Stochastic weather generator

Anastassia Baxevani

Avignon, 17 Sept 2014

Joint work with: Jan Lennartsson

Outline

Motivation

Motivation

Weather generator

Motivation

Weather generator

- Weak sense similarity

Motivation

Weather generator

- Weak sense similarity
- Extremes

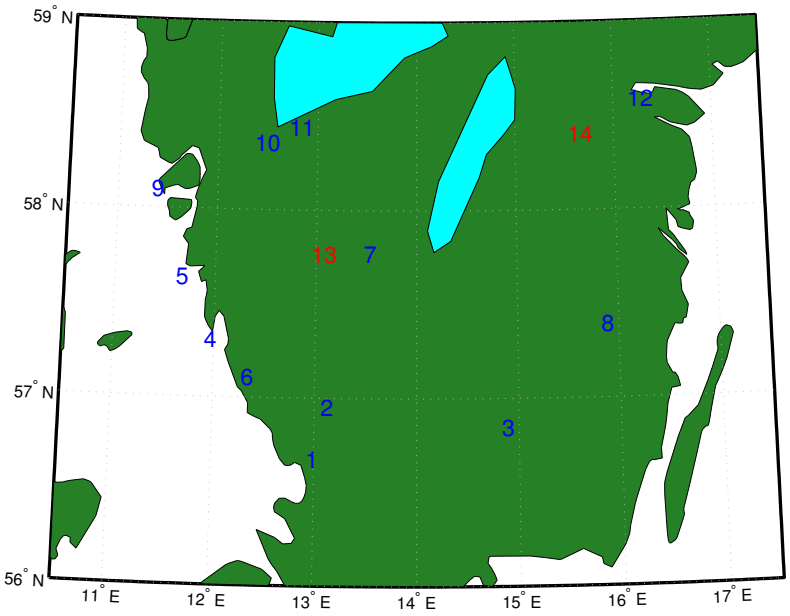
Motivation

Weather generator

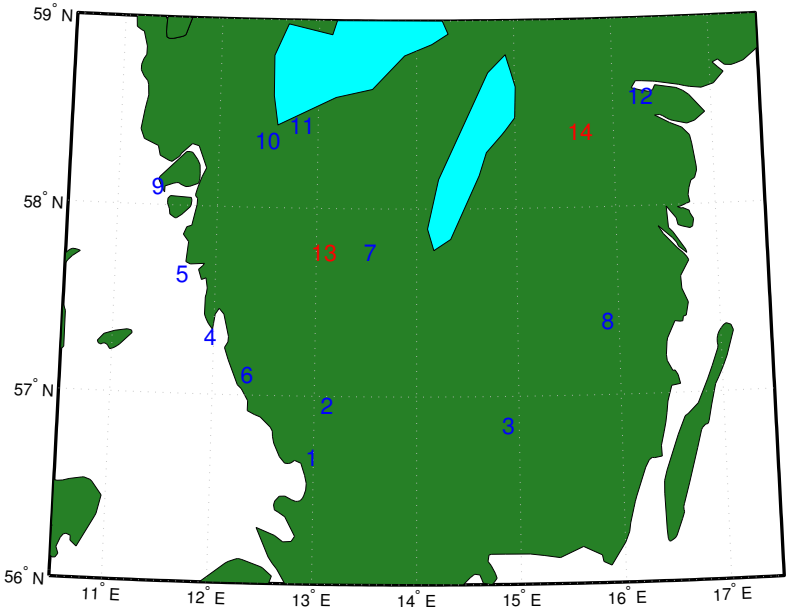
- Weak sense similarity
- Extremes

Data

Data


59° N
58° N
57° N
56° N
11° E
12° E
13° E
14° E
15° E
16° E
17° E
1
2
3
4
5
6
7
8
9
10
11
12
13
14

Data



- location $\mathbf{s} = (x, y)$
- time t measured in days

Model specification

Latent Gaussian field $W(\mathbf{s}, t)$

Model specification

Latent Gaussian field $W(\mathbf{s}, t)$

- For $Z \sim N(\mu, 1)$ denote distribution of $Z|Z > 0$ by Φ^+_μ

Model specification

Latent Gaussian field $W(\mathbf{s}, t)$

- For $Z \sim N(\mu, 1)$ denote distribution of $Z|Z > 0$ by Φ_{μ}^{+}
- $F_{\mathbf{s},t}$ is marginal distribution of intensities in wet days

Model specification

Latent Gaussian field $W(\mathbf{s}, t)$

- For $Z \sim N(\mu, 1)$ denote distribution of $Z|Z > 0$ by Φ_{μ}^{+}
- $F_{\mathbf{s},t}$ is marginal distribution of intensities in wet days ($F_{\mathbf{s},t}(0) = 0$)

Model specification

Latent Gaussian field $W(\mathbf{s}, t)$

- For $Z \sim N(\mu, 1)$ denote distribution of $Z|Z > 0$ by Φ^+_μ
- $F_{\mathbf{s},t}$ is marginal distribution of intensities in wet days ($F_{\mathbf{s},t}(0) = 0$)

$$Y(\mathbf{s}, t) = F^{-1}_{\mathbf{s},t} \circ \Phi^+_\mu(W(\mathbf{s}, t))$$

where μ is mean of W - is amount of precipitation at location $\mathbf{x}$ and time t.

