GRAFFAS: Generation of RAinfall Fields From Areal Statistics.

Philippe Cantet, Benjamin Renard, Jean-Alain Fine, Patrick Arnaud

Water Resources Research

RESEARCH ARTICLE
10.1029/2024WR038562

Key Points:

- A model to generate rainfall fields from areal statistics is built by augmenting an existing model with spatial heterogeneity and anisotropy

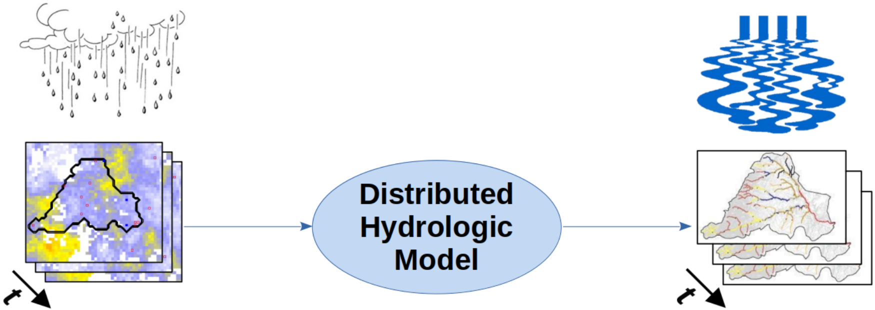
Space-Time Simulation of Hourly Rainfall From Areal Statistics Accounting for Spatial Heterogeneity and Anisotropy

P. Cantet[1], B. Renard[2], J. A. Fine[1], and P. Arnaud[2]

[1]HYDRIS Hydrologie, Montpellier, France, [2]RECOVER, INRAE, Aix-Marseille University, Aix-en-Provence, France

Objective

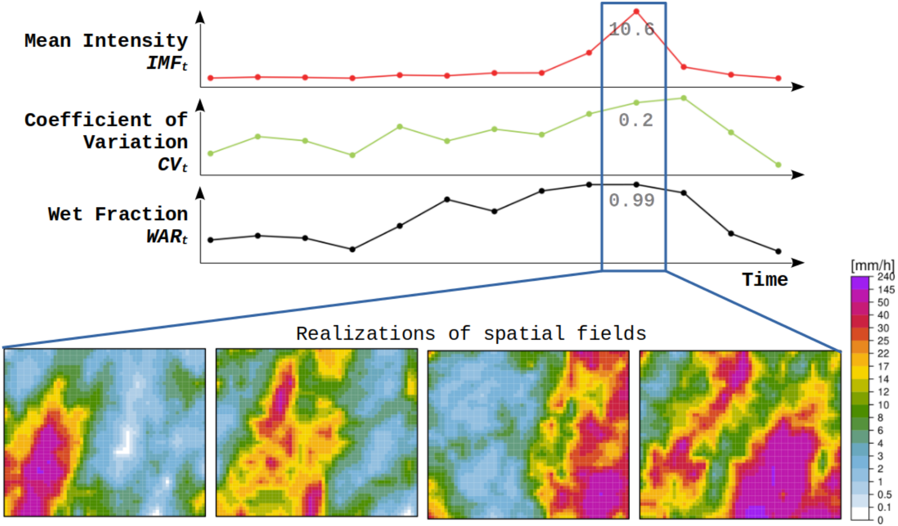
Aim: estimate streamflow extremes everywhere on the hydrologic network using continuous simulation



Objective

Requirement: Simulate long time series of rainfall fields (hourly, 1km²)

GRAFFAS: **G**eneration of **Ra**infall **F**ields **F**rom **A**real **S**tatistics



Simulation

[1] Simulation of Areal Statistics

IMF_t, WAR_t & CV_t

[2] Simulation of the Gaussian process MAR(1)

$$G_t(\mathbf{x}) = A.G_{t-1}(\mathbf{x}) + E_t(\mathbf{x})$$

[3] Transformation to Rainfall (fields) $R_t(\mathbf{x})$

$$R_t(x_i) = \begin{cases} 0 \ \text{ if } \Phi(G_t(x_i)) < 1 - WAR_t \\ R_t^{+}(x_i, IMF_t, CV_t) \quad \text{otherwise.} \end{cases}$$

