Stochastic rainfall generation over complex topography

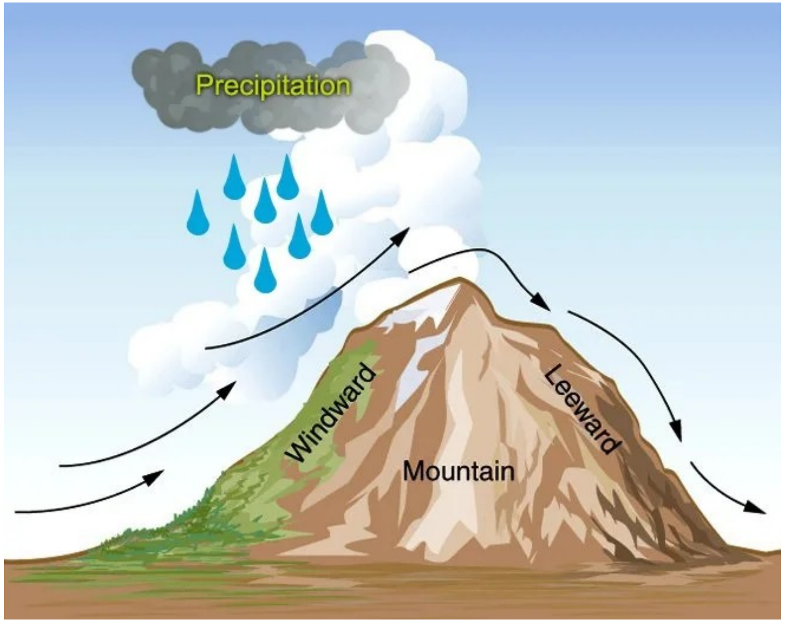
Lionel Benoit, Denis Allard

Matthew Lucas, Keri Kodama, Thomas Giambelluca

INRAE BIOSP GEOLEARNING CHAIRE /// Data Science for the Environment

UNIVERSITY of HAWAI'I at MĀNOA WATER RESOURCES RESEARCH CENTER

Orographic rainfall:



Orographic precipitation is ubiquitous:

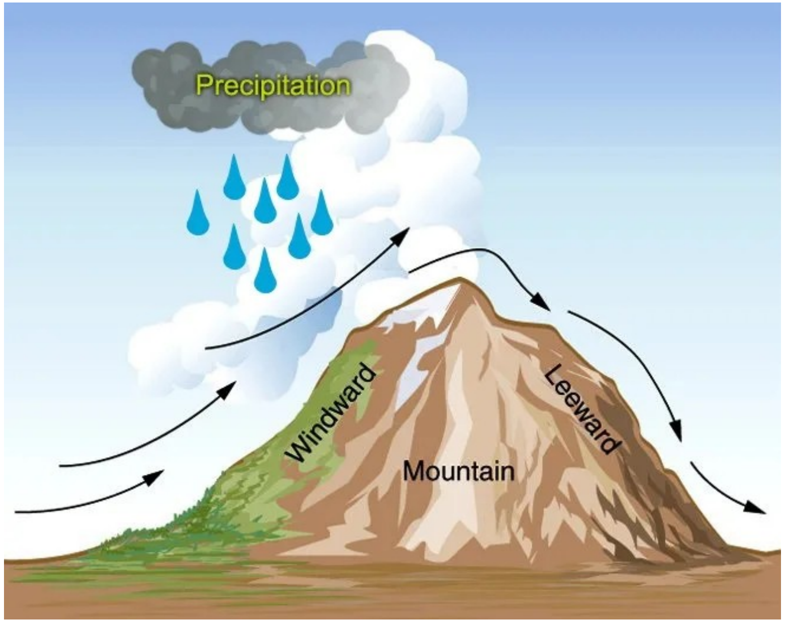
→ *Mediterranean climate: Sierra Navada (Spain & USA)*

→ *Temperate climate: Alps*

→ *Tropical climate: Andes & High tropical islands*

Introduction (1/3): orographic rainfall

Orographic rainfall:



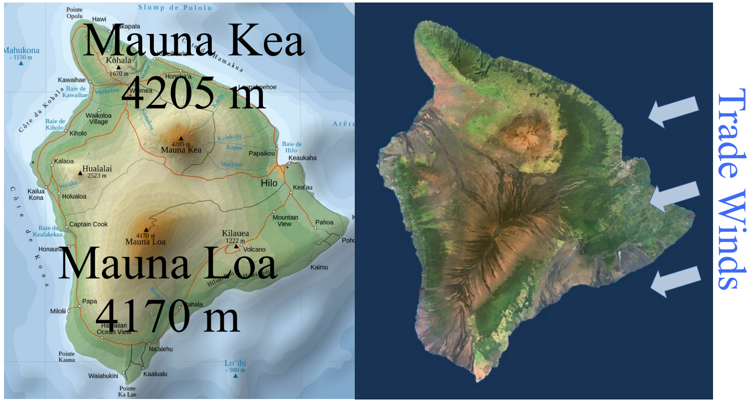
And impacts rainfall statistics:

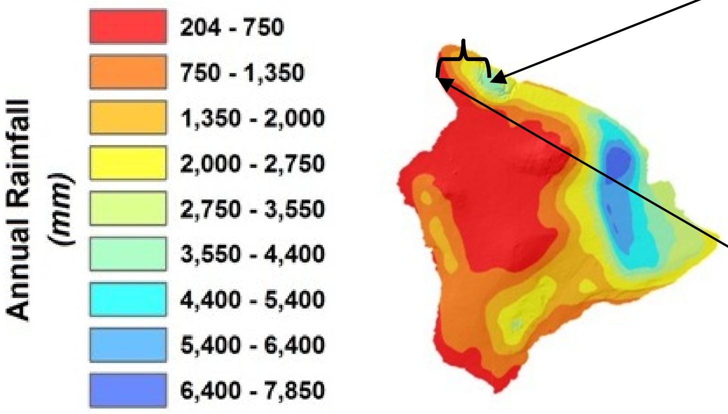
→ *Rainfall statistics vary in space (occurrence, intensity, spatial dependence)*

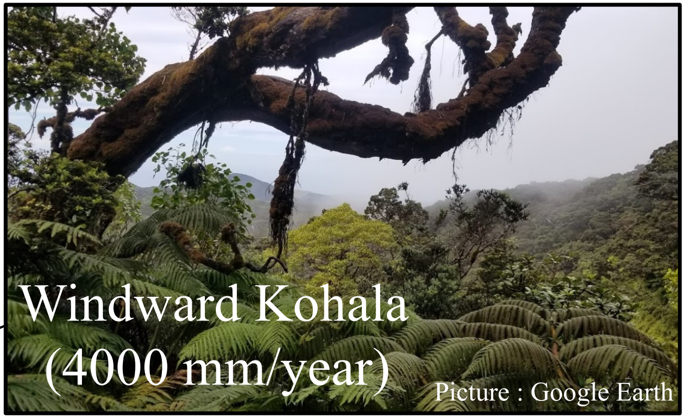
→ *No straightforward link with covariates (elevation, slope, weather)*