Model identification

Model identification

- $F_{\mathbf{s},t}$

Model identification

- $F_{\mathbf{s},t}$
- mean and covariance functions of Gaussian field

Model identification

- $F_{\mathbf{s},t}$
- mean and covariance functions of Gaussian field

Problem : Observe transformed version of censored $W(\mathbf{s}, t)$

Model identification

- $F_{\mathbf{s},t}$
- mean and covariance functions of Gaussian field

Problem : Observe transformed version of censored $W(\mathbf{s}, t)$

Goal : Estimate mean and covariance function of W from observations of $W \vee 0$.

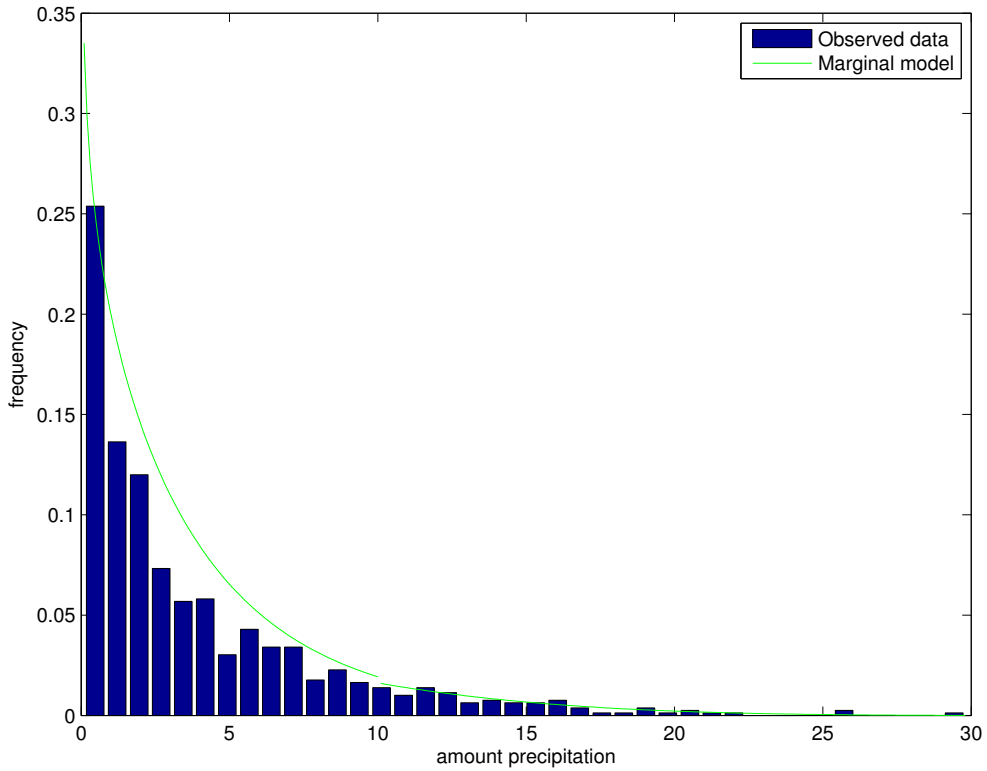
Marginal distributions

Marginal distributions

$$F_{\mathbf{s},t}(x) = F_\gamma(x \wedge u) + (1 - F_\gamma(u))F_u(x)$$

Marginal distributions

$$F_{\mathbf{s},t}(x) = F_\gamma(x \wedge u) + (1 - F_\gamma(u))F_u(x)$$



Fitting by modified likelihood

Mean function

Mean function

Link between precipitation $Y(\mathbf{s}, t)$ and Gaussian $W(\mathbf{s}, t)$

Mean function

Link between precipitation $Y(\mathbf{s}, t)$ and Gaussian $W(\mathbf{s}, t)$

$$\mathbf{1}_{Y>0} = \mathbf{1}_{W>0}$$

Mean function

Link between precipitation $Y(\mathbf{s}, t)$ and Gaussian $W(\mathbf{s}, t)$

$$\mathbf{1}_{Y>0} = \mathbf{1}_{W>0}$$

$$P(\textit{wet day}) = \mathbb{E}[\mathbf{1}_{Y>0}]$$

Mean function

Link between precipitation $Y(\mathbf{s}, t)$ and Gaussian $W(\mathbf{s}, t)$

$$\mathbf{1}_{Y>0} = \mathbf{1}_{W>0}$$

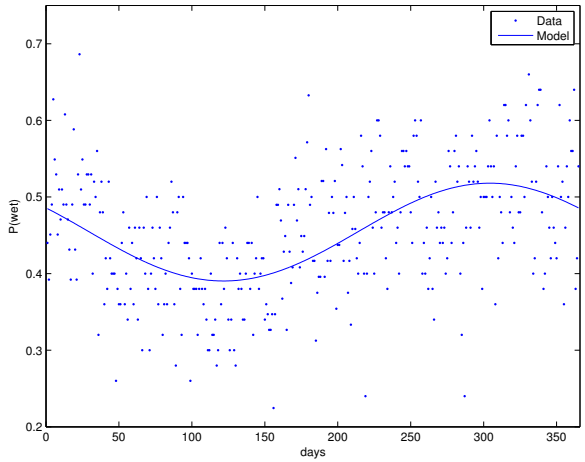
$$P(\textit{wet day}) = \mathbb{E}[\mathbf{1}_{Y>0}] = \Phi(\mu)$$

Mean function

Link between precipitation $Y(\mathbf{s}, t)$ and Gaussian $W(\mathbf{s}, t)$

$$\mathbf{1}_{Y>0} = \mathbf{1}_{W>0}$$

$$P(\textit{wet day}) = \mathbb{E}[\mathbf{1}_{Y>0}] = \Phi(\mu)$$



Mean function

Interpolation in time and space

Mean function

Interpolation in time and space

Interpolation in time

Mean function

Interpolation in time and space

Interpolation in time by Fourier series

Mean function

Interpolation in time and space

Interpolation in time by Fourier series

Interpolation in space

Mean function

Interpolation in time and space

Interpolation in time by Fourier series

Interpolation in space by linear regression on location + altitude

Covariance function

Implicit computations

Covariance function

Implicit computations

(Z_1, Z_2) bivariate Gaussian with variance 1 and covariance ρ.