GRAFFAS

Simulation

[1] Simulation of Areal Statistics

$IMF_t,\ WAR_t\ \&\ CV_t$

[2] Simulation of the Gaussian process MAR(1)

$$G_t(\mathbf{x}) = A.G_{t-1}(\mathbf{x}) + E_t(\mathbf{x})$$

[3] Transformation to Rainfall (fields) $R_t(\mathbf{x})$

$$R_t(x_i) = \begin{cases} 0 & \text{if } \Phi(G_t(x_i)) < 1 - WAR_t \\ R_t^+(x_i, IMF_t, CV_t) & \text{otherwise.} \end{cases}$$

GRAFFAS

1. Approach introduced by Paschalis et al. (2013)
2. Extremal behavior & temporal structure are mostly handled in step [1]
3. Spatial structure is handled in steps [2]-[3]

Simulation

Step [2]: Simulation of a Gaussian Field over N pixels $\mathbf{x}$:

$$G_t(\mathbf{x}) = A.G_{t-1}(\mathbf{x}) + E_t(\mathbf{x}) \text{ with } E_t(\mathbf{x}) \sim \mathcal{N}_N(0, (1 - A^2)\Sigma)$$

Simulation

Step [2]: Simulation of a Gaussian Field over N pixels $\mathbf{x}$:

$$G_t(\mathbf{x}) = A.G_{t-1}(\mathbf{x}) + E_t(\mathbf{x}) \text{ with } E_t(\mathbf{x}) \sim \mathcal{N}_N(0, (1 - A^2)\Sigma)$$

Assumptions

1. A diagonal $\implies$ autocorrelation $a_{i,i}$ at each pixel but no advection
2. Σ specified using a covariance function depending on the inter-pixel distance τ_{ij}:
   $\Sigma_{ij} = \rho(\tau_{ij}) = \exp\left[-\left(\frac{\tau_{ij}}{\lambda}\right)\right]$ for $i, j \in \{1, .., N\}$
3. To allow for anisotropy:
   $\tau_{ij} = \sqrt{(x_i - x_j)^T M^T M (x_i - x_j)}$ with $M = \begin{pmatrix} \cos\psi & \sin\psi \\ -b\sin\psi & b\cos\psi \end{pmatrix}$

Simulation

Step [3]: From Gaussian Fields to Rainfall fields:

$$R_t(x_i) = \begin{cases} 0 & \text{if } \Phi(G_t(x_i)) < 1 - WAR_t \\ F^{-1}\left(\frac{\Phi(G_t(x_i)) - 1 + WAR_t}{WAR_t}, \mu_t, \sigma_t\right) & \text{otherwise.} \end{cases}$$

With:

1. $F(., \mu, \sigma)$ the cdf of a $\mathcal{LN}(\mu, \sigma)$
2. $\sigma_t = \sqrt{\log(1 + {CV_t}^2)}$; $\mu_t = \log(IMF_t / WAR_t) - 0.5.\sigma_t^2$

Simulation

Step [3]: From Gaussian Fields to Rainfall fields:

$$R_t(x_i) = \begin{cases} 0 & \text{if } \Phi(G_t(x_i)) < 1 - WAR_t \\ F^{-1}\left(\frac{\Phi(G_t(x_i)) - 1 + WAR_t}{WAR_t}, \mu_t, \sigma_t\right) & \text{otherwise.} \end{cases}$$

With:

1. $F(., \mu, \sigma)$ the cdf of a $\mathcal{LN}(\mu, \sigma)$
2. $\sigma_t = \sqrt{\log(1 + CV_t{}^2)}$; $\mu_t = \log(IMF_t / WAR_t) - 0.5.\sigma_t^2$

Introducing spatial heterogeneity:

1. Introduce heterogeneity in μ_t by using:
   $\widehat{IMF}_t(x_i) = k_i^{IMF} \times IMF_t$

Simulation

Step [3]: From Gaussian Fields to Rainfall fields:

$$R_t(x_i) = \begin{cases} 0 & \text{if } \Phi(G_t(x_i)) < 1 - WAR_t \\ F^{-1}\left(\frac{\Phi(G_t(x_i)) - 1 + WAR_t}{WAR_t}, \mu_t, \sigma_t\right) & \text{otherwise.} \end{cases}$$

