipitation
- Autocorrelation of precipitation

Statistical model

- Occurrence of precipitation
- Quantity of precipitation
- Autocorrelation of precipitation
- Affected by the model fields

Statistical model

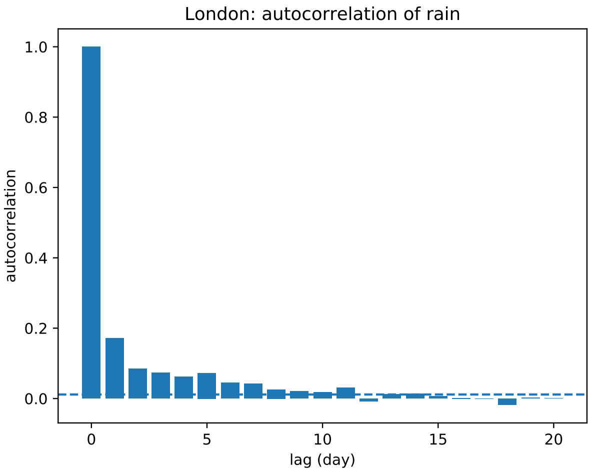


Figure: Autocorrelation of precipitation

Statistical model

We are dealing with a *zero-inflated* random variable.

Statistical model

We are dealing with a *zero-inflated* random variable. Sometimes, the precipitation is *exactly* 0 mm,

Statistical model

We are dealing with a *zero-inflated* random variable. Sometimes, the precipitation is *exactly* 0 mm, sometimes it is a positive number.

Statistical model

We are dealing with a *zero-inflated* random variable. Sometimes, the precipitation is *exactly* 0 mm, sometimes it is a positive number.

- We introduce the *compound-Poisson*.

Statistical model

$$Z_t = \text{number of times it rains at day } t$$

Statistical model

$$Z_t = \text{number of times it rains at day } t$$

$$R_t^{(i)} = \text{amount of precipitation in a rain event at day } t$$

Statistical model

$$Z_t = \text{number of times it rains at day } t$$

$$R_t^{(i)} = \text{amount of precipitation in a rain event at day } t$$

$$Y_t = \text{amount of precipitation at day } t$$

Statistical model

$$Z_t = \text{number of times it rains at day } t$$

$$R_t^{(i)} = \text{amount of precipitation in a rain event at day } t$$

$$Y_t = \text{amount of precipitation at day } t$$

$$Y_t | Z_t = \sum_{i=1}^{Z_t} R_t^{(i)}$$

Statistical model

$$Z_t \sim \text{Poisson}(\lambda_t)$$

Statistical model

$$Z_t \sim \text{Poisson}(\lambda_t)$$

$$R_t^{(i)} \sim \text{Gamma}(1/\omega_t, 1/(\omega_t\mu_t))$$

Statistical model

$$Z_t \sim \text{Poisson}(\lambda_t)$$

$$R_t^{(i)} \sim \text{Gamma}(1/\omega_t, 1/(\omega_t \mu_t))$$

$$Y_t | Z_t = \sum_{i=1}^{Z_t} R_t^{(i)} \sim \text{Gamma}(Z_t/\omega_t, 1/(\omega_t \mu_t))$$

Statistical model

$$Z_t \sim \text{Poisson}(\lambda_t)$$

$$R_t^{(i)} \sim \text{Gamma}(1/\omega_t, 1/(\omega_t \mu_t))$$

$$Y_t | Z_t = \sum_{i=1}^{Z_t} R_t^{(i)} \sim \text{Gamma}(Z_t/\omega_t, 1/(\omega_t \mu_t))$$

$$Y_t \sim \text{Compound-Poisson}$$

Statistical model

- We make λ_t, μ_t, and ω_t a function of the model fields $\mathbf{x_t}$.

Statistical model

- We make λ_t, μ_t, and ω_t a function of the model fields $\mathbf{x_t}$.
- We make Z_t and $Y_t|Z_t$ undergo an autoregressive and moving average process.