Covariance function

Implicit computations

(Z_1, Z_2) bivariate Gaussian with variance 1 and covariance ρ.

$$E(Z_1 \vee 0)(Z_2 \vee 0) = \int_0^\infty g(x; \boldsymbol{\mu}, \rho)\phi(x)\ dx$$

$$g(x; \boldsymbol{\mu}, \rho) = x\phi(x-\mu_1)\cdot\left[(\rho(x-\mu_2)+\mu_1)\cdot\Phi\left(\frac{\mu_1+\rho(x-\mu_2)}{\sqrt{1-\rho^2}}\right)+\right.$$

$$\left.\sqrt{1-\rho^2}\phi\left(\frac{\mu_1+\rho x(x-\mu_2)}{\sqrt{1-\rho^2}}\right)\right].$$

Covariance function

Implicit computations

(Z_1, Z_2) bivariate Gaussian with variance 1 and covariance ρ.

$$E(Z_1 \vee 0)(Z_2 \vee 0) = \int_0^\infty g(x; \boldsymbol{\mu}, \rho)\phi(x)\ dx$$

$$g(x; \boldsymbol{\mu}, \rho) = x\phi(x-\mu_1)\cdot\left[(\rho(x-\mu_2)+\mu_1)\cdot\Phi\left(\frac{\mu_1+\rho(x-\mu_2)}{\sqrt{1-\rho^2}}\right)+\right.$$

$$\left.\sqrt{1-\rho^2}\phi\left(\frac{\mu_1+\rho x(x-\mu_2)}{\sqrt{1-\rho^2}}\right)\right].$$

Correlation $\rho_{ij}(\tau)$, estimated by minimizing

Covariance function

Implicit computations

(Z_1, Z_2) bivariate Gaussian with variance 1 and covariance ρ.

$$E(Z_1 \vee 0)(Z_2 \vee 0) = \int_0^\infty g(x; \boldsymbol{\mu}, \rho)\phi(x)\ dx$$

$$g(x; \boldsymbol{\mu}, \rho) = x\phi(x-\mu_1)\cdot\left[(\rho(x-\mu_2)+\mu_1)\cdot\Phi\left(\frac{\mu_1+\rho(x-\mu_2)}{\sqrt{1-\rho^2}}\right)+\right.$$

$$\left.\sqrt{1-\rho^2}\phi\left(\frac{\mu_1+\rho x(x-\mu_2)}{\sqrt{1-\rho^2}}\right)\right].$$

Correlation $\rho_{ij}(\tau)$, estimated by minimizing

$$min_\rho \left| \overline{z_i^+ \cdot z_j^+} - \int_0^\infty g(x; \mu_{ij}, \rho_{ij})\phi(x)\ dx \right|,$$

Covariance function

Covariance function

$C(\mathbf{h}, \tau)$

Covariance function

$C(\mathbf{h}, \tau)$

- spatial lag $\mathbf{h}$
- temporal lag τ

Covariance function

$C(\mathbf{h}, \tau)$

- spatial lag $\mathbf{h}$
- temporal lag τ

W Gaussian $\Rightarrow \Big(W(\mathbf{s}, t), W(\mathbf{s}+\mathbf{h}, t+\tau)\Big) \sim N(\mu, C(\mathbf{h}, \tau))$

Covariance function

$C(\mathbf{h}, \tau)$

- spatial lag $\mathbf{h}$
- temporal lag τ

W Gaussian $\Rightarrow \Big(W(\mathbf{s}, t), W(\mathbf{s}+\mathbf{h}, t+\tau)\Big) \sim N(\mu, C(\mathbf{h}, \tau))$

$$\mathbb{E}[(W(\mathbf{s}_i, t) \vee 0)(W(\mathbf{s_j}, t+\tau) \vee 0)]$$

Covariance function

$C(\mathbf{h}, \tau)$

- spatial lag $\mathbf{h}$
- temporal lag τ

W Gaussian $\Rightarrow \Big(W(\mathbf{s}, t), W(\mathbf{s}+\mathbf{h}, t+\tau)\Big) \sim N(\mu, C(\mathbf{h}, \tau))$

$$\mathbb{E}[(W(\mathbf{s}_i, t) \vee 0)(W(\mathbf{s_j}, t+\tau) \vee 0)] = f(C(\mathbf{h}_{ij}, \tau), \mu_{\mathbf{s}_i}, \mu_{\mathbf{s}_j})$$

