Improving rainfall gradients modeling by conditioning daily rainfall maps to monthly totals

Lionel Benoit, Denis Allard

Matthew Lucas, Keri Kodama, Thomas Giambelluca

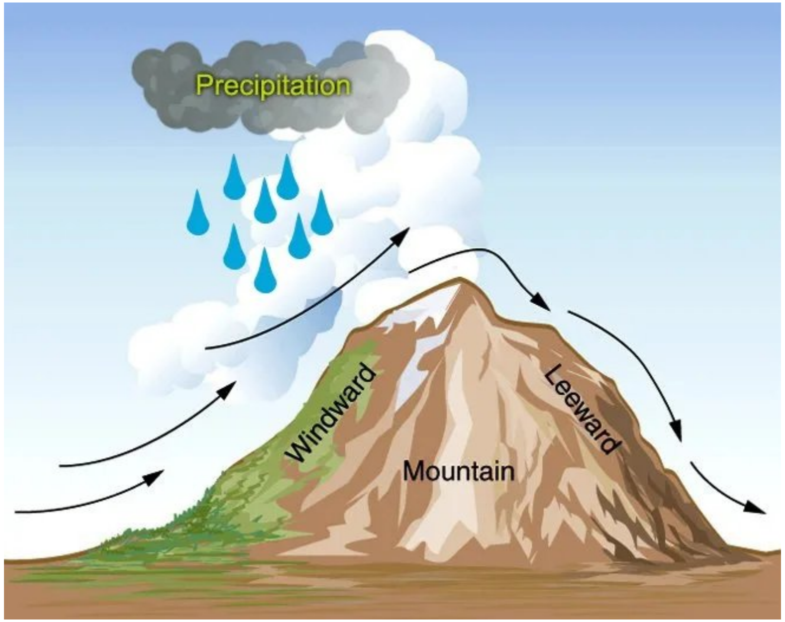
INRAE BIO/Π

UNIVERSITY of HAWAI'I at MĀNOA
WATER RESOURCES RESEARCH CENTER

April 2025 - Workshop GEOLEARNING CHAIRE /// Data Science for the Environment with support from ANDRA BNP PARIBAS CCR SCOR FONDATION POUR LA SCIENCE

Introduction (1/5): orographic rainfall leads to strong spatial gradients

Orographic rainfall:



Picture: Wikipedia

Orographic precipitation is ubiquitous:

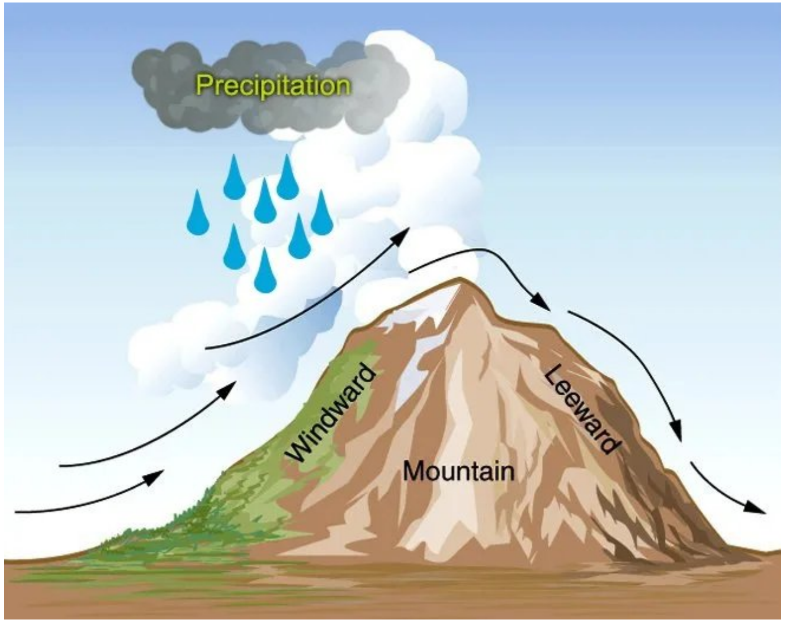
→ *Mediterranean climate: Sierra Navada (Spain & USA)*

→ *Temperate climate: Alpes*

→ *Tropical climate: Andes & High tropical islands*

Introduction (1/5): orographic rainfall leads to strong spatial gradients

Orographic rainfall:



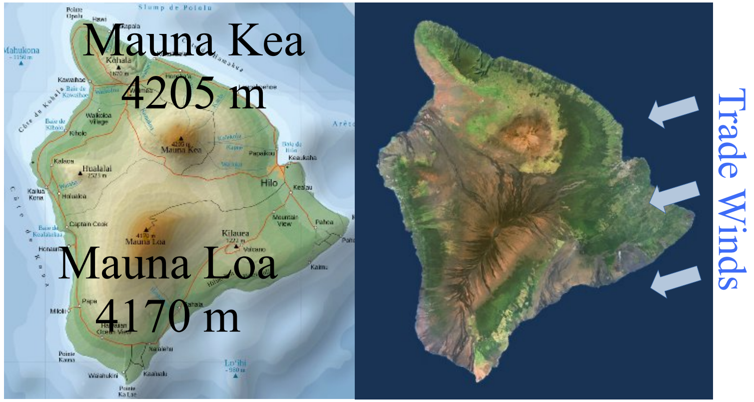
And impacts rainfall statistics:

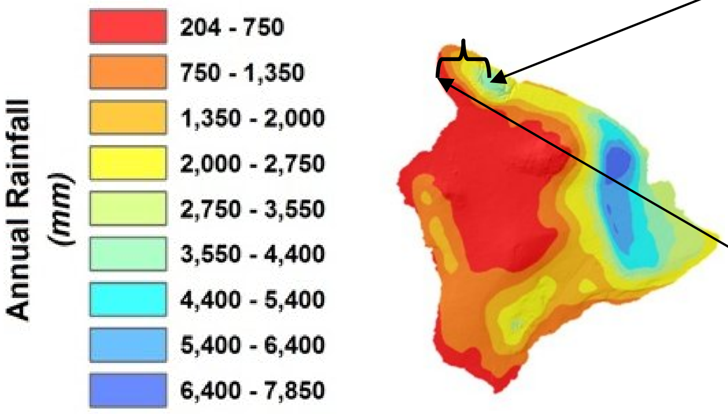
→ *Rainfall statistics vary in space (occurrence, intensity, dependencies)*

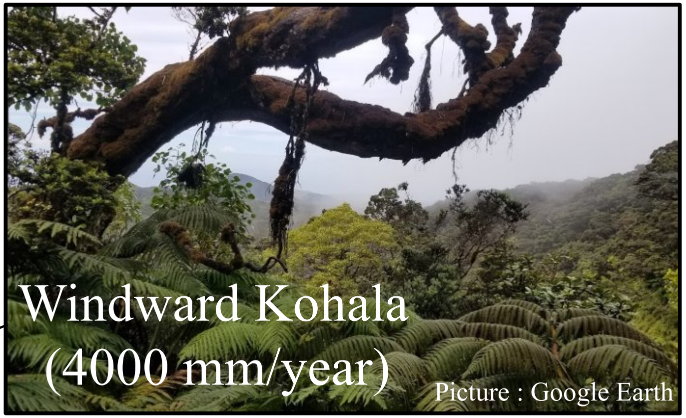
→ *Not simply related to covariates (e.g., elevation, slope, weather)*

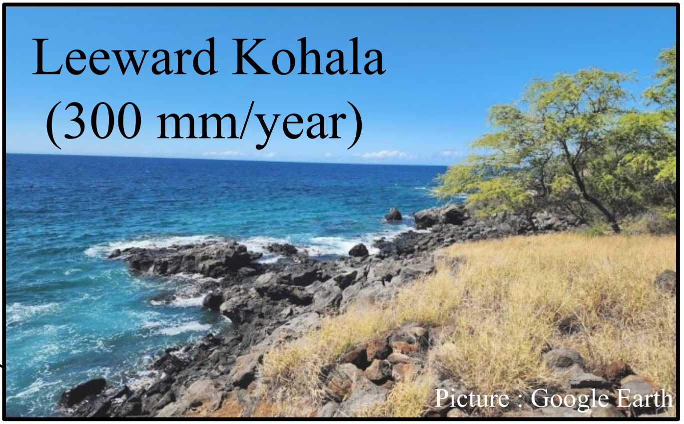
Introduction (2/5): the example of Hawai‘i Island