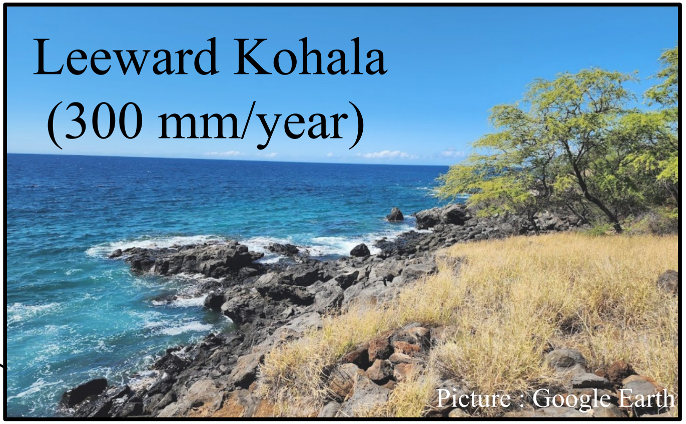
Introduction (2/3): the example of Hawai‘i Island

Hawai‘i is a textbook study area for orographic rainfall:









Distance = 20 km

Introduction (3/3): objectives of the study

Stochastic rainfall model accounting for orographic effects

→ *Non-stationary spatial model to capture the variation of rainfall statistics*

→ *Fully non-stationary: marginal distribution and spatial dependencies*

Data-driven inference of non-stationary parameters

→ *No use of covariates in the spatial model*

→ *Leverage geostatistics to learn a continuous spatial model from sparse obs.*

Set-up

→ *Daily resolution and focus on spatial statistics*

→ *Designed for Hawai‘i, but (hopefully) applicable to any mountain range*

Rainfall model (1/4): overview

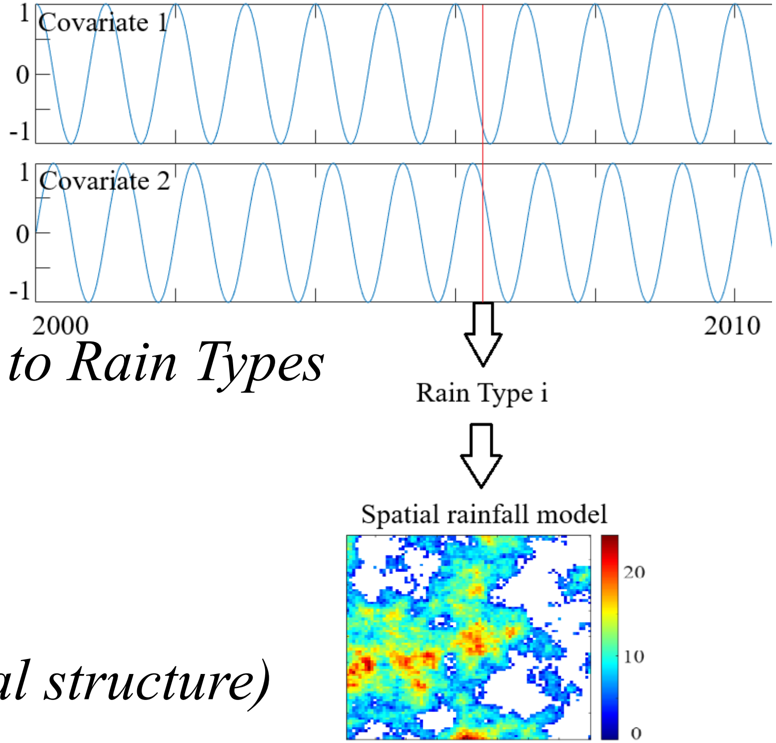
Split modeling into two separate components: temporal and spatial

Temporal model [very basic]

→ *The time line is clustered into Rain Types*

→ *Rain type occurrence conditioned to covariates*

→ *Daily rainfall fields are independent conditional to Rain Types*

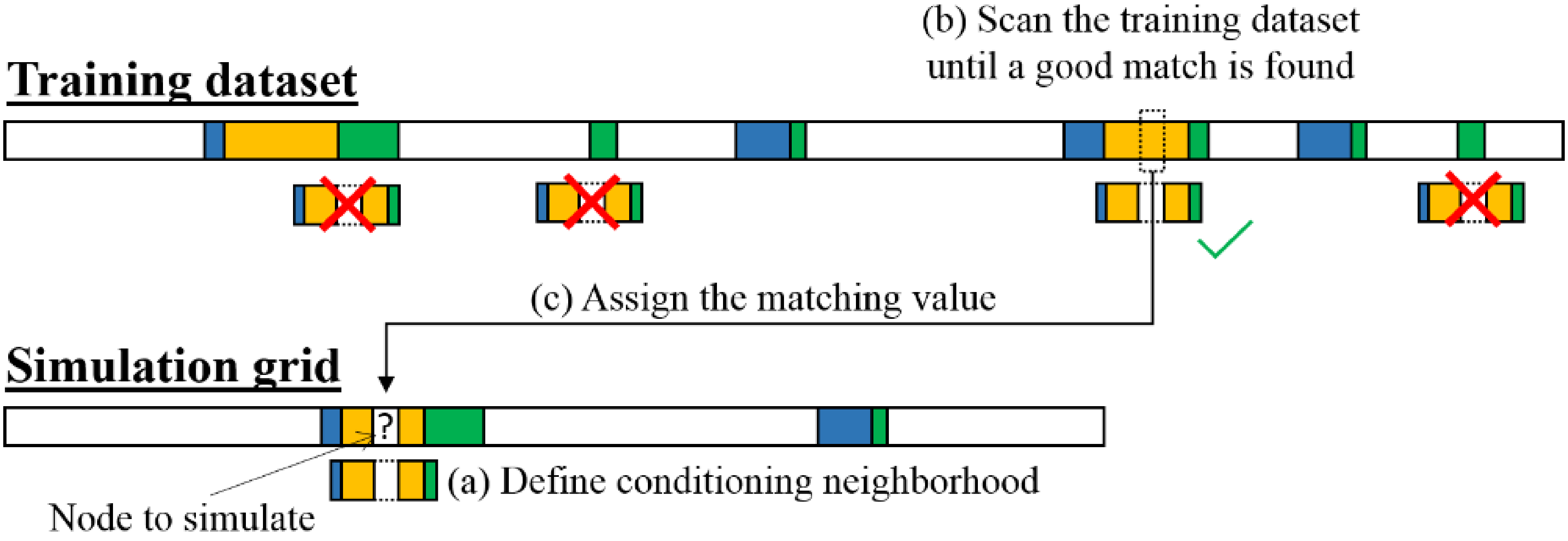


Spatial model [focus of the talk, more comprehensive]

→ *Trans-Gaussian geostatistics*

→ *Fully non-stationary model (marginal and spatial structure)*

Rain type time series are simulated by Multiple-Points Statistics simulation (MPS ≈ pattern-based resampling of a training dataset)



From Benoit et al. (2020) Nonstationary stochastic rain type generation: accounting for climate drivers.