Covariance function

$C(\mathbf{h}, \tau)$

- spatial lag $\mathbf{h}$
- temporal lag τ

W Gaussian $\Rightarrow \Big(W(\mathbf{s}, t), W(\mathbf{s}+\mathbf{h}, t+\tau)\Big) \sim N(\mu, C(\mathbf{h}, \tau))$

$$\mathbb{E}[(W(\mathbf{s}_i, t) \vee 0)(W(\mathbf{s_j}, t+\tau) \vee 0)] = f(C(\mathbf{h}_{ij}, \tau), \mu_{\mathbf{s}_i}, \mu_{\mathbf{s}_j})$$

where $\mathbf{h}_{ij} = \mathbf{s}_i - \mathbf{s}_j$

Covariance function

$C(\mathbf{h}, \tau)$

- spatial lag $\mathbf{h}$
- temporal lag τ

W Gaussian $\Rightarrow \Big(W(\mathbf{s}, t), W(\mathbf{s}+\mathbf{h}, t+\tau)\Big) \sim N(\mu, C(\mathbf{h}, \tau))$

$$\mathbb{E}[(W(\mathbf{s}_i, t) \vee 0)(W(\mathbf{s_j}, t+\tau) \vee 0)] = f(C(\mathbf{h}_{ij}, \tau), \mu_{\mathbf{s}_i}, \mu_{\mathbf{s}_j})$$

where $\mathbf{h}_{ij} = \mathbf{s}_i - \mathbf{s}_j$

Covariance model

Covariance model

Non-separable spatio-temporal covariance function

Covariance model

Non-separable spatio-temporal covariance function

$$C(\mathbf{h}, \tau) = \frac{1}{a|\tau| + 1} e^{-\frac{b||\mathbf{h}||^2}{a|\tau|+1}}$$

Covariance model

Non-separable spatio-temporal covariance function

$$C(\mathbf{h}, \tau) = \frac{1}{a|\tau| + 1} e^{-\frac{b||\mathbf{h}||^2}{a|\tau|+1}}$$

- nugget effect

Covariance model

Non-separable spatio-temporal covariance function

$$C(\mathbf{h}, \tau) = \frac{1}{a|\tau| + 1} e^{-\frac{b||\mathbf{h}||^2}{a|\tau|+1}}$$

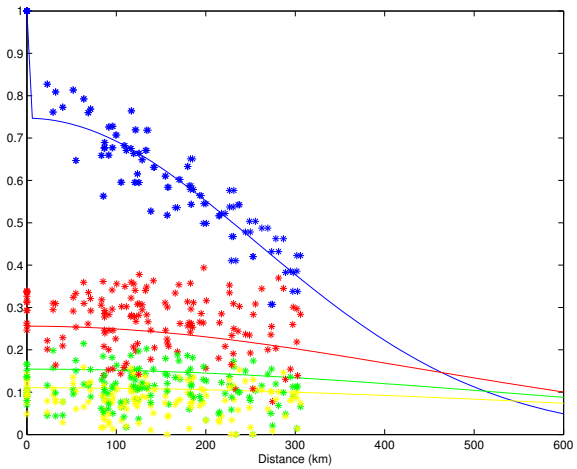
- nugget effect
- anisotropic distance $\mathbf{h}$

Covariance model

Non-separable spatio-temporal covariance function

$$C(\mathbf{h}, \tau) = \frac{1}{a|\tau| + 1} e^{-\frac{b||\mathbf{h}||^2}{a|\tau|+1}}$$

- nugget effect
- anisotropic distance $\mathbf{h}$



Validation

Validation

How good is the stochastic generator?

- temporal
- spatial
- spatio-temporal

Validation

How good is the stochastic generator?

- temporal
- spatial
- spatio-temporal

Simulations!

Validation

Temporal dependence structure

Validation

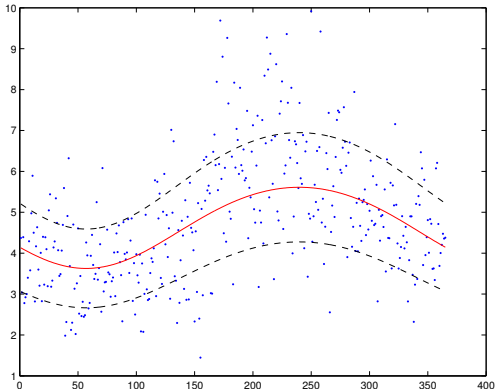
Temporal dependence structure

- daily mean intensitites, conditional proportions, mean intensities
- monthly qq-plots
- weather indices
- dry/wet spells

Validation

Temporal dependence structure

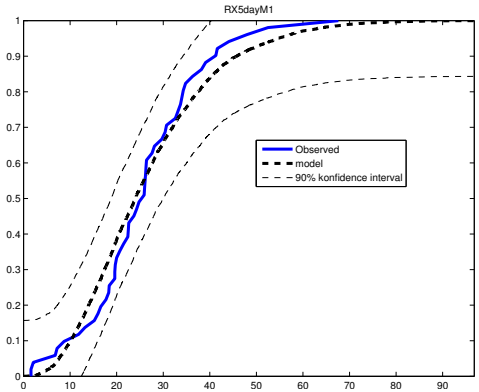
- daily mean intensitites, conditional proportions, mean intensities
- monthly qq-plots
- weather indices
- dry/wet spells



Validation

Temporal dependence structure

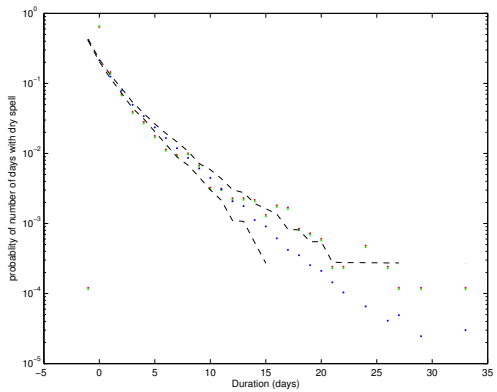
- daily mean intensitites, conditional proportions, mean intensities
- monthly qq-plots
- weather indices
- dry/wet spells