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







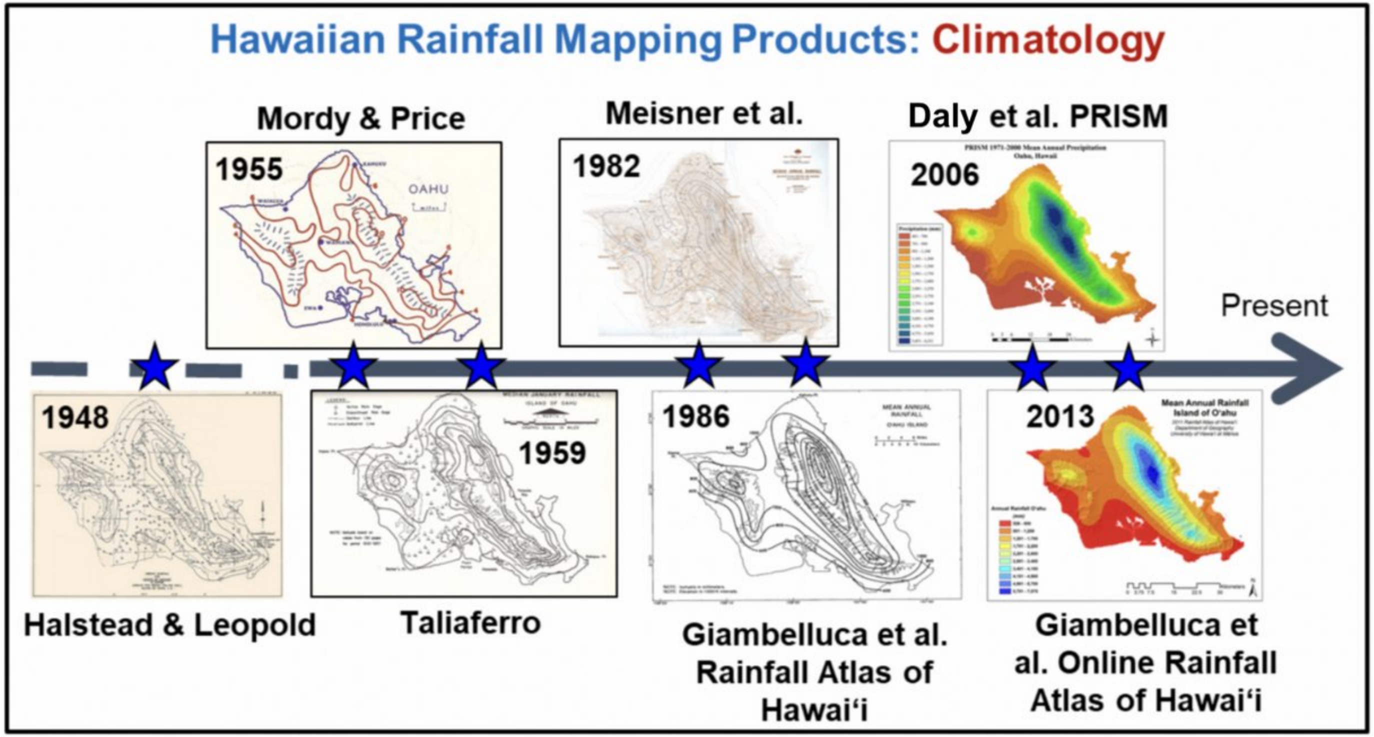


Distance = 20 km

Climatological rainfall maps:

→ *Long time series of rainfall observations (rain gauges only)*

→ *Vegetation proxies to complement direct observations in poorly gauged areas*

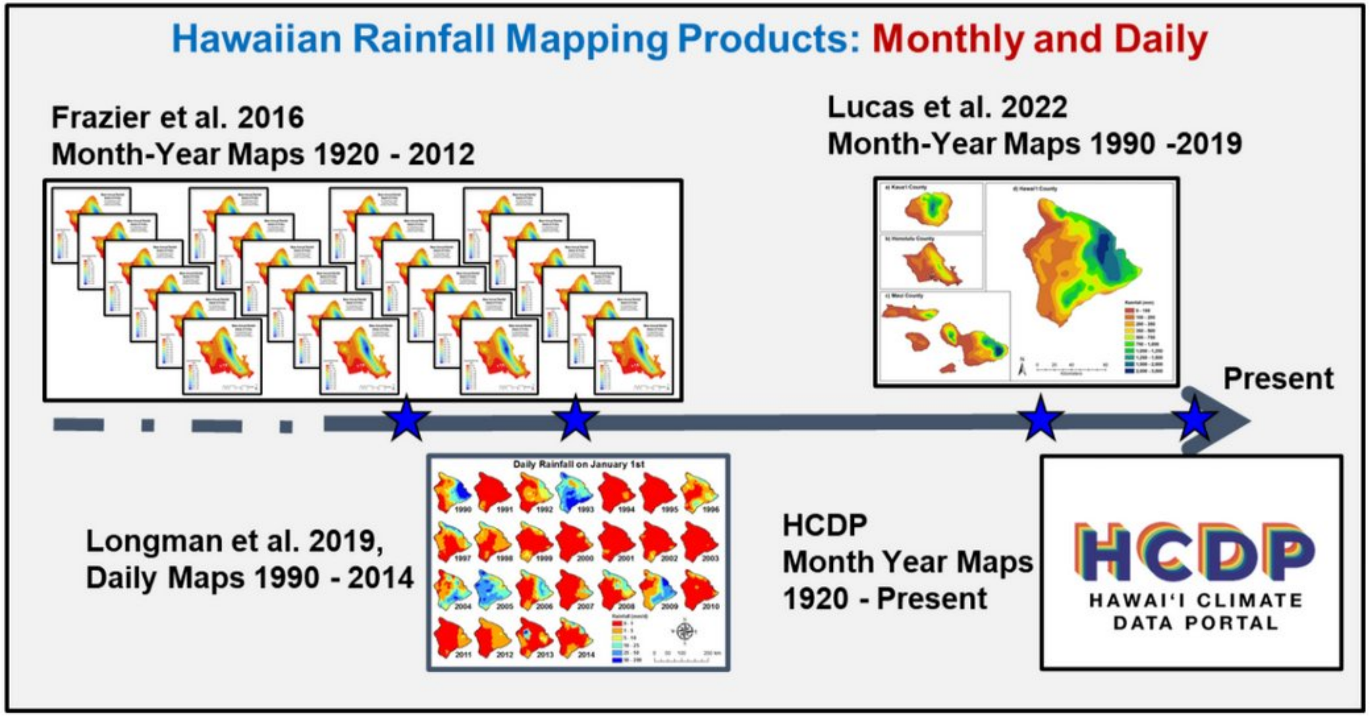


Picture : www.hawaii.edu/climate-data-portal

Monthly and daily rainfall maps:

→ *Input data = rain gauge observations + climatological maps*

→ *Climatologically aided interpolation (i.e., KED with drift from the climatology)*



Picture : www.hawaii.edu/climate-data-portal

Introduction (5/5): objectives of this work

Improve uncertainty quantification

→ *Replace Kriging by conditional simulations*

Account for non-stationary rain statistics

→ *Spatial model accounting for non-stationary marginals and dependencies*

→ *Parameter inference from rain gauge observations (spatially sparse)*

Condition daily rainfall maps to monthly totals

→ *Better capture the spatial gradients in poorly gauged areas*

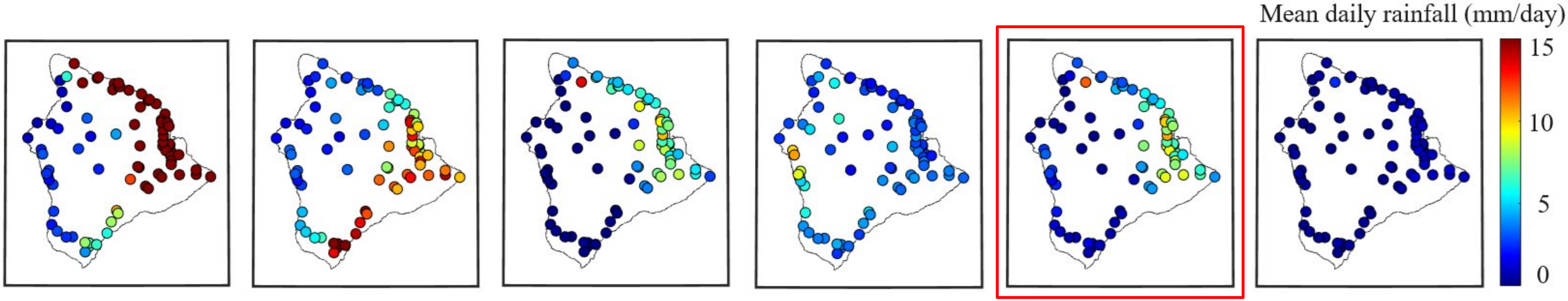
→ *Ensure consistency between monthly and daily maps*

Rainfall model (1/5): rain types

Orographic effects depend on atmospheric conditions:

→ Days with similar rainfall are pooled into **rain types** and **processed separately.**

→ One [purely spatial] stochastic model is set-up for each rain type.



Trade wind conditions (used for illustration)

=> shallow convection & distinctive orographic effects

=> around 50% of the dataset

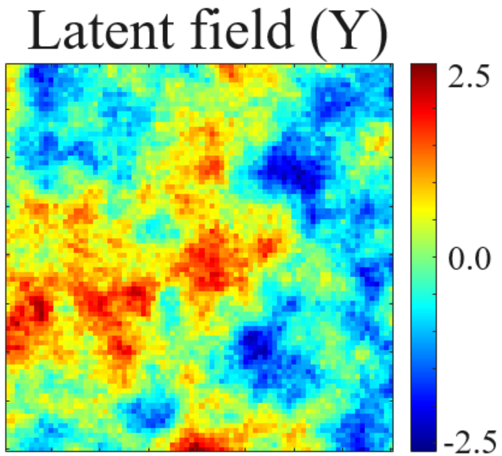
Rainfall model (2/5): trans-Gaussian Random Field

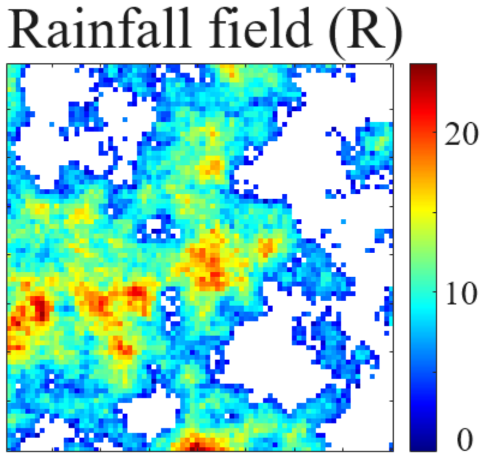
Trans-Gaussian geostatistics split the rain signal (R) in two components:

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

Latent field:
Models spatial dependencies through the covariance function C_Y

Transform function:
Models rainfall occurrence and intensity





Rainfall model (2/5): trans-Gaussian Random Field

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 (3/5): making the model non-stationary