Statistical model

- We make λ_t, μ_t, and ω_t a function of the model fields $\mathbf{x_t}$.
- We make Z_t and $Y_t|Z_t$ undergo an autoregressive and moving average process.

Statistical model

- We make λ_t, μ_t, and ω_t a function of the model fields $\mathbf{x_t}$.
- We make Z_t and $Y_t|Z_t$ undergo an autoregressive and moving average process.

For example

$$\lambda_t = \exp\left[\boldsymbol{\beta}_\lambda \mathbf{x_t}\right.$$

Statistical model

- We make λ_t, μ_t, and ω_t a function of the model fields $\mathbf{x_t}$.
- We make Z_t and $Y_t|Z_t$ undergo an autoregressive and moving average process.

For example

$$\lambda_t = \exp\Bigg[\boldsymbol{\beta}_\lambda \mathbf{x_t} + \phi_\lambda(\ln \lambda_{t-1} - k_\lambda)$$

Statistical model

- We make λ_t, μ_t, and ω_t a function of the model fields $\mathbf{x_t}$.
- We make Z_t and $Y_t | Z_t$ undergo an autoregressive and moving average process.

For example

$$\lambda_t = \exp\left[\boldsymbol{\beta}_\lambda \mathbf{x_t} + \phi_\lambda(\ln \lambda_{t-1} - k_\lambda) + \theta_\lambda \frac{Z_{t-1} - \lambda_{t-1}}{\sqrt{\lambda_{t-1}}}\right.$$

Statistical model

- We make λ_t, μ_t, and ω_t a function of the model fields $\mathbf{x_t}$.
- We make Z_t and $Y_t|Z_t$ undergo an autoregressive and moving average process.

For example

$$\lambda_t = \exp\left[\boldsymbol{\beta}_\lambda \mathbf{x_t} + \phi_\lambda(\ln \lambda_{t-1} - k_\lambda) + \theta_\lambda \frac{Z_{t-1} - \lambda_{t-1}}{\sqrt{\lambda_{t-1}}} + k_\lambda\right]$$

Statistical model

- We make λ_t, μ_t, and ω_t a function of the model fields $\mathbf{x_t}$.
- We make Z_t and $Y_t|Z_t$ undergo an autoregressive and moving average process.

For example

$$\lambda_t = \exp\left[\boldsymbol{\beta}_\lambda \mathbf{x_t} + \phi_\lambda(\ln\lambda_{t-1} - k_\lambda) + \theta_\lambda\frac{Z_{t-1} - \lambda_{t-1}}{\sqrt{\lambda_{t-1}}} + k_\lambda\right]$$

$$\mu_t = \exp\left[\boldsymbol{\beta}_\mu \mathbf{x_t} + \phi_\mu(\ln\mu_{t-1} - k_\mu) + \theta_\mu\frac{y_{t-1} - Z_{t-1}\mu_{t-1}}{\mu_{t-1}\sqrt{Z_{t-1}\omega_{t-1}}} + k_\mu\right]$$

Statistical model

- We make λ_t, μ_t, and ω_t a function of the model fields $\mathbf{x_t}$.
- We make Z_t and $Y_t|Z_t$ undergo an autoregressive and moving average process.

For example

$$\lambda_t = \exp\left[\boldsymbol{\beta}_\lambda \mathbf{x_t} + \phi_\lambda(\ln \lambda_{t-1} - k_\lambda) + \theta_\lambda \frac{Z_{t-1} - \lambda_{t-1}}{\sqrt{\lambda_{t-1}}} + k_\lambda\right]$$

$$\mu_t = \exp\left[\boldsymbol{\beta}_\mu \mathbf{x_t} + \phi_\mu(\ln \mu_{t-1} - k_\mu) + \theta_\mu \frac{y_{t-1} - Z_{t-1}\mu_{t-1}}{\mu_{t-1}\sqrt{Z_{t-1}\omega_{t-1}}} + k_\mu\right]$$

$$\omega_t = \exp\left[\boldsymbol{\beta}_\omega \mathbf{x_t} + k_\omega\right]$$