Validation

Temporal dependence structure

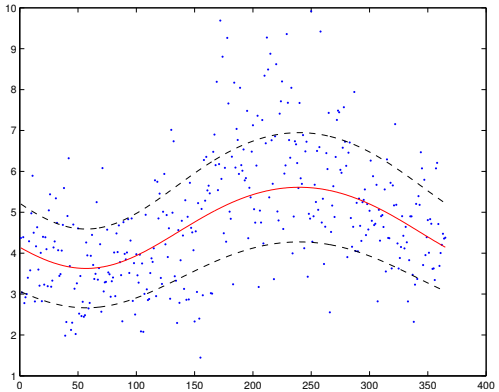
- daily mean intensitites, conditional proportions, mean intensities
- monthly qq-plots
- weather indices
- dry/wet spells

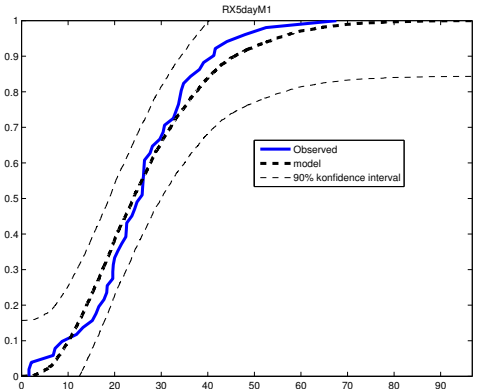


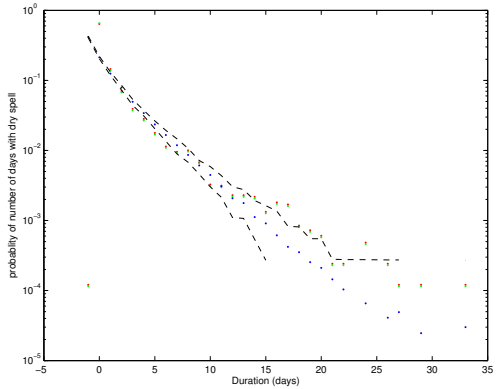
Validation

Temporal dependence structure

- daily mean intensitites, conditional proportions, mean intensities
- monthly qq-plots
- weather indices
- dry/wet spells







Validation

Spatial dependence

Validation

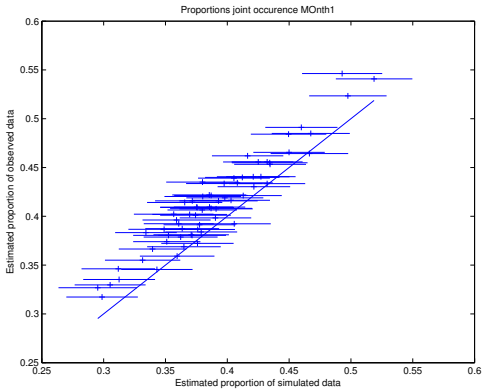
Spatial dependence

- proportions simultaneous occurrences of dry, wet days
- correlation of intensity
- no simultaneously wet stations

Validation

Spatial dependence

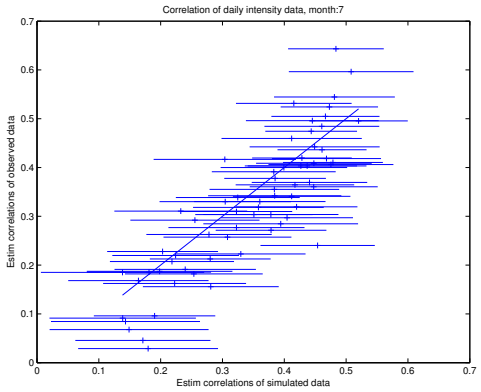
- proportions simultaneous occurrences of dry, wet days
- correlation of intensity
- no simultaneously wet stations



Validation

Spatial dependence

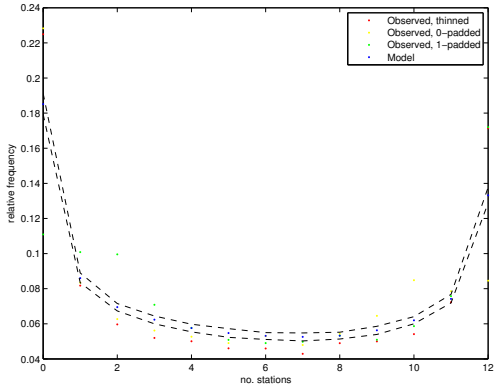
- proportions simultaneous occurrences of dry, wet days
- correlation of intensity
- no simultaneously wet stations



Validation

Spatial dependence

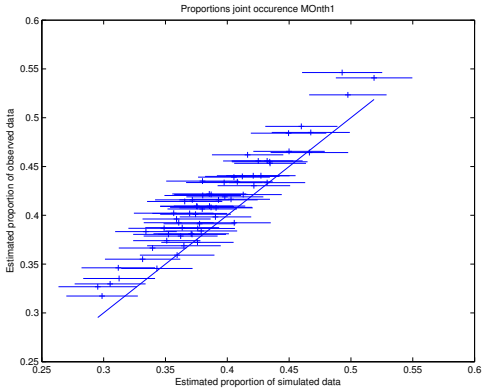
- proportions simultaneous occurrences of dry, wet days
- correlation of intensity
- no simultaneously wet stations