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

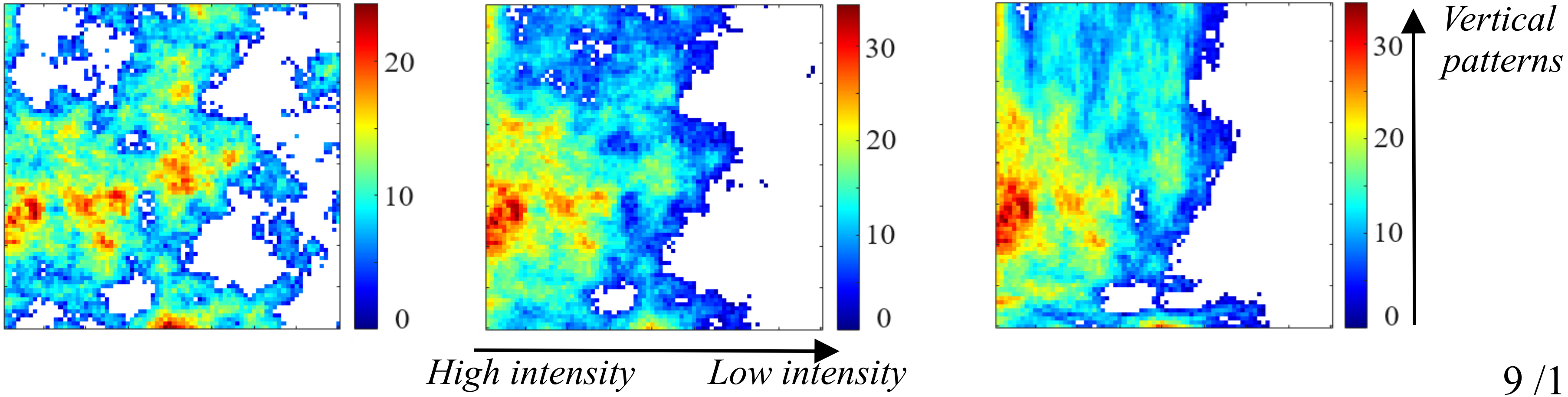
=> Model parameters are made location-dependent:

→ $\psi \Rightarrow \psi_{\mathbf{s}}$ *and* $C_Y \Rightarrow C_{Ys}$ *(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*]

Stationary model — Non-stationary ψ — Non-stationary ψ and C_Y



Rainfall model (3/5): making the model non-stationary

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

=> Model parameters are made location-dependent:

→ $\psi \Rightarrow \psi_{\mathbf{s}}$ *and* $C_Y \Rightarrow C_{Y\mathrm{s}}$ *(with* **s** *the location of interest)*

→ $C_{Y\mathrm{s}}$ *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*]

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

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

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

Rainfall model (4/5): parameter estimation

Model parameters are estimated by likelihood maximization

→ *Overall 9 parameters*

Estimation of model parameters from sparse observations

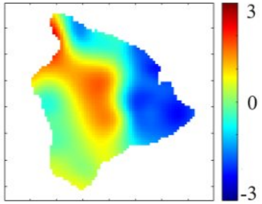
→ *Marginal distribution (ψ_s): estimation at gauge locations + Ordinary Kriging*

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

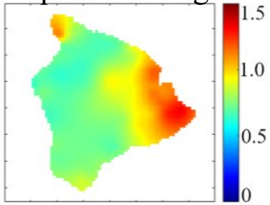
Example for RainType 5

Parameters of the marginal distr. ψ

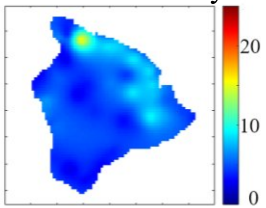
$\mathbf{a_0} \approx$ proba. dry day



$\mathbf{k} \approx$ proba. strong rain



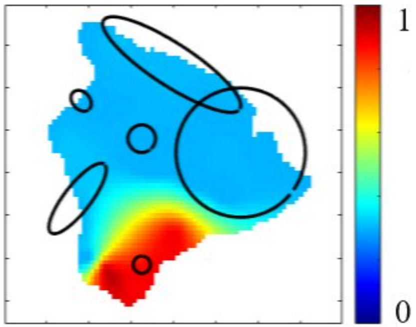
$\boldsymbol{\theta} \approx$ rain intensity



Parameters of the covariance C_Y

Background :

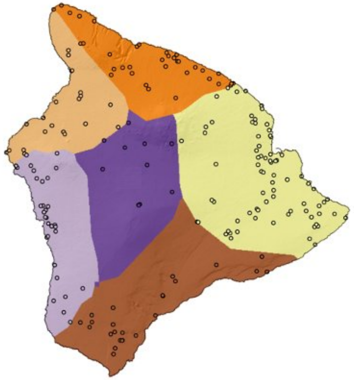
$\boldsymbol{\nu} \approx$ short lag variability



On top:

Ellipses of anisotropy ≈ correlation distance

Hawai‘i climate divisions



Rainfall model (5/5): simulation (unconditional)