Rainfall model (3/4): trans-Gaussian geostatistics

Trans-Gaussian geostatistics split the rain signal (R) in two components:

$Y \sim MVN(0,1,C_Y)$ and $R = \psi(Y)$

Latent field:

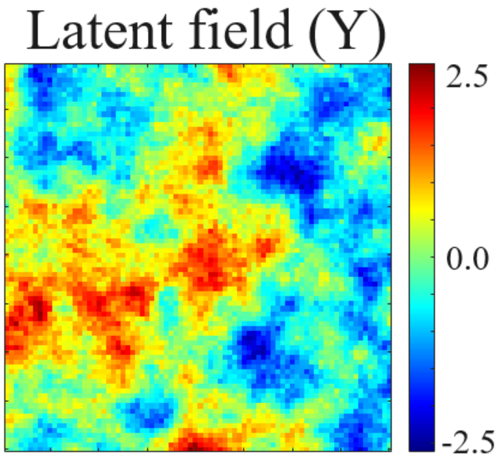
$Y \sim MVN(0,1,C_Y)$

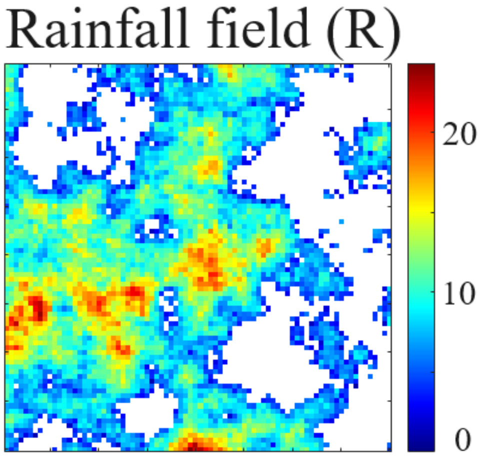
Models spatial dependencies through the covariance function C_Y

Transform function:

$R = \psi(Y)$

Models rainfall occurrence and intensity





Rainfall model (3/4): trans-Gaussian geostatistics

Trans-Gaussian geostatistics split the rain signal (R) in two components:

$Y \sim MVN(0,1,C_Y)$ and $R = \psi(Y)$

Latent field:
Models spatial dependencies through the covariance function C_Y

Transform function:
Models rainfall occurrence and intensity

C_Y parameterized by a Matérn covariance function

$$C_Y(||\mathbf{h}||;\nu,\rho) = \frac{1}{\Gamma(\nu)2^{\nu-1}}\left(\frac{||\mathbf{h}||}{\rho}\right)^{\nu}\mathcal{K}_\nu\left(\frac{||\mathbf{h}||}{\rho}\right)$$

ψ parameterized by the mixture of an atom of zeros (truncation) and a Gamma distribution

$$R(\mathbf{s}) = 0 \quad \text{if} \quad Y(\mathbf{s}) \leq a_0$$
$$R(\mathbf{s}) = \text{Gamma}^{-1}(\Phi(Y(\mathbf{s}));k,\theta) \quad \text{if} \quad Y(\mathbf{s}) > a_0$$

Rainfall model (4/4): making the spatial model non-stationary

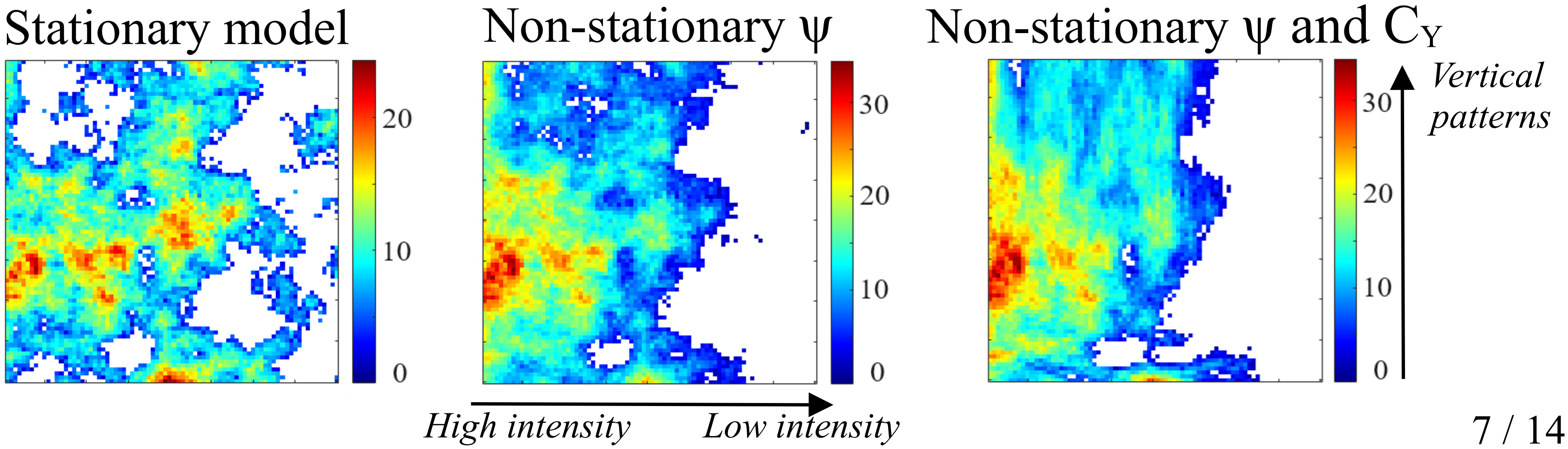
Non-stationary model to capture the spatial variation of rain statistics:

=> Model parameters are made location-dependent:

→ ψ => $\psi_{\mathbf{s}}$ *and* C_Y => C_{Ys} *(with* **s** *the location of interest)*

→ C_{Ys} *requires a valid model of non-stationary covariance*

[*Paciorek & Schervish, 2006, spatial modelling using a new class of non-stationary covariance functions, Environmetric, 17:483-506*]



Rainfall model (4/4): making the spatial model non-stationary

Non-stationary model to capture the spatial variation of rain statistics:

=> Model parameters are made location-dependent:

→ ψ => $\psi_{\mathbf{s}}$ *and* C_Y => C_{Ys} *(with* **s** *the location of interest)*