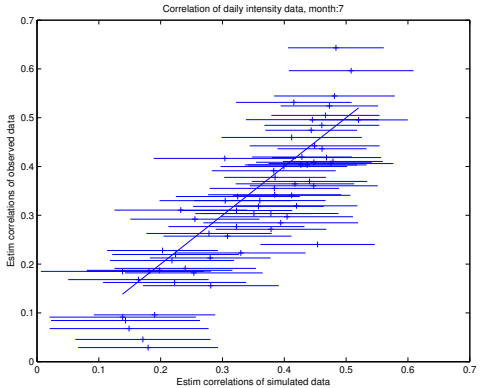


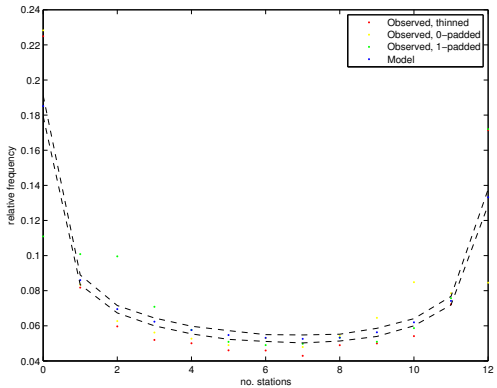
Validation

Spatial dependence

- proportions simultaneous occurrences of dry, wet days
- correlation of intensity
- no simultaneously wet stations







Validation

Spatio-temporal

Validation

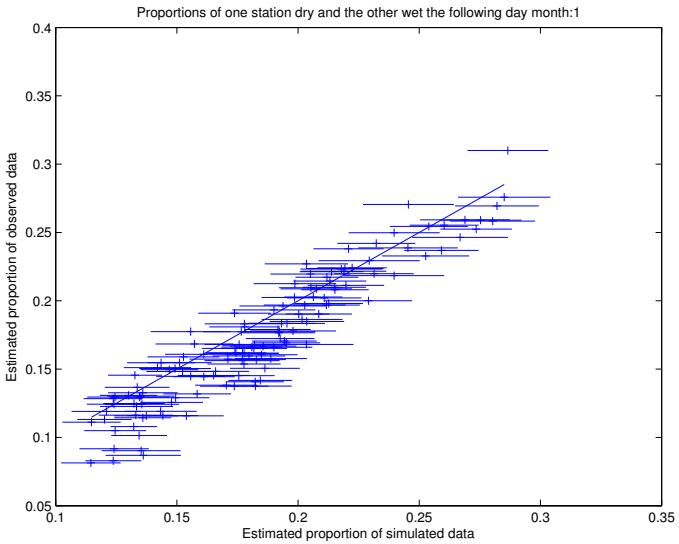
Spatio-temporal

- pairwise lagged occurrences
- weather indices of spatially aggregated data

Validation

Spatio-temporal

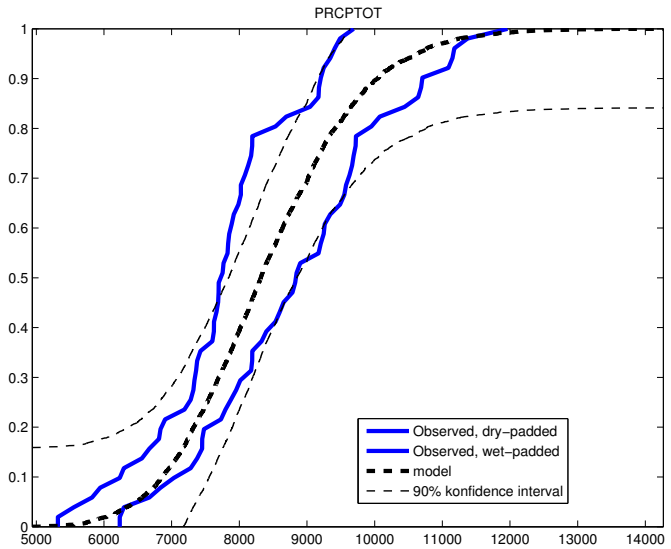
- pairwise lagged occurrences
- weather indices of spatially aggregated data



Validation

Spatio-temporal

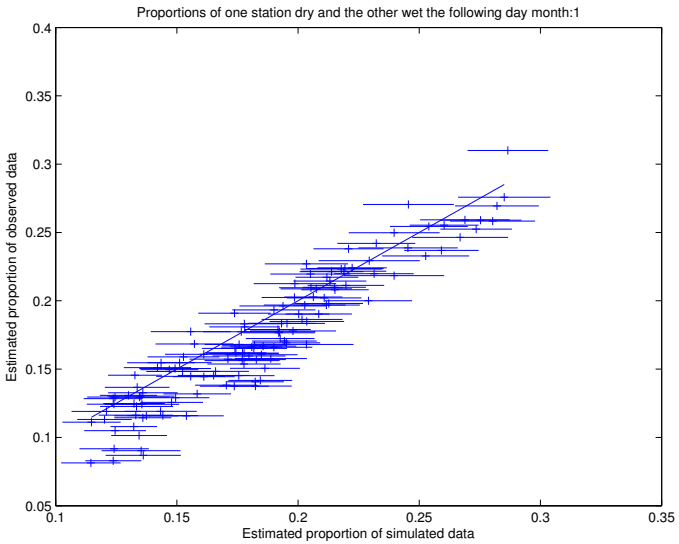
- pairwise lagged occurrences
- weather indices of spatially aggregated data