With:

1. $F(., \mu, \sigma)$ the cdf of a $\mathcal{LN}(\mu, \sigma)$
2. $\sigma_t = \sqrt{\log(1 + CV_t{}^2)}$; $\mu_t = \log(IMF_t / WAR_t) - 0.5.\sigma_t^2$

Introducing spatial heterogeneity:

1. Introduce heterogeneity in μ_t by using: $\widehat{IMF}_t(x_i) = k_i^{IMF} \times IMF_t$
2. Currently, $P_{wet}(x_i) = Pr(R_t(x_i) > 0) = WAR_t$

Simulation

Step [3]: From Gaussian Fields to Rainfall fields:

$$R_t(x_i) = \begin{cases} 0 & \text{if } \Phi(G_t(x_i)) < 1 - WAR_t \\ F^{-1}\left(\frac{\Phi(G_t(x_i)) - 1 + WAR_t}{WAR_t}, \mu_t, \sigma_t\right) & \text{otherwise.} \end{cases}$$

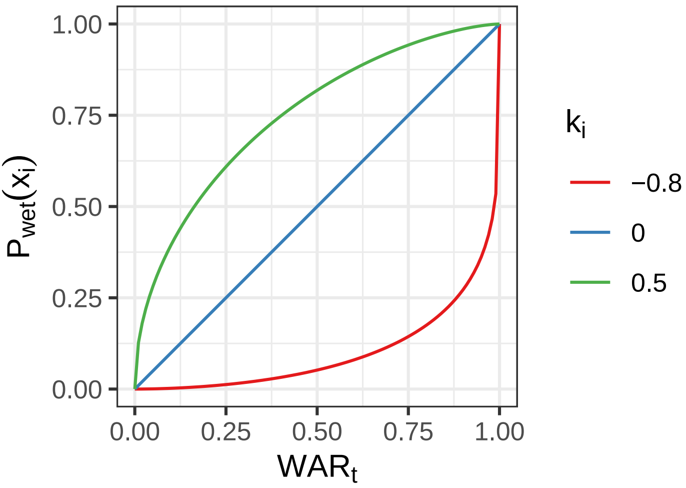
With:

1. $F(., \mu, \sigma)$ the cdf of a $\mathcal{LN}(\mu, \sigma)$
2. $\sigma_t = \sqrt{\log(1 + {CV_t}^2)}$; $\mu_t = \log(IMF_t / WAR_t) - 0.5.\sigma_t^2$

Introducing spatial heterogeneity:

1. Introduce heterogeneity in μ_t by using:
   $\widehat{IMF}_t(x_i) = k_i^{IMF} \times IMF_t$
2. Currently, $P_{wet}(x_i) = Pr(R_t(x_i) > 0) = WAR_t$
   Modify as: $P_{wet}(x_i) = B(WAR_t, 1 - k_i^{WAR}, 1 + k_i^{WAR})$

Simulation



Estimation

In summary...

Notation	Description	Step
$\{a_{i,i}\}_{i=1..N}$	Auto-regression matrix A	[2]
$\{\lambda, b, \psi\}$	Spatial covariance function	[2]
$\{k_i^{WAR}\}_{i=1..N}$	Spatial heterogeneity for occurrence	[3]
$\{k_i^{IMF}\}_{i=1..N}$	Spatial heterogeneity for intensity	[3]

Estimation

In summary...

Notation	Description	Step
$\{a_{i,i}\}_{i=1..N}$	Auto-regression matrix A	[2]
$\{\lambda, b, \psi\}$	Spatial covariance function	[2]
$\{k_i^{WAR}\}_{i=1..N}$	Spatial heterogeneity for occurrence	[3]
$\{k_i^{IMF}\}_{i=1..N}$	Spatial heterogeneity for intensity	[3]

Given an observed dataset $R_t(x_i)$ (N pixels $\times$ T time steps), yielding observed time series IMF_t, WAR_t and CV_t, estimation proceeds as follows:

Estimation

In summary...