→ C_{Ys} *requires a valid model of non-stationary covariance*

[*Paciorek & Schervish, 2006, spatial modelling using a new class of non-stationary covariance functions, Environmetric, 17:483-506*]

$$C_Y(\mathbf{s}_i, \mathbf{s}_j) = \frac{2^{1-(\nu_i+\nu_j)/2}}{\sqrt{\Gamma(\nu_i)\Gamma(\nu_j)}} |\Sigma_i|^{\frac{1}{4}} |\Sigma_j|^{\frac{1}{4}} \left|\frac{\Sigma_i + \Sigma_j}{2}\right|^{-\frac{1}{2}} \left(\sqrt{Q_{ij}}\right)^{\frac{\nu_i+\nu_j}{2}} \mathcal{K}_{\frac{\nu_i+\nu_j}{2}}\left(\sqrt{Q_{ij}}\right)$$

$$\text{with} \quad \Sigma_i = \mathbf{V}_i \times \mathbf{\Lambda}_i \times \mathbf{V}_i^T; \quad \mathbf{V}_i = \begin{bmatrix} \frac{\gamma_{1,i}}{\sqrt{\gamma_{1,i}^2+\gamma_{2,i}^2}} & -\frac{\gamma_{2,i}}{\sqrt{\gamma_{1,i}^2+\gamma_{2,i}^2}} \\ \frac{\gamma_{2,i}}{\sqrt{\gamma_{1,i}^2+\gamma_{2,i}^2}} & \frac{\gamma_{1,i}}{\sqrt{\gamma_{1,i}^2+\gamma_{2,i}^2}} \end{bmatrix}; \quad \mathbf{\Lambda}_\mathbf{i} = \begin{bmatrix} \gamma_{1,i}^2 + \gamma_{2,i}^2 & 0 \\ 0 & \lambda_{2,i} \end{bmatrix}$$

$$\text{and with} \quad Q_{ij} = (\mathbf{s}_i - \mathbf{s}_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2}\right)^{-1} (\mathbf{s}_i - \mathbf{s}_j)$$

In practice (1/2): model calibration

Rain types are delineated by GMM clustering based on rain gauge obs.

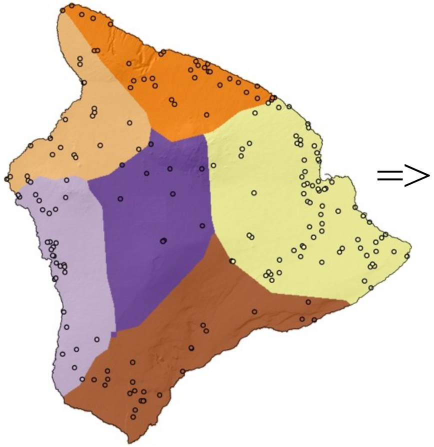
Parameters of the spatial model are estimated by likelihood maximization

→ *Overall 9 parameters (stationary case)*

Estimation of model parameters from sparse observations

→ *Marginal distribution (ψ_s): estimation at gauge locations + Ordinary Kriging*

→ *Covariance function (C_{Y_S}): estimation within climate division + Spline interp.*



Hawai'i climate divisions

=> hypothesis of local stationarity of the covariance

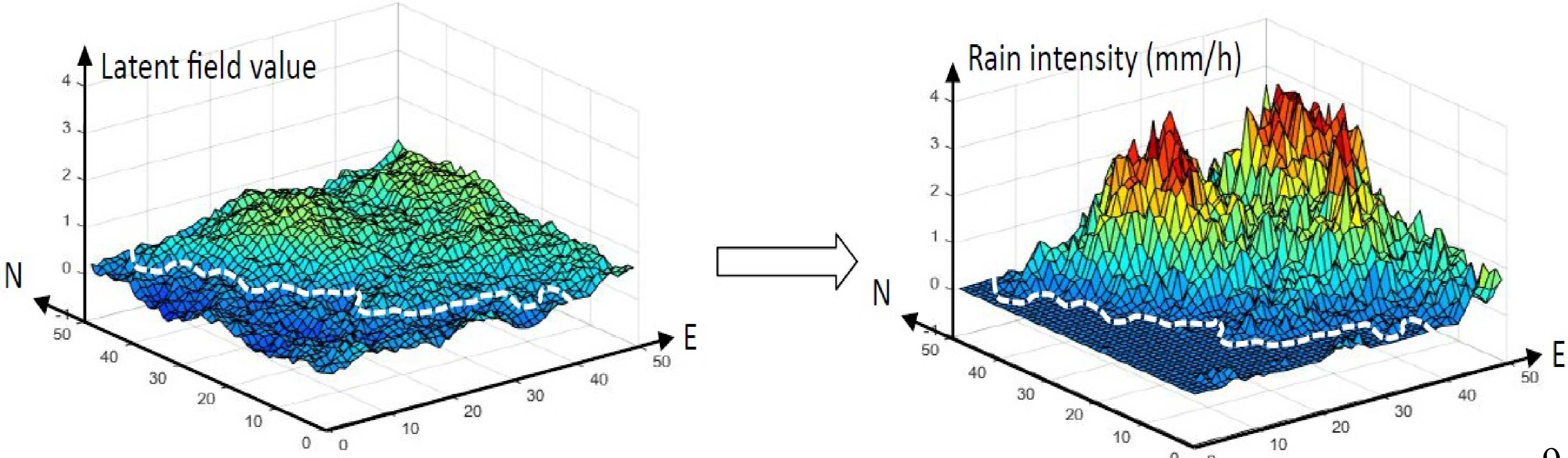
In practice (2/2): stochastic rainfall generation

(1) Unconditional simulation of the latent field (Y)

→ *Cholesky decomposition of the covariance matrix (small simulation grids)*

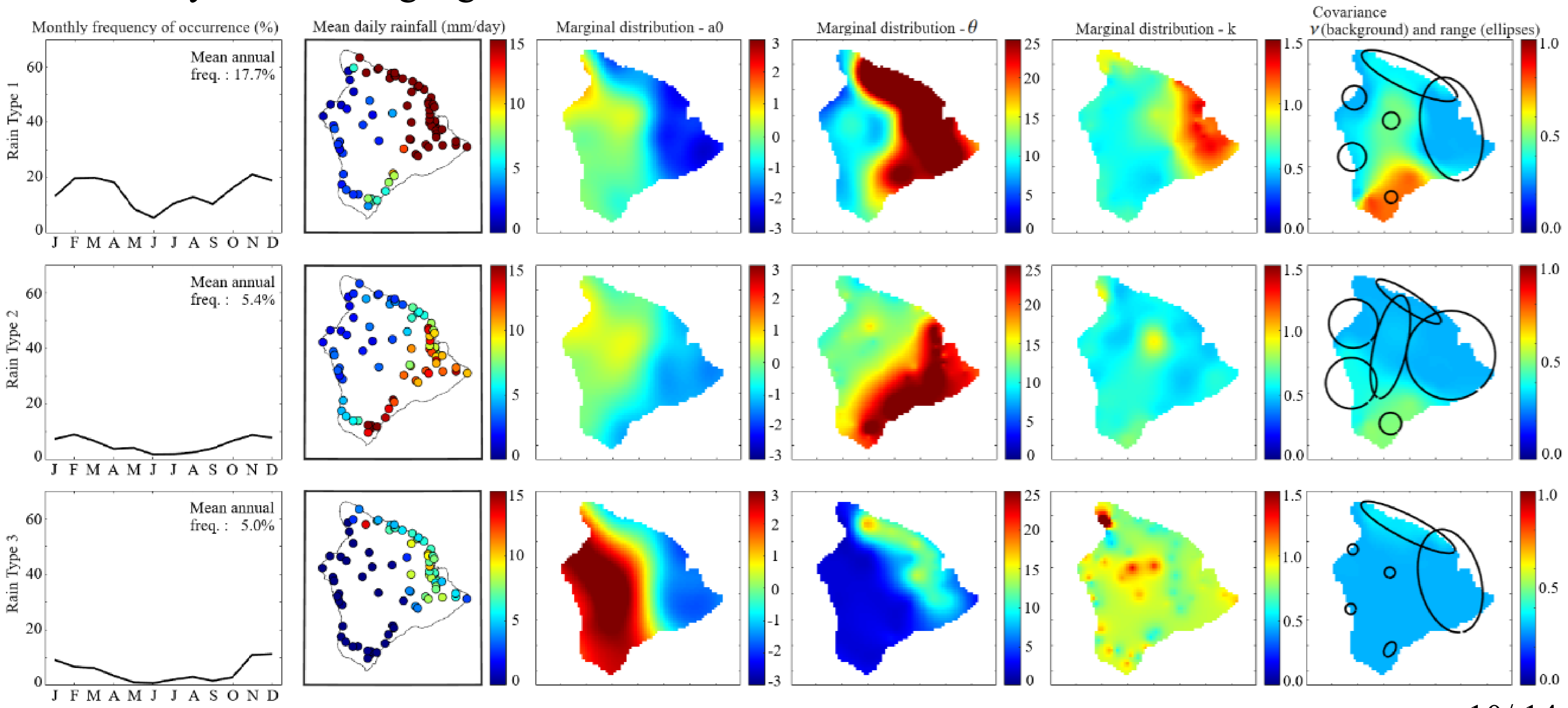
→ *Spectral simulation - Gaussian mixture approach (large simulation grids) [cf. Talk Denis Allard on Tuesday].*