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

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

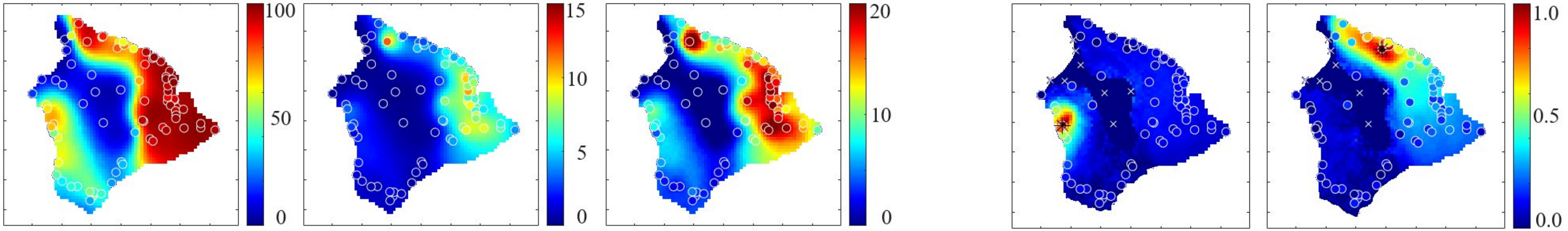
= *Stochastic rainfall generation*

Example for RainType 5

Statistics of the marginal distribution

Statistics of the spatial structure

Rain occurrence (%) | Mean rainfall (mm/day) | $q_{90\%}$ rainfall (mm/day) | Correlation with point *



dots = observations; background maps = simulations

Rainfall model (5/5): simulation (conditional)

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

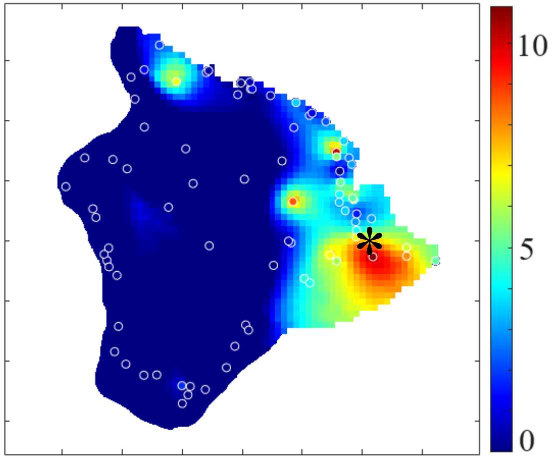
(2) Simulation of censored latent values (dry obs.) by Gibbs sampler

(3) Conditioning by conditional Kriging

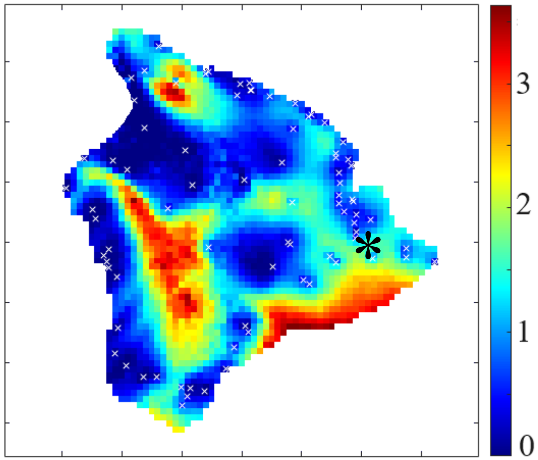
(4) Transformation of the latent field into rain intensities: $R(\mathbf{s}) = \Psi(Y(\mathbf{s}))$

Example for RainType 5

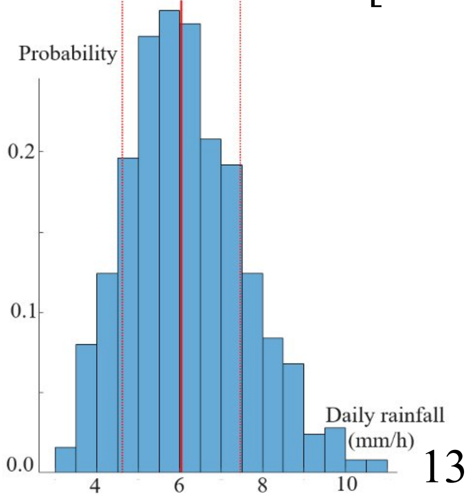
Estimated rainfall [$q_{50\%}$ sim.] (mm/day)



Uncertainty [1σ] (mm/day)



Probabilistic forecast [loc. *]



Conditioning to monthly totals (1/4): motivation

Uncertainties increase with the distance to rain gauges

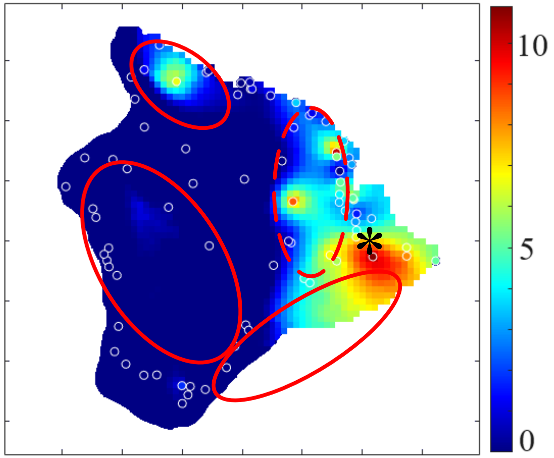
→ *Large uncertainties in poorly gauged areas and at the edges of the domain*