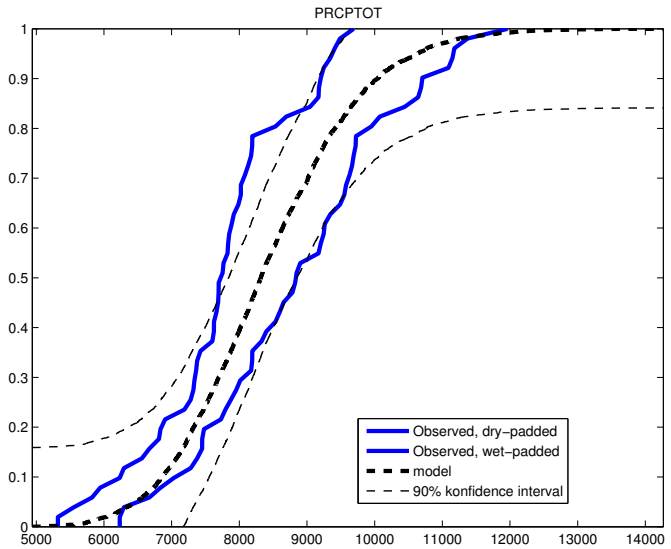


Validation

Spatio-temporal

- pairwise lagged occurrences
- weather indices of spatially aggregated data





Cross validation

Cross validation

- Location regression mean

Cross validation

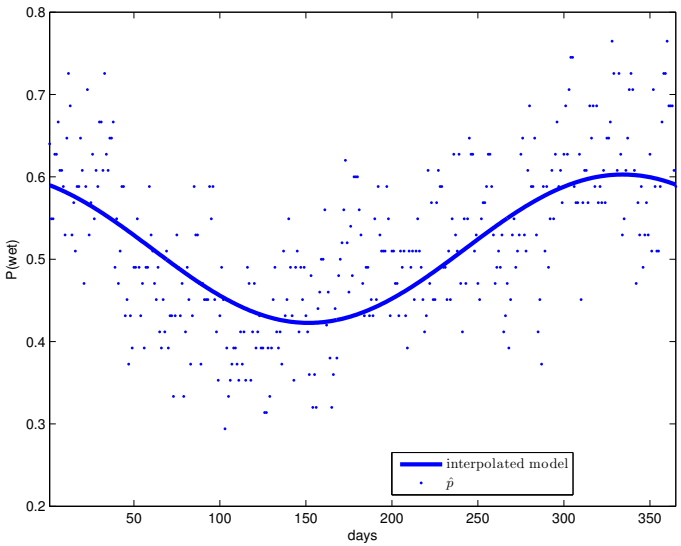
- Location regression mean
- empirical covariances between design and validation set stations

Cross validation

- Location regression mean
- empirical covariances between design and validation set stations

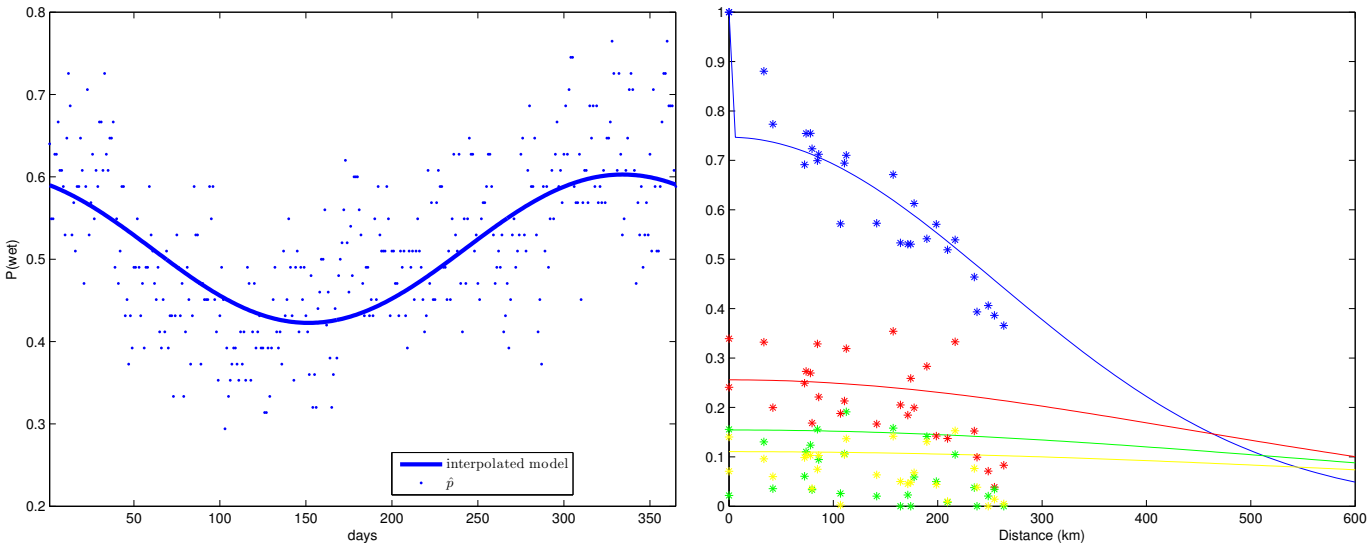
Cross validation

- Location regression mean
- empirical covariances between design and validation set stations