(2) Transformation of the latent field into rain intensities: $R(\mathbf{s}) = \Psi(Y(\mathbf{s}))$



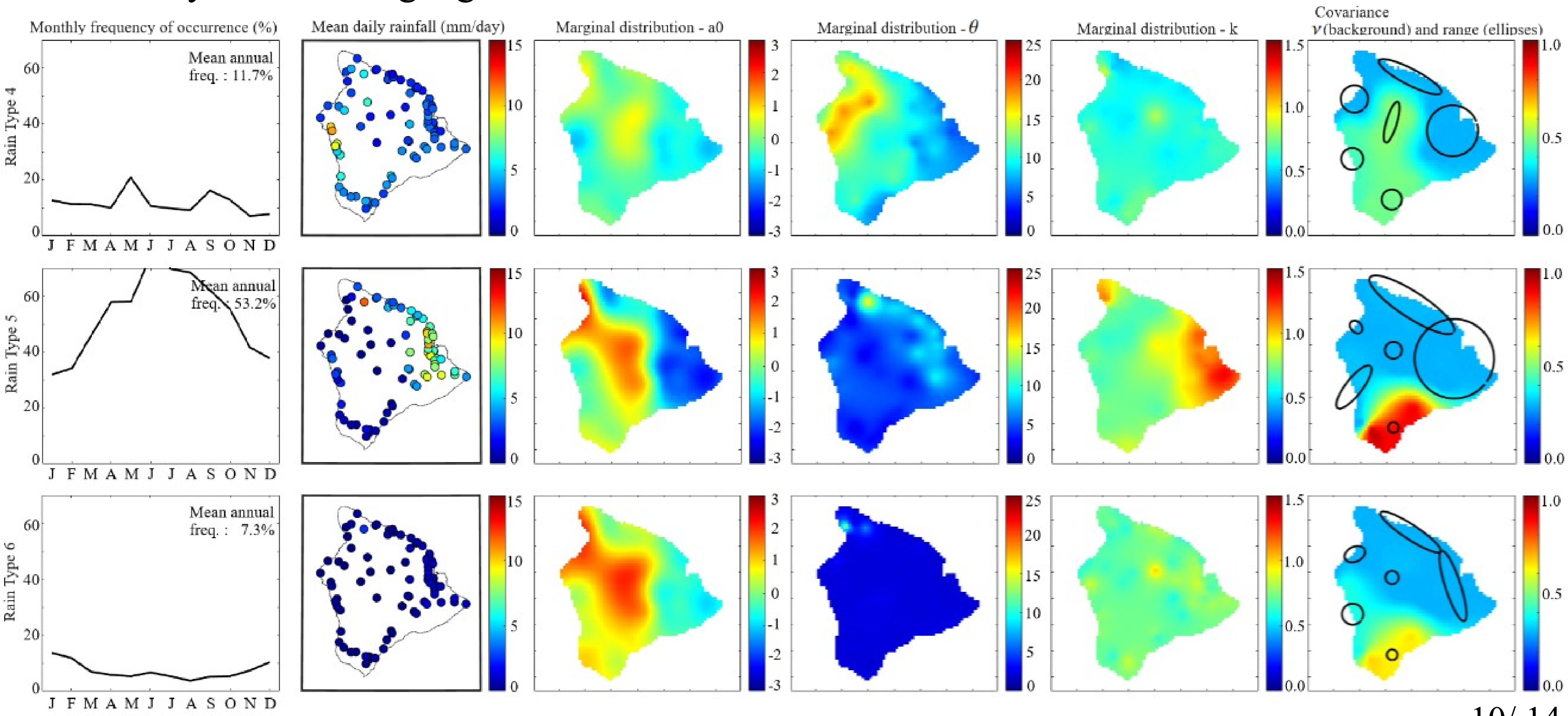
Results (1/4): The 6 Rain types of Hawai‘i (and associated parameters)