→ *Spatial gradients may be over-smoothed*

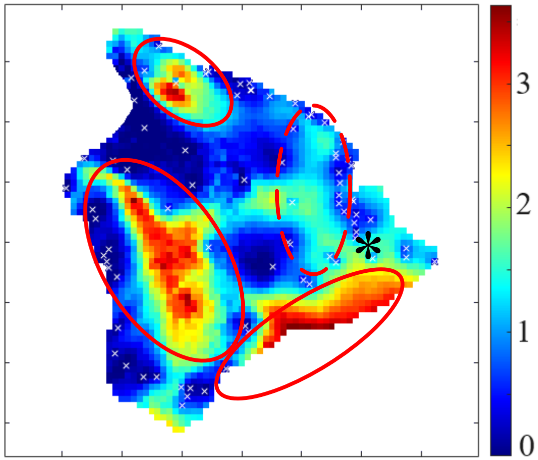
Use the spatial information from monthly maps to constrain the daily maps

Example for RainType 5

Estimated rainfall [$q_{50\%}$ sim.] (mm/day)



Uncertainty [1σ] (mm/day)



Conditioning to monthly totals (2/4): method

Challenge: different spatial models (i.e., rain types) during the same month

=> **Metropolis within Gibbs**

(1) Select target locations and initialize their latent values

(2) For each location and each day:

→ Gibbs sampling conditional to (i) latent obs. and (ii) simu. at other target locations
→ Transformation to get daily rainfall simulation at target location
→ Acceptance following a Metropolis rule applied to monthly sum

(3) Iterate (2) for warm-up period + sampling daily rainfall at target locations

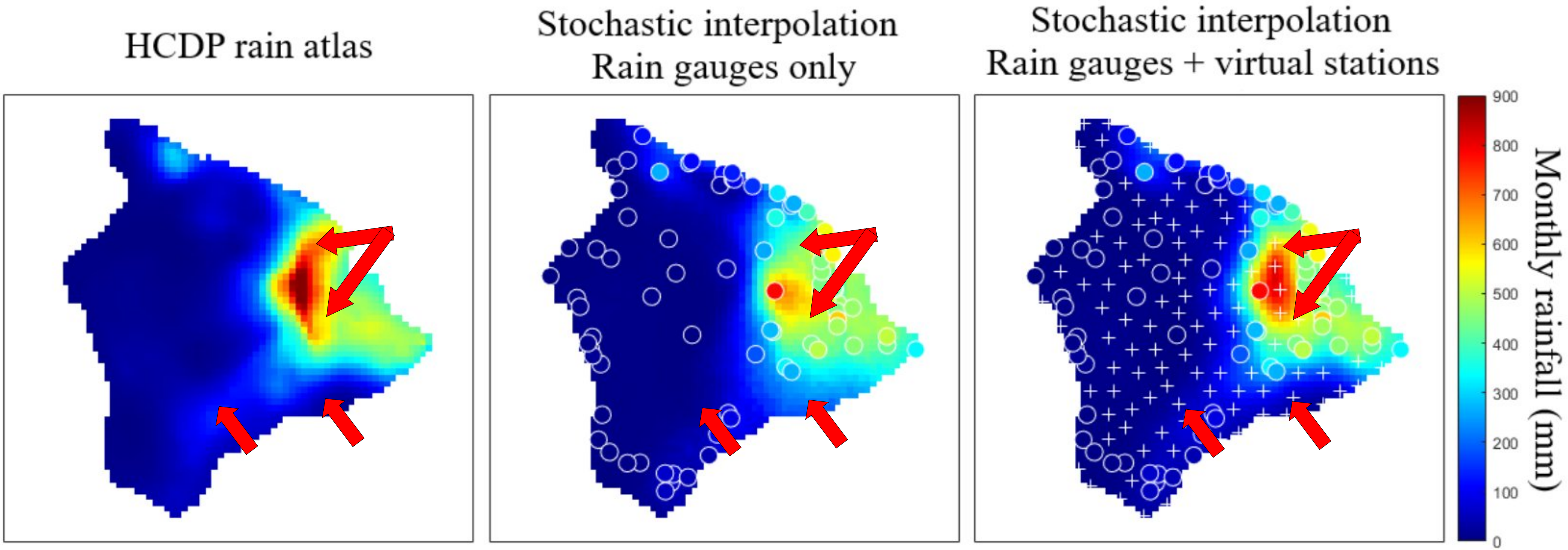
Other problem: low computational efficiency

=> **Metropolis within Gibbs to simulate a small set of 'virtual stations'** (≈100)

\+ *Conditional simulation using (i) daily rain gauge observations*
(ii) virtual stations pseudo-observations