Statistical model

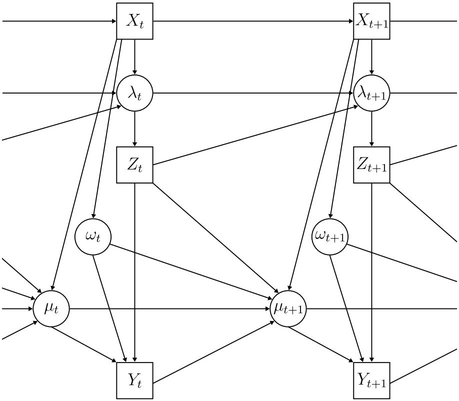
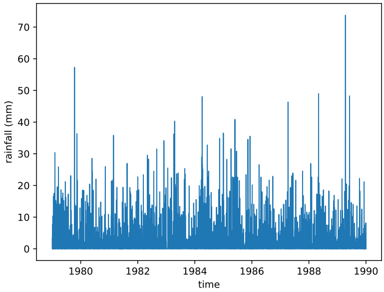
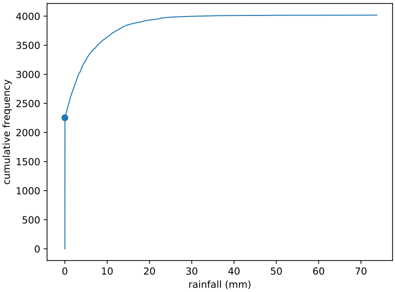


Figure: Graphical model

Statistical model



(a) Time series



(b) Cumulative

Figure: Simulated data

Statistical model

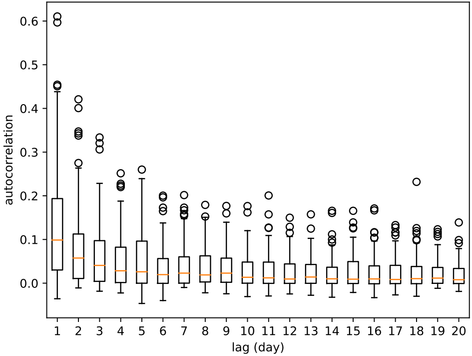


Figure: Simulated data autocorrelation ARMA(5,5)

Markov chain Monte Carlo

Markov chains Monte Carlo

- This model is wrong so it is important to quantify how wrong the model is.

Markov chains Monte Carlo

- This model is wrong so it is important to quantify how wrong the model is.
- The problem of modelling fitting can be done using Bayesian inference to get *uncertainity quantification*.

Markov chains Monte Carlo

- This model is wrong so it is important to quantify how wrong the model is.
- The problem of modelling fitting can be done using Bayesian inference to get *uncertainity quantification*.
- This was done using Markov chains Monte Carlo (MCMC).

Markov chains Monte Carlo

- This model is wrong so it is important to quantify how wrong the model is.
- The problem of modelling fitting can be done using Bayesian inference to get *uncertainity quantification*.
- This was done using Markov chains Monte Carlo (MCMC).
- We use Monte Carlo which uses random numbers to solve a model fitting problem. Flucations in solutions, (samples), reflect the uncertainity.

Markov chains Monte Carlo

- This model is wrong so it is important to quantify how wrong the model is.
- The problem of modelling fitting can be done using Bayesian inference to get *uncertainity quantification*.
- This was done using Markov chains Monte Carlo (MCMC).
- We use Monte Carlo which uses random numbers to solve a model fitting problem. Flucations in solutions, (samples), reflect the uncertainity.
- We use a Markov chain to draw dependent samples in such a way they converge to the right answer.