From 20 years of rain gauge observations : 2000 - 2019

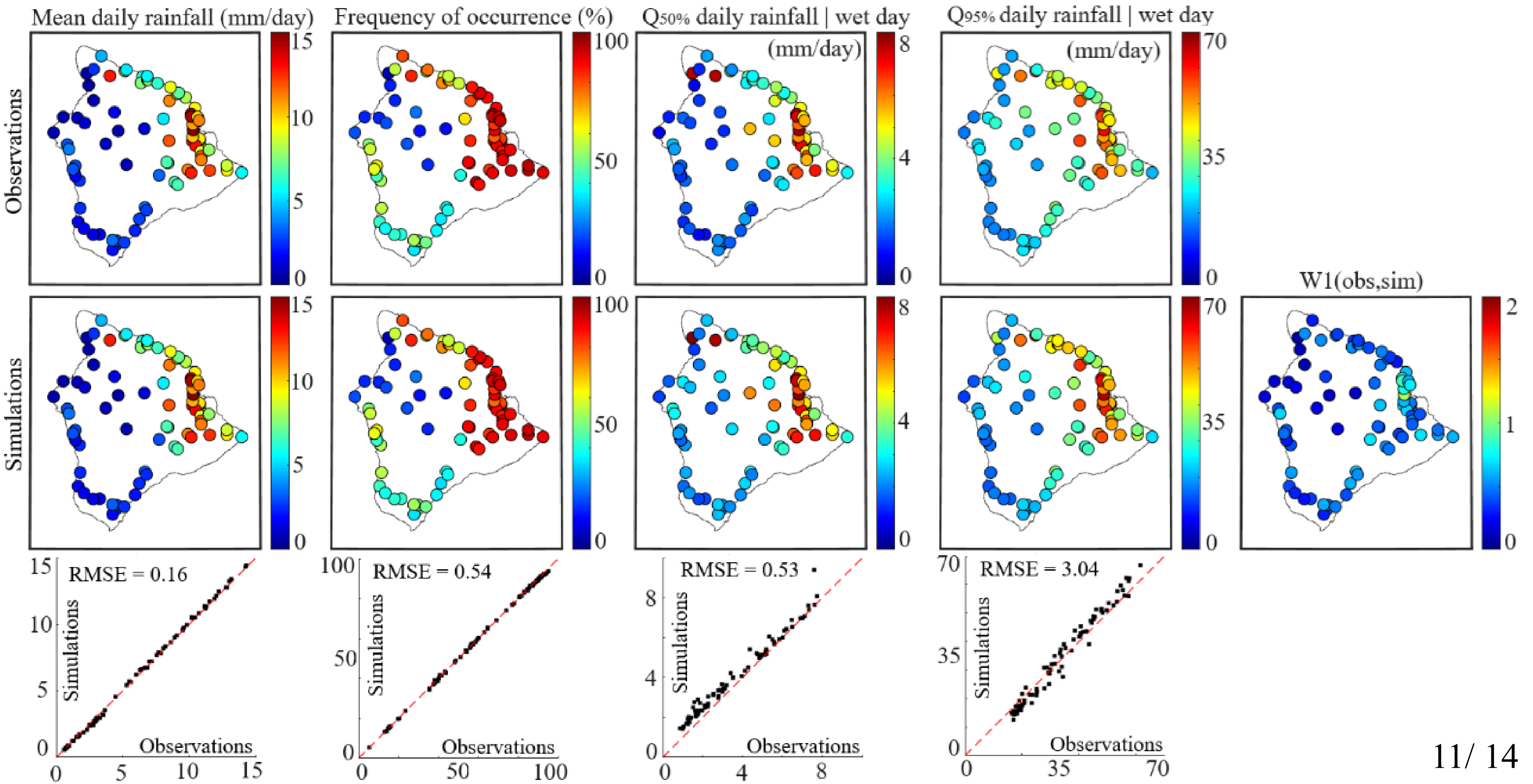


 The 6 Rain types of Hawai‘i (and associated parameters)