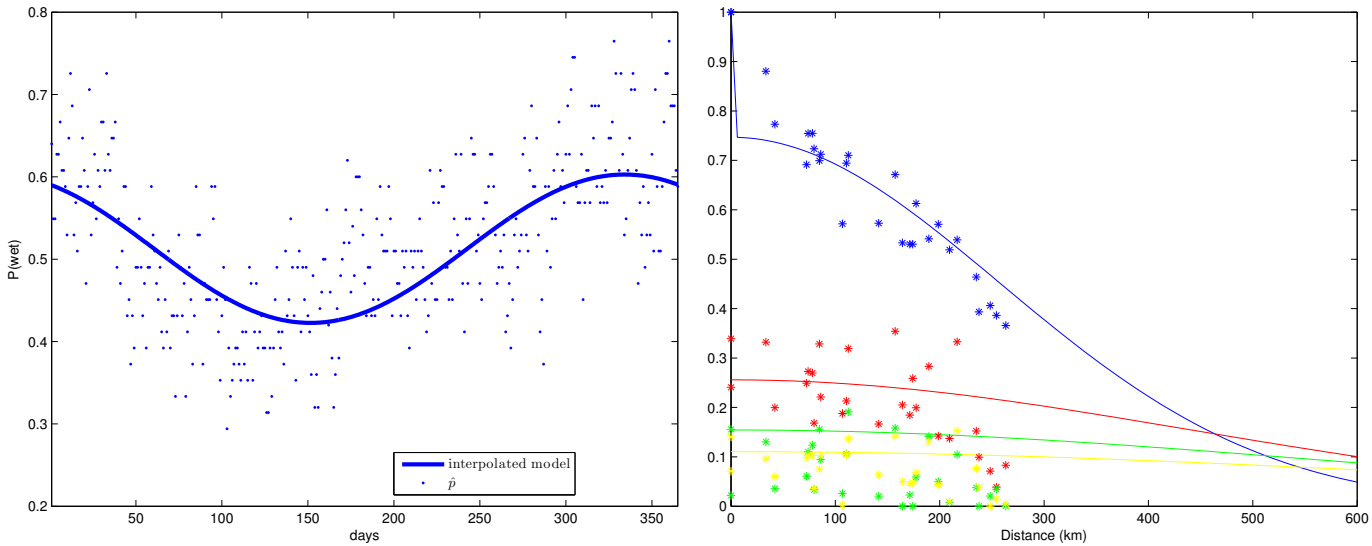
Cross validation

- Location regression mean
- empirical covariances between design and validation set stations



Cross validation

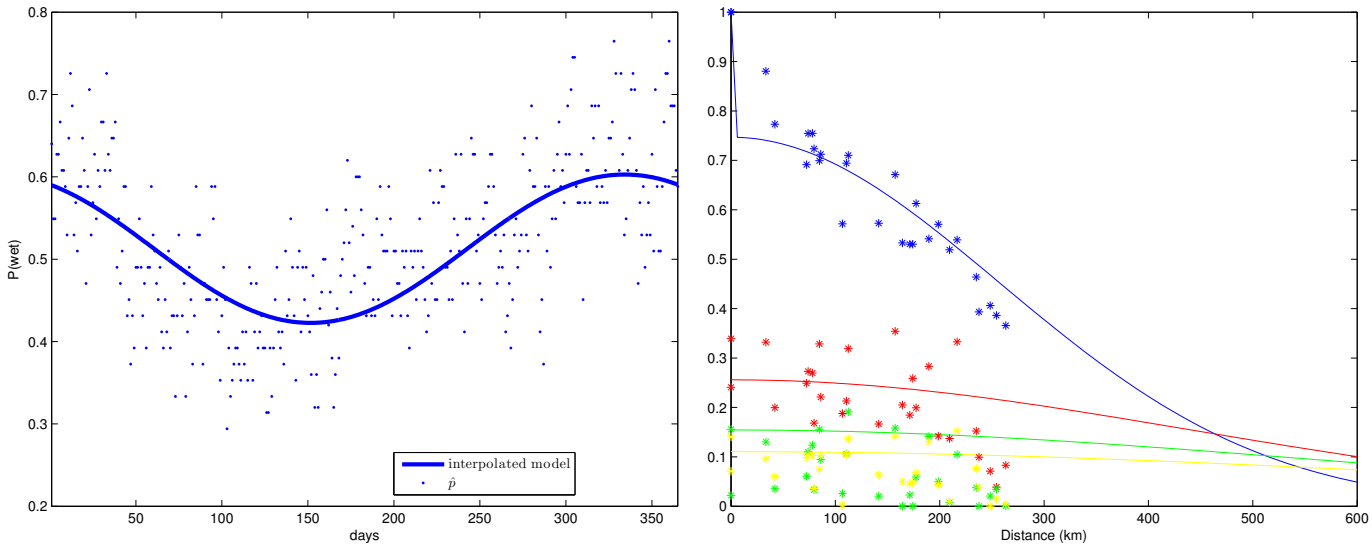
- Location regression mean
- empirical covariances between design and validation set stations



- marginal distributions:

Cross validation

- Location regression mean
- empirical covariances between design and validation set stations



- marginal distributions: location regression

Conclusion

Conclusion

- Model with few parameters that accurately replicate measures of observed data.

Conclusion

- Model with few parameters that accurately replicate measures of observed data.
- Extrapolate in space to locations where there is no available data and:

Conclusion

- Model with few parameters that accurately replicate measures of observed data.
- Extrapolate in space to locations where there is no available data and:
  - accurately replicates dry/wet behavior

Conclusion

- Model with few parameters that accurately replicate measures of observed data.
- Extrapolate in space to locations where there is no available data and:
  - accurately replicates dry/wet behavior
  - moderately well replicates measures dependent of marginal distributions

Conclusion

- Model with few parameters that accurately replicate measures of observed data.
- Extrapolate in space to locations where there is no available data and:
  - accurately replicates dry/wet behavior
  - moderately well replicates measures dependent of marginal distributions