Markov chains Monte Carlo

In Bayesian inference, we are interested in studying the posterior distribution

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

Markov chains Monte Carlo

In Bayesian inference, we are interested in studying the posterior distribution

$$\text{posterior} \propto \text{likelihood} \times \text{prior}$$

$$\underbrace{\pi(\beta, z_{1:T} | x_{1:T}, y_{1:T})}_{\text{posterior}} \propto \underbrace{p(y_{1:T}, z_{1:T} | x_{1:T}, \beta)}_{\text{likelihood}} \times \underbrace{\pi(\beta)}_{\text{prior}}$$

Markov chains Monte Carlo

$$\underbrace{p(\text{data}|\text{parameters})}_{\text{likelihood}}$$

Markov chains Monte Carlo

$$\underbrace{p(\text{data}|\text{parameters})}_{\text{likelihood}}$$

$$\underbrace{p(\text{parameters}|\text{data})}_{\text{posterior}}$$

Markov chains Monte Carlo

$$\underbrace{p(\text{data}|\text{parameters})}_{\text{likelihood}}$$

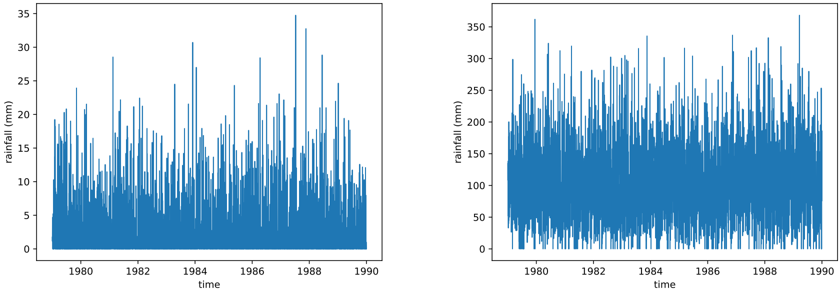
$$\underbrace{p(\text{parameters}|\text{data})}_{\text{posterior}} \propto \underbrace{p(\text{data}|\text{parameters})}_{\text{likelihood}} \times \underbrace{p(\text{parameters})}_{\text{prior}}$$

Markov chains Monte Carlo

But the prior is subjective!

Markov chains Monte Carlo

But the prior is subjective!



(a) Sensible (b) Not sensible

Figure: Simulated data

Markov chains Monte Carlo

We use Gibbs sampling to split the posterior into two parts

- Sample $Z_t | \beta, Z_{1:T\setminus t}, y_{1:T}, x_{1:T}$

Markov chains Monte Carlo

We use Gibbs sampling to split the posterior into two parts

- Sample $Z_t|\beta, Z_{1:T\setminus t}, y_{1:T}, x_{1:T}$
- Sample $\beta|y_{1:T}, Z_{1:T}, x_{1:T}$

Markov chains Monte Carlo

We use Gibbs sampling to split the posterior into two parts

- Sample $Z_t|\beta, Z_{1:T\setminus t}, y_{1:T}, x_{1:T}$
- Sample $\beta|y_{1:T}, Z_{1:T}, x_{1:T}$

We use *slice sampling* and *elliptical slice sampling*.

Markov chains Monte Carlo

We use slice sampling to sample $Z_t | \beta, Z_{1:T \setminus t}, y_{1:T}, x_{1:T}$.

Markov chains Monte Carlo

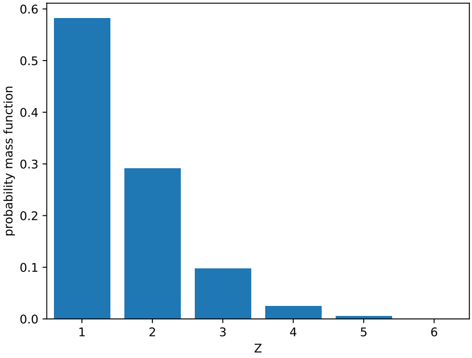
We use slice sampling to sample $Z_t|\beta, Z_{1:T\setminus t}, y_{1:T}, x_{1:T}$.

- $y_t = 0 \Rightarrow Z_t = 0$

Markov chains Monte Carlo

Slice sampling:

- $s \sim \text{Uniform}(0, \pi(z^{(i)}))$
- $z^{(i+1)} \sim \text{Uniform}(z : \pi(z) > s)$



Markov chains Monte Carlo

We use *elliptical* slice sampling to sample β.