From 20 years of rain gauge observations : 2000 - 2019

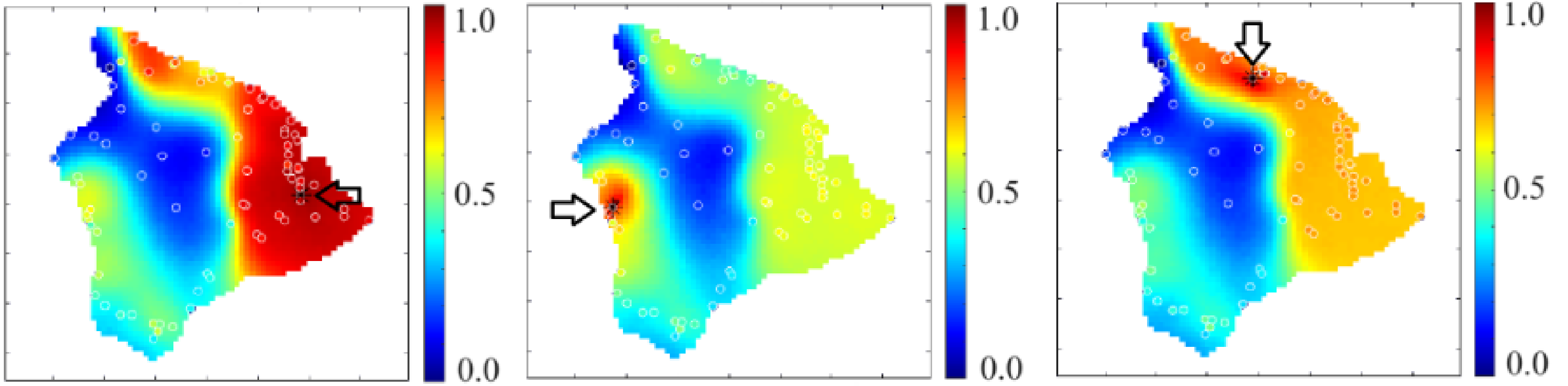


Results (2/4): Point-scale rainfall statistics

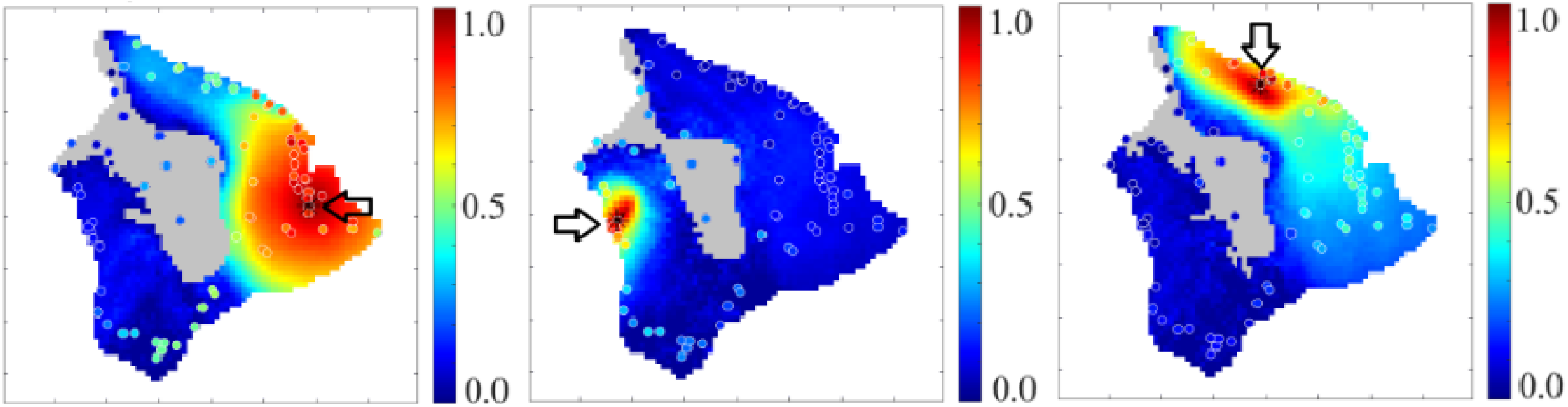


Results (3/4): Spatial dependencies within rainfall fields

Spatial patterns of rain occurrence: Jaccard index with point *

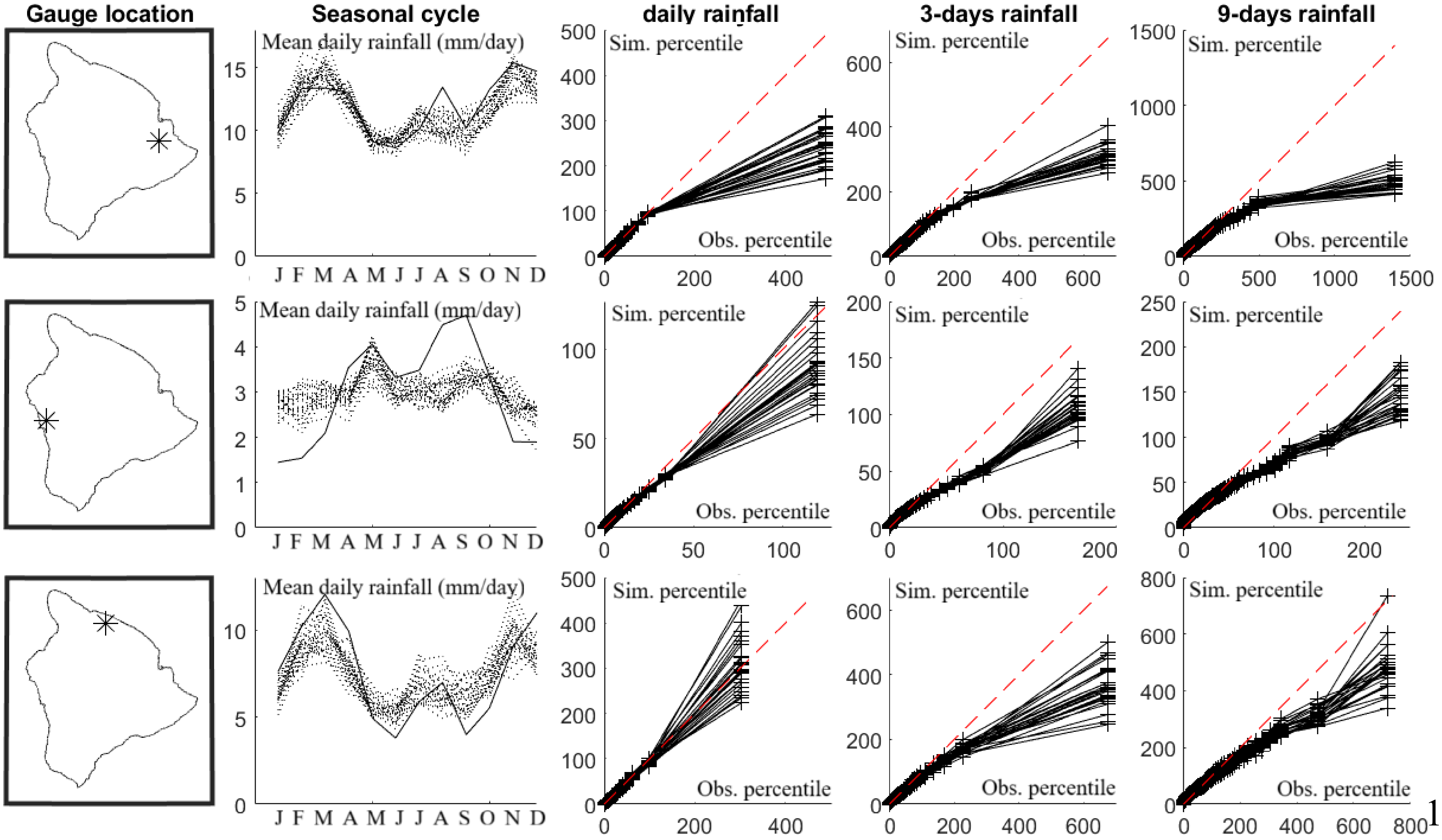


Spatial patterns of rain intensity: Paerson correlation with point *



Background color = simulations, Dots = observations

Results (4/4): Spatial variability of the temporal behavior of rainfall



Results (4/4): Spatial variability of the temporal behavior of rainfall

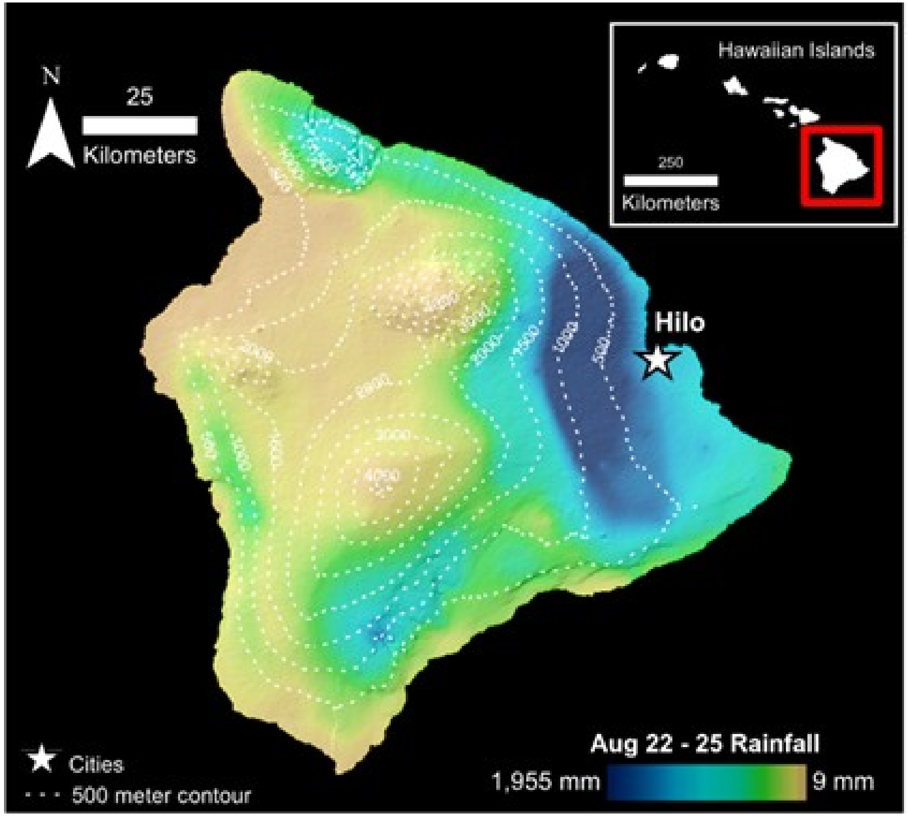
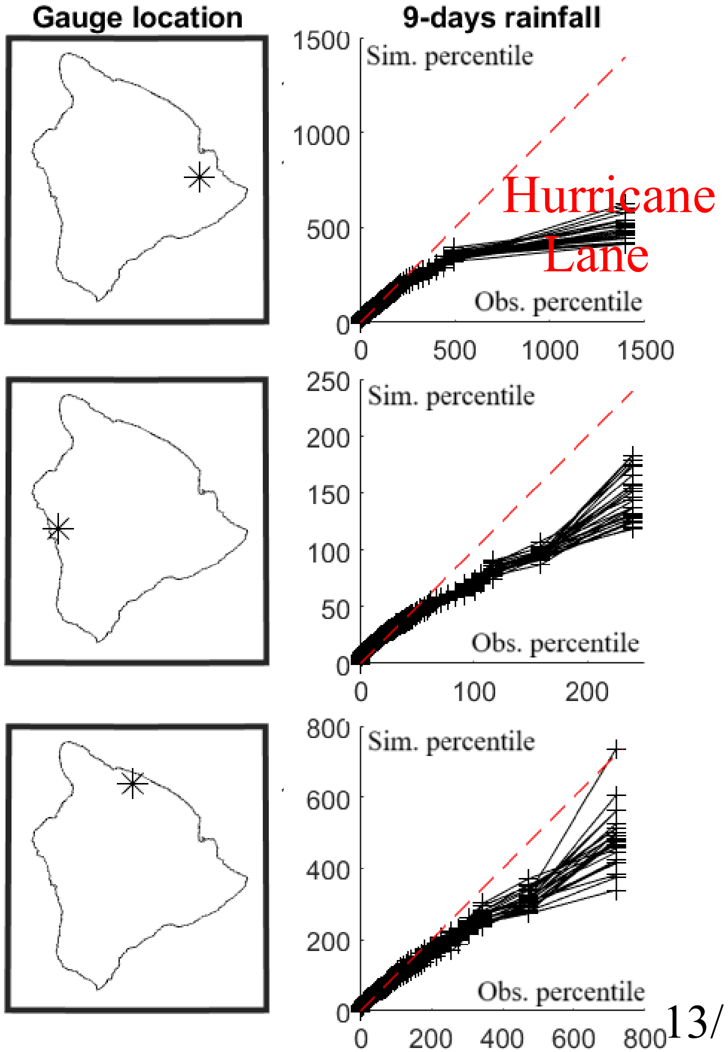