Notation	Description	Step
$\{a_{i,i}\}_{i=1..N}$	Auto-regression matrix A	[2]
$\{\lambda, b, \psi\}$	Spatial covariance function	[2]
$\{k_i^{WAR}\}_{i=1..N}$	Spatial heterogeneity for occurrence	[3]
$\{k_i^{IMF}\}_{i=1..N}$	Spatial heterogeneity for intensity	[3]

Given an observed dataset $R_t(x_i)$ (N pixels $\times$ T time steps), yielding observed time series IMF_t, WAR_t and CV_t, estimation proceeds as follows:

1. k_i^{WAR} estimated independently at each pixel (Bernoulli likelihood)

Estimation

In summary...

Notation	Description	Step
$\{a_{i,i}\}_{i=1..N}$	Auto-regression matrix A	[2]
$\{\lambda, b, \psi\}$	Spatial covariance function	[2]
$\{k_i^{WAR}\}_{i=1..N}$	Spatial heterogeneity for occurrence	[3]
$\{k_i^{IMF}\}_{i=1..N}$	Spatial heterogeneity for intensity	[3]

Given an observed dataset $R_t(x_i)$ (N pixels $\times$ T time steps), yielding observed time series IMF_t, WAR_t and CV_t, estimation proceeds as follows:

1. k_i^{WAR} estimated independently at each pixel (Bernoulli likelihood)
2. k_i^{IMF} estimated independently at each pixel (Log-Normal lkh + tricks)

Estimation

In summary...

Notation	Description	Step
$\{a_{i,i}\}_{i=1..N}$	Auto-regression matrix A	[2]
$\{\lambda, b, \psi\}$	Spatial covariance function	[2]
$\{k_i^{WAR}\}_{i=1..N}$	Spatial heterogeneity for occurrence	[3]
$\{k_i^{IMF}\}_{i=1..N}$	Spatial heterogeneity for intensity	[3]

Given an observed dataset $R_t(x_i)$ (N pixels $\times$ T time steps), yielding observed time series IMF_t, WAR_t and CV_t, estimation proceeds as follows:

1. k_i^{WAR} estimated independently at each pixel (Bernoulli likelihood)
2. k_i^{IMF} estimated independently at each pixel (Log-Normal lkh + tricks)
3. A censored version of the Gaussian fields can then be computed $\Longrightarrow$ censored likelihood approach of Vaittinada Ayar et al. 2020

Estimation

In summary...

Notation	Description	Step
$\{a_{i,i}\}_{i=1..N}$	Auto-regression matrix A	[2]
$\{\lambda, b, \psi\}$	Spatial covariance function	[2]
$\{k_i^{WAR}\}_{i=1..N}$	Spatial heterogeneity for occurrence	[3]
$\{k_i^{IMF}\}_{i=1..N}$	Spatial heterogeneity for intensity	[3]

Given an observed dataset $R_t(x_i)$ (N pixels $\times$ T time steps), yielding observed time series IMF_t, WAR_t and CV_t, estimation proceeds as follows:

1. k_i^{WAR} estimated independently at each pixel (Bernoulli likelihood)
2. k_i^{IMF} estimated independently at each pixel (Log-Normal lkh + tricks)
3. A censored version of the Gaussian fields can then be computed $\Longrightarrow$ censored likelihood approach of Vaittinada Ayar et al. 2020
4. $a_{i,i}$ estimated independently at each pixel

Estimation

In summary...

Notation	Description	Step
$\{a_{i,i}\}_{i=1..N}$	Auto-regression matrix A	[2]
$\{\lambda, b, \psi\}$	Spatial covariance function	[2]
$\{k_i^{WAR}\}_{i=1..N}$	Spatial heterogeneity for occurrence	[3]
$\{k_i^{IMF}\}_{i=1..N}$	Spatial heterogeneity for intensity	[3]

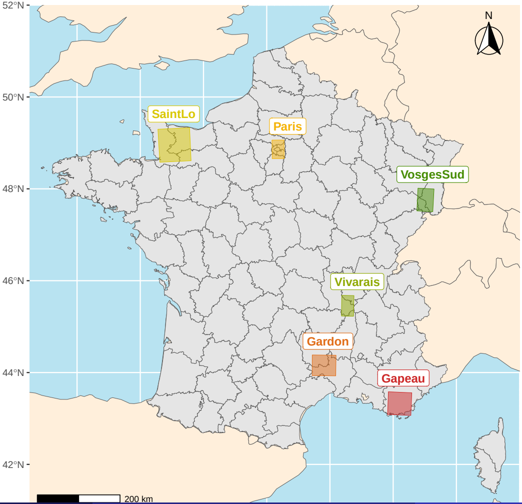
Given an observed dataset $R_t(x_i)$ (N pixels $\times$ T time steps), yielding observed time series IMF_t, WAR_t and CV_t, estimation proceeds as follows:

1. k_i^{WAR} estimated independently at each pixel (Bernoulli likelihood)
2. k_i^{IMF} estimated independently at each pixel (Log-Normal lkh + tricks)
3. A censored version of the Gaussian fields can then be computed $\Longrightarrow$ censored likelihood approach of Vaittinada Ayar et al. 2020
4. $a_{i,i}$ estimated independently at each pixel
5. $\{\lambda, b, \psi\}$ estimated from all pixels

Case Study

Data

COMEPHORE: 1h $\times$ 1km fields over continental France, 1997-present.
Merging of meteorological radar and rain gauge data (Tabary et al., 2012).

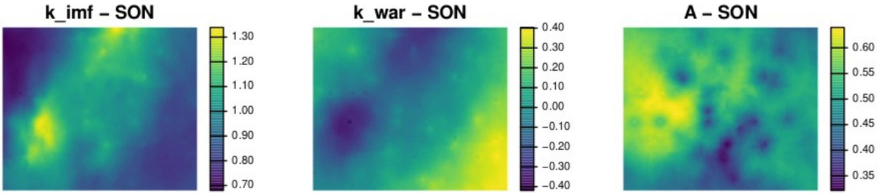


Case Study

Estimation

Exemple of estimated parameters k_i^{IMF}, k_i^{WAR} and $a_{i,i}$

Gardon area, fall season (SON)

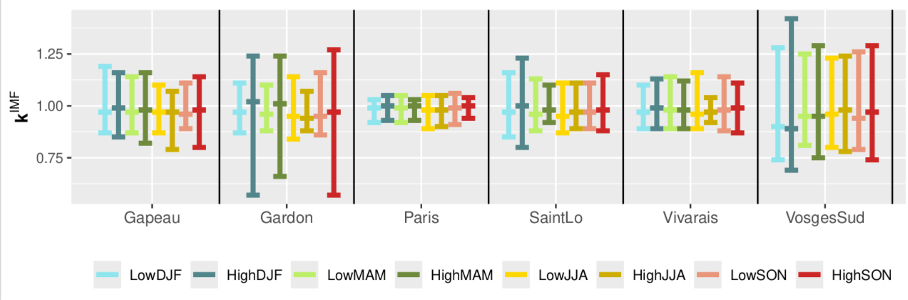


Estimation

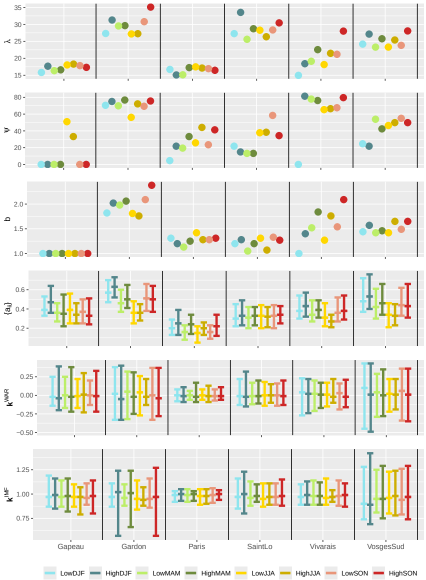
Need to cluster events into "rainfall types"

Here season $\times$ 2-group clustering of (IMF_t, WAR_t, CV_t)

Exemple for parameters k_i^{IMF}:

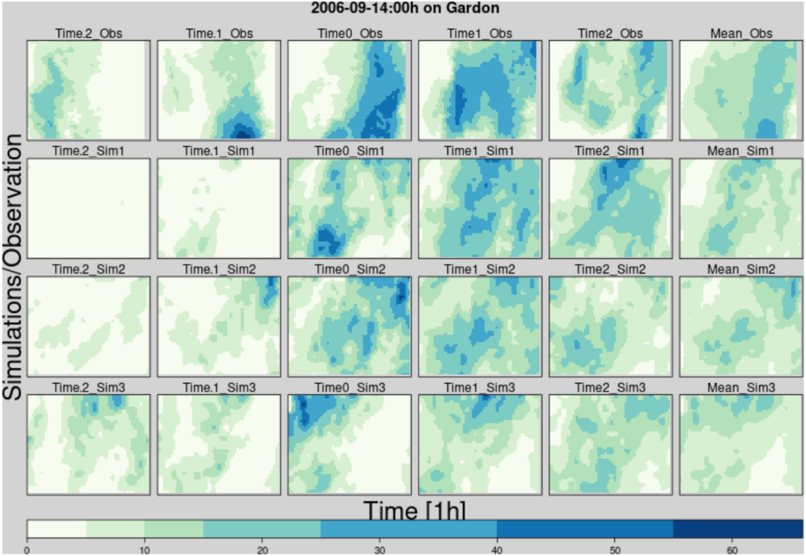


Estimation


λ
ψ
b
{a_ij}
k^WAR
k^IMF
Gapeau
Gardon
Paris
SaintLo
Vivarais
VosgesSud
LowDJF
HighDJF
LowMAM
HighMAM
LowJJA
HighJJA
LowSON
HighSON

Case Study

Example of simulated fields

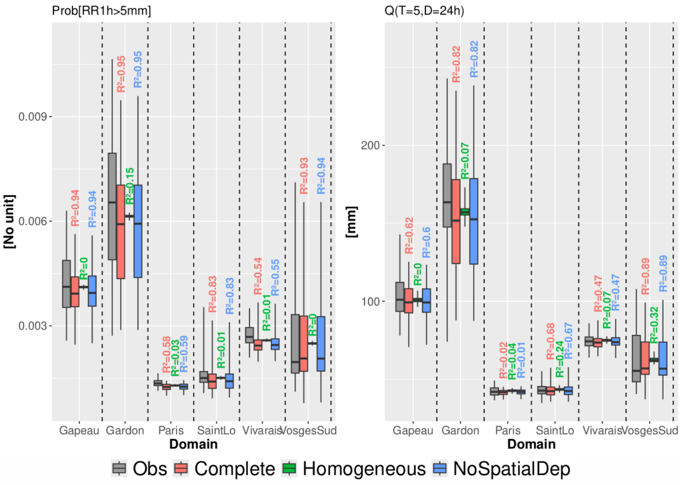


Case Study

Model checking

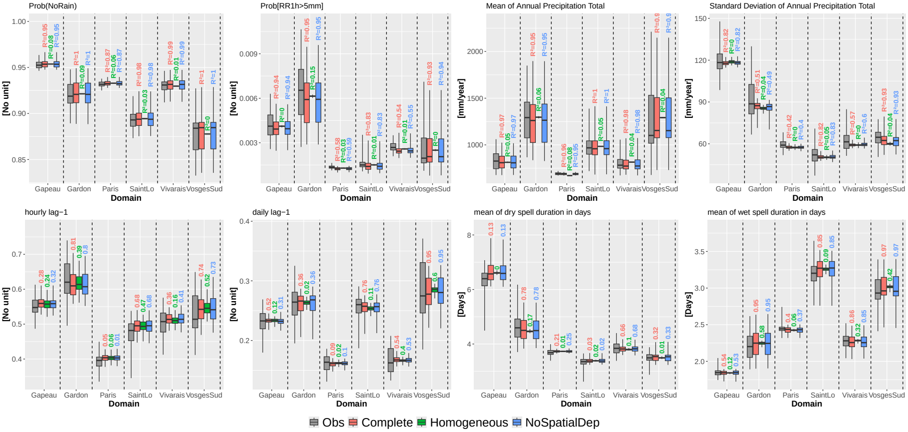
Occurrence and Extreme metrics for the six domains