Markov chains Monte Carlo

We use *elliptical* slice sampling to sample β. It can sample a posterior with a Normal prior.

Markov chains Monte Carlo

We use *elliptical* slice sampling to sample β. It can sample a posterior with a Normal prior. For the technical:

- Chain at $\beta^{(i)}$
- $\nu \sim$ prior
- $s \sim \text{Uniform}(0, \text{posterior}(\beta^{(i)}))$
- $\theta \sim \text{Uniform}(0, 2\pi)$
- $\beta^{(i+1)} = \beta^{(i)} \cos\theta + \nu \sin\theta$ if posterior$(\beta^{(i+1)}) > s$, else try another θ with a smaller support containing zero

Markov chains Monte Carlo

Even more technical:

- Initial: $[\theta_{\text{min}}, \theta_{\text{max}}] = [\theta - 2\pi, \theta]$
- If: $\theta < 0$, $\theta_{\text{min}} = \theta$
- Else: $\theta_{\text{max}} = \theta$

Markov chains Monte Carlo

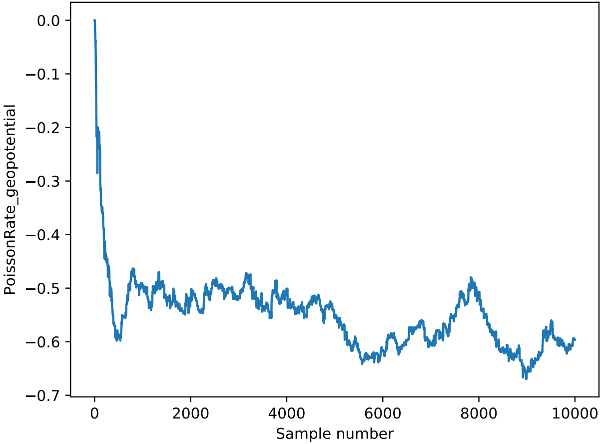


Figure: Example of a chain

Forecasting

Forecasting

Simulate the future using different samples from the posterior (MCMC chain). Any variation reflect the uncertainity.

Forecasting

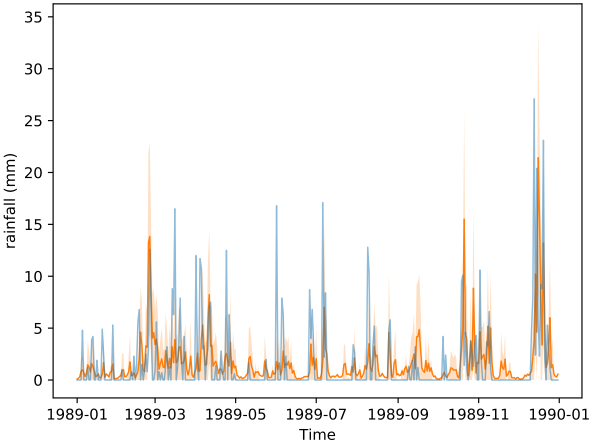


Figure: Forecast (orange) and observed (blue)