Fig. 5. Total accumulated rainfall (mm) during Hurricane Lane, from 22 to 25 Aug 2018 on the island of Hawai'i. Hillshade is included along with white dashed contours indicating terrain height every 500 m.

From Nugent et al. (2020) Fire and Rain: The Legacy of Hurricane Lane in Hawai'i



Conclusion & perspectives:

Non-stationary trans-Gaussian geostatistical simulations:

→ *Reproduce the statistical signature of daily rainfall in mountains*

→ *If a dense rain gauge network is available can learn spatial non-stationarity in a data-driven way (no need for covariates)*

=> Very flexible model for daily rainfall fields

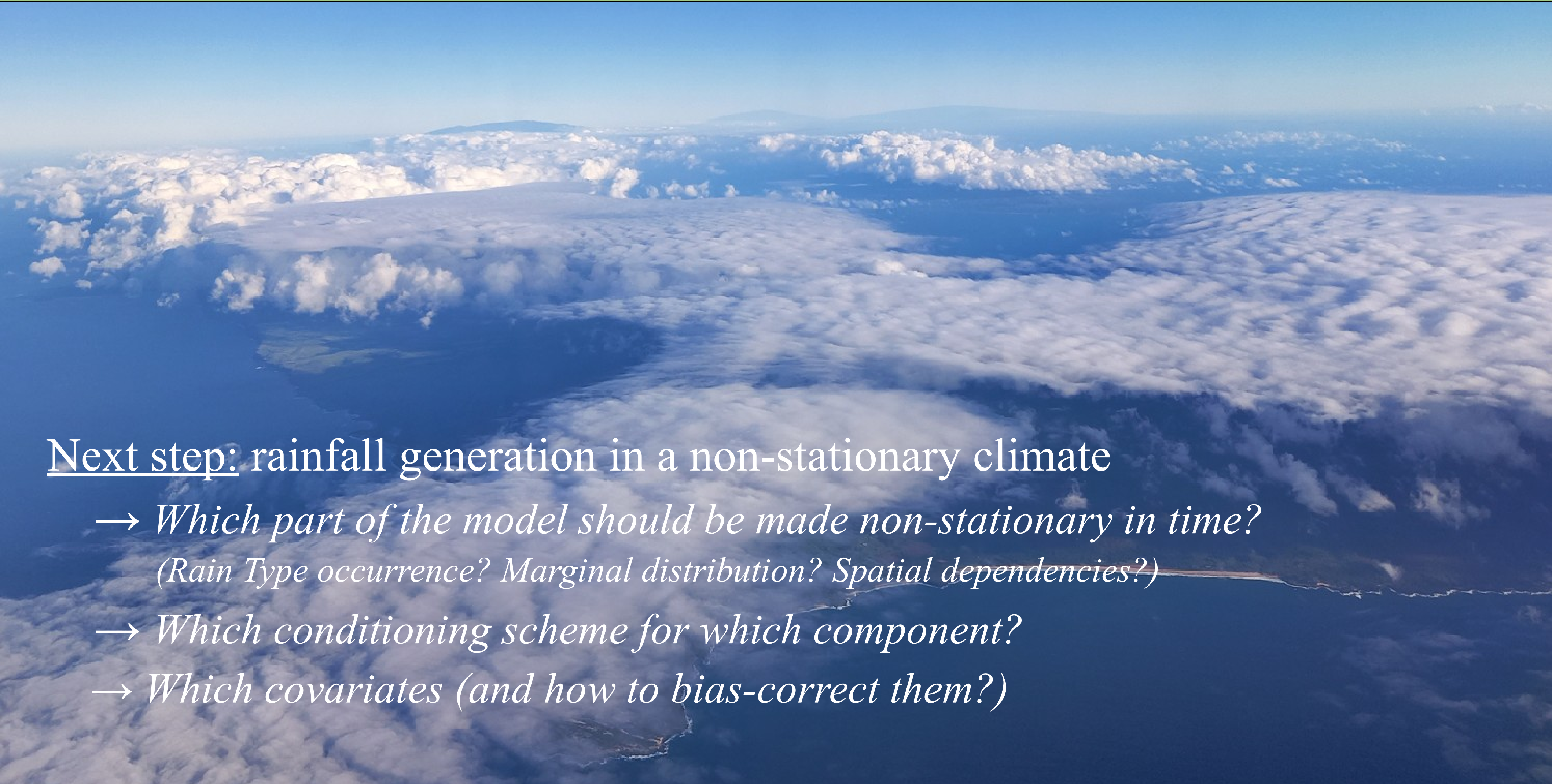
Next step: rainfall generation in a non-stationary climate

→ *Which part of the model should be made non-stationary in time?*
(Rain Type occurrence? Marginal distribution? Spatial dependencies?)

→ *Which conditioning scheme for which component?*

→ *Which covariates (and how to bias-correct them?)*

Thank you for your attention :-)



<u>Next step:</u> rainfall generation in a non-stationary climate

→ *Which part of the model should be made non-stationary in time?*

(Rain Type occurrence? Marginal distribution? Spatial dependencies?)

→ *Which conditioning scheme for which component?*

→ *Which covariates (and how to bias-correct them?)*