Each boxplots represents spatial variability across N pixels



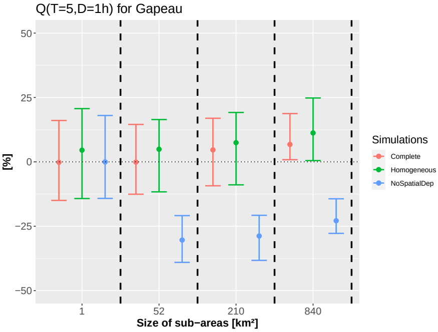
Model checking

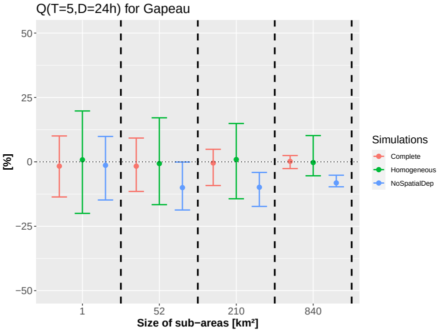
More metrics...



Model checking

Error (in %) between observed and simulated 5-year return levels for rainfall fields aggregated at various spatial scales





Case Study

Key findings

1. In many regions, the introduction of spatial heterogeneity is key to reproduce the spatial variability of hydrologically relevant rainfall metrics.
2. Spatial dependence is key to reproduce extremes of spatially-aggregated quantities (important for hydrology).
3. Overall, temporal persistence metrics are well reproduced (autocorrelations, duration of wet/dry spells, etc.)
4. Overall, properties of temporally-aggregated rainfall are better reproduced than those of hourly rainfall. This might be due to several reasons (choice of distribution, missing advection, ...)

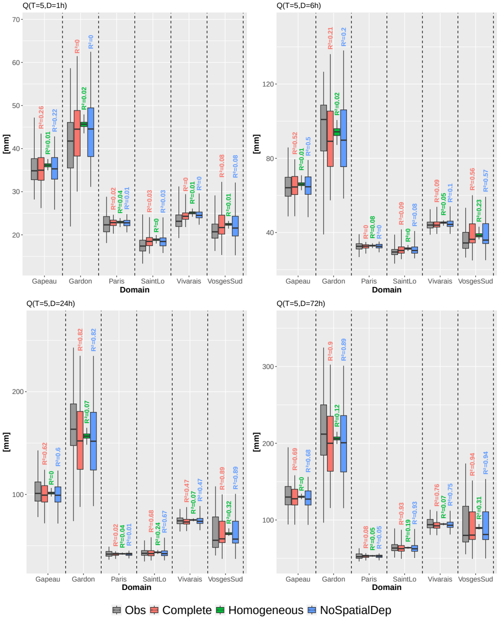
Conclusion

To be explored in SHARE:

1. Introduce advection
2. Link with EGPD
3. Link with tri-variate generator of (IMF_t, WAR_t, CV_t)
4. Develop validation metrics and strategy
5. Use distributed hydrologic model to validate on river streamflow

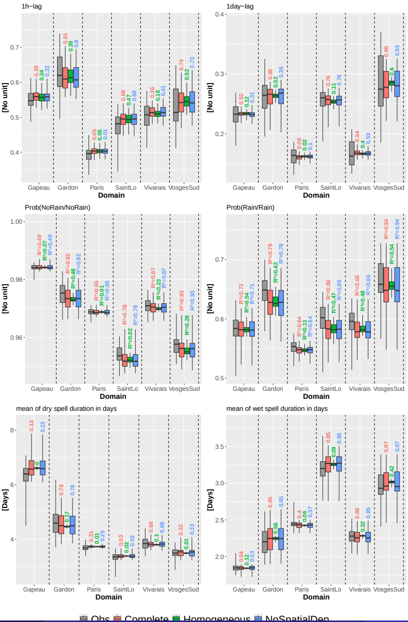
The End

Case Study



Q(T=5,D=1h)
Q(T=5,D=6h)
Q(T=5,D=24h)
Q(T=5,D=72h)
[mm]
Domain
Gapeau
Gardon
Paris
SaintLo
Vivarais
VosgesSud
Obs
Complete
Homogeneous
NoSpatialDep

Case Study



1h−lag
1day−lag
Prob(NoRain/NoRain)
Prob(Rain/Rain)
mean of dry spell duration in days
mean of wet spell duration in days
[No unit]
[Days]
Domain
Gapeau
Gardon
Paris
SaintLo
Vivarais
VosgesSud
Obs
Complete
Homogeneous
NoSpatialDep

Case Study

Spatial correlation plot for hourly precipitation