Forecasting

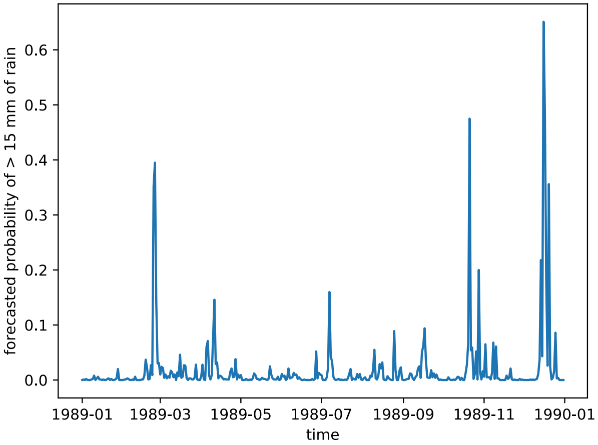


Figure: Probability precipitation $> 15\,\mathrm{mm}$

Forecasting

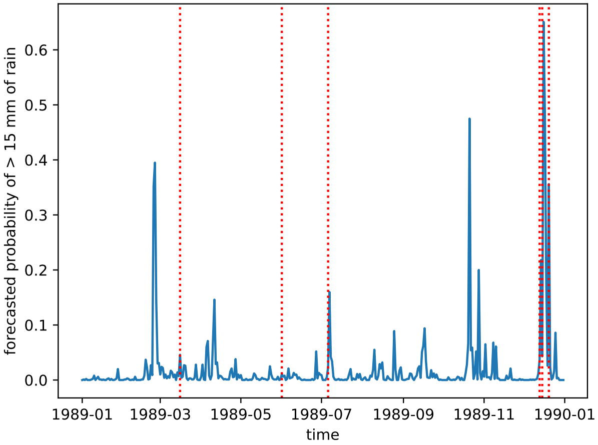


Figure: Probability precipitation $> 15\,\text{mm}$

Forecasting

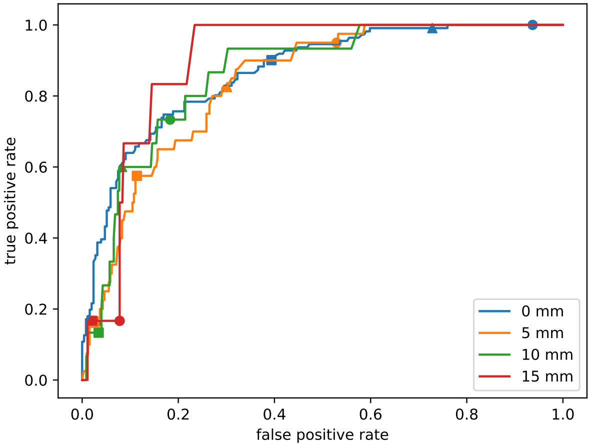
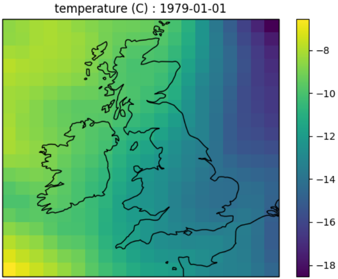


Figure: ROC curve

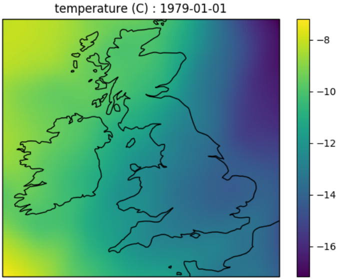
Downscaling

Downscaling

We interpolate the model fields for now



(a) Coarse grid



(b) Fine grid

Figure: Interpolation

Downscaling

We impose a prior on β where neighbouring locations with similar topography have similar values

Downscaling

We impose a prior on β where neighbouring locations with similar topography have similar values

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{pmatrix} \sim \mathrm{N}\left(0, \tau^{-1}\mathbf{K}\right)$$

where $[\mathbf{K}]_{i,j} = \exp\left[-\frac{\nu}{2}\left(w_i - w_j\right)^2\right]$ and w_i is the topography