Conditioning to monthly totals (3/4): results - impact on monthly totals

January 2018



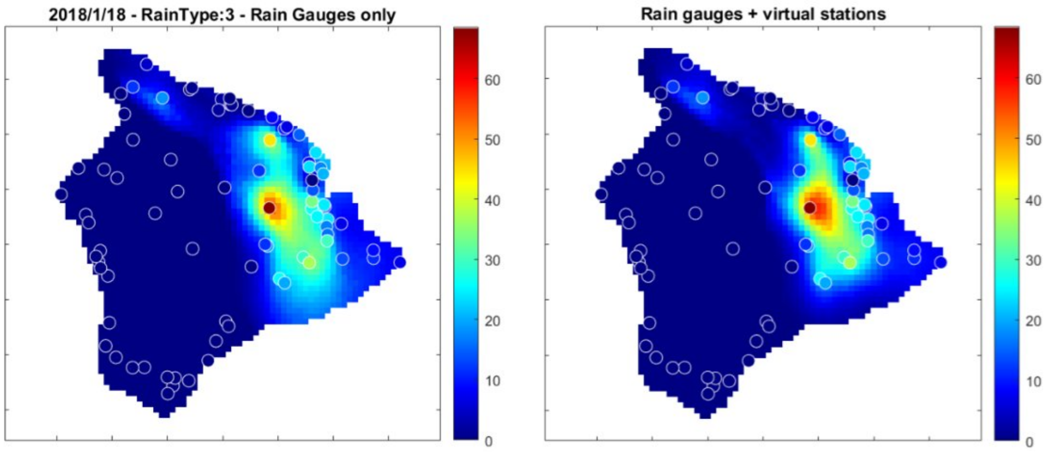
o Rain gauge

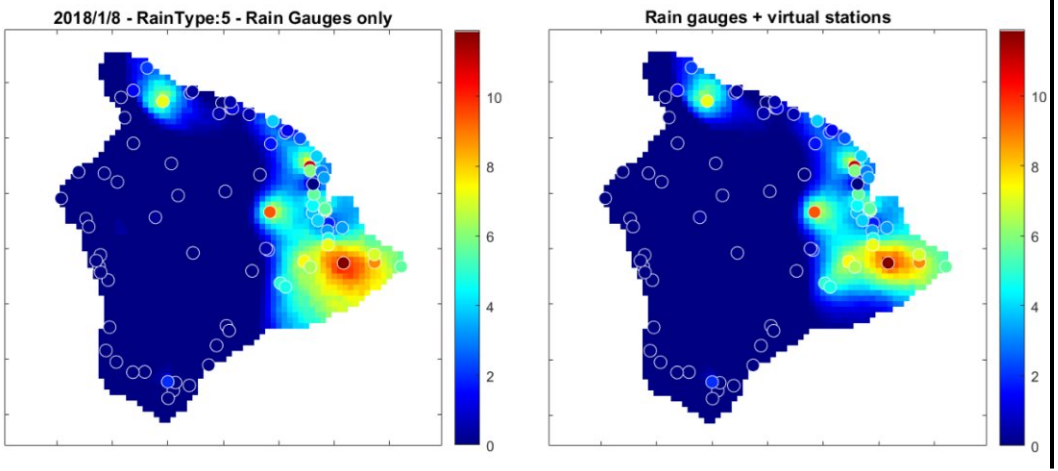
+ Virtual station

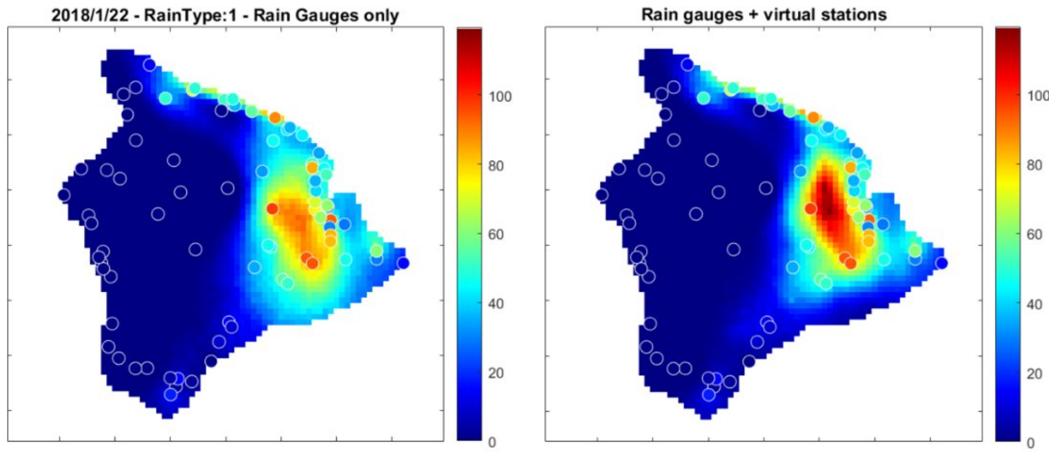
Conditioning to monthly totals (4/4): results - impact on daily maps

Improve gradients also in daily rainfall maps

Can simulate daily maximum between rain gauges







Conclusion & perspectives:

Non-stationary trans-Gaussian geostatistics:

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

→ *Applications: (i) stochastic rainfall generator, (ii) rainfall mapping*

Conditioning to monthly totals:

→ *Ensures consistency between rainfall maps at different time scales*

→ *Improves the mapping of spatial gradients at the daily scale*

Next step: non-stationary space-time model for sub-daily orographic rainfall

→ *Non-stationary space-time covariance functions (cf. presentation Denis)*

→ *Diurnal cycle*

Thank you for your attention :-)