Downscaling

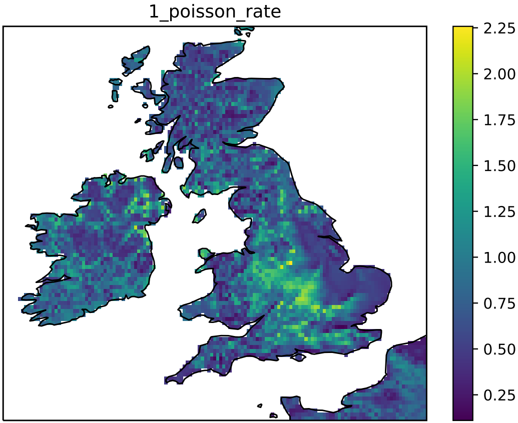


Figure: Simulation from prior

Downscaling

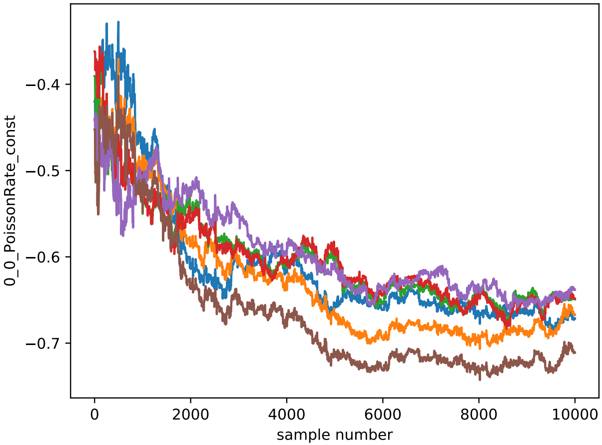


Figure: MCMC chains

Downscaling

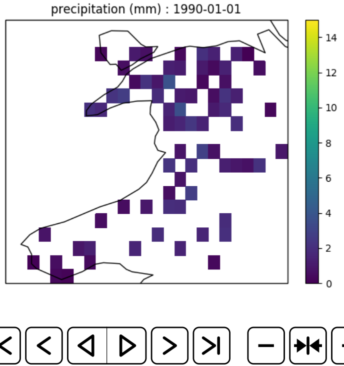


Figure: Forecast at Wales

Downscaling

Downscaling

Instead of interpolating the model fields, we can learn the model fields on the coarse grid. We use a *Gaussian process*.

Downscaling

- $x'_{i,t}$ = model fields at time t, location i on *coarse* grid
- $x_{j,t}$ = model fields at time t, location j on *fine* grid

$$\begin{pmatrix} x'_{1,t} \\ x'_{2,t} \\ \vdots \\ x'_{M,t} \end{pmatrix} \sim \mathrm{N}(0, \tau^{-1}\mathbf{K})$$

Downscaling

We can sample the model fields on the fine grid *given*:

- model fields on the coarse grid
- precipitation & Z on the fine grid

Downscaling

We can sample the model fields on the fine grid *given*:

- model fields on the coarse grid
- precipitation & Z on the fine grid

$$\begin{pmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{N,t} \end{pmatrix} \Bigg| \begin{pmatrix} x'_{1,t} \\ x'_{2,t} \\ \vdots \\ x'_{M,t} \end{pmatrix}, \{\mathbf{Y}_{1:T}\}, \{\mathbf{Z}_{1:T}\}$$

Conclusion

Conclusion

- We used the model fields on a *coarse grid* to forecast the precipitation on the *fine grid*.

Conclusion

- We used the model fields on a *coarse grid* to forecast the precipitation on the *fine grid*.
- We have quantified uncertainity.

Conclusion

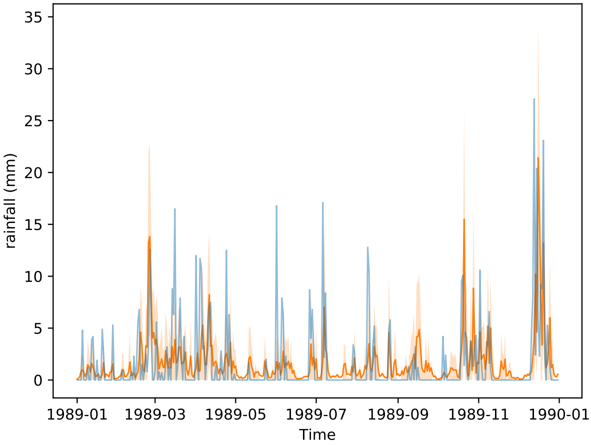


Figure: Forecast (orange) and observed (blue)

Conclusion

- The task of downscaling is computational expensive.

Conclusion

- To transistion from weather forecasting to climate prediction, we need to forecast many years into the future.

Conclusion

Thank you