EGU26 – General Assembly,
3 - 8 May 2026, Vienna, Austria

How Reliable Are Rainfall Observations? Assessing Credible Intervals with Bilinear Surface Smoothing

Nikolaos Malamos[1], *Theano Iliopoulou*[2,3], *Panagiotis D. Oikonomou*[1] *and Demetris Koutsoyiannis*[2]

[1]*Department of Natural Resources Development & Agricultural Engineering, Agricultural University of Athens, Greece*

[2]*Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Greece*

[3]*Department of Geography, Harokopio University, Athens, Greece*

Introduction

- Rainfall regionalization refers to a broader spatial modeling process that transforms point measurements into reliable continuous fields, incorporating additional information.
- Yet the fidelity of the resulting continuous surface is strongly influenced by the quality of the underlying data, as well as by the density and spatial configuration of the observational network
- This contribution addresses the question of how reliable rainfall data are when evaluated against a regionalized rainfall surface, by extending the Bilinear Surface Smoothing with Explanatory variable (BSSE) framework to explicitly incorporate Bayesian credible intervals

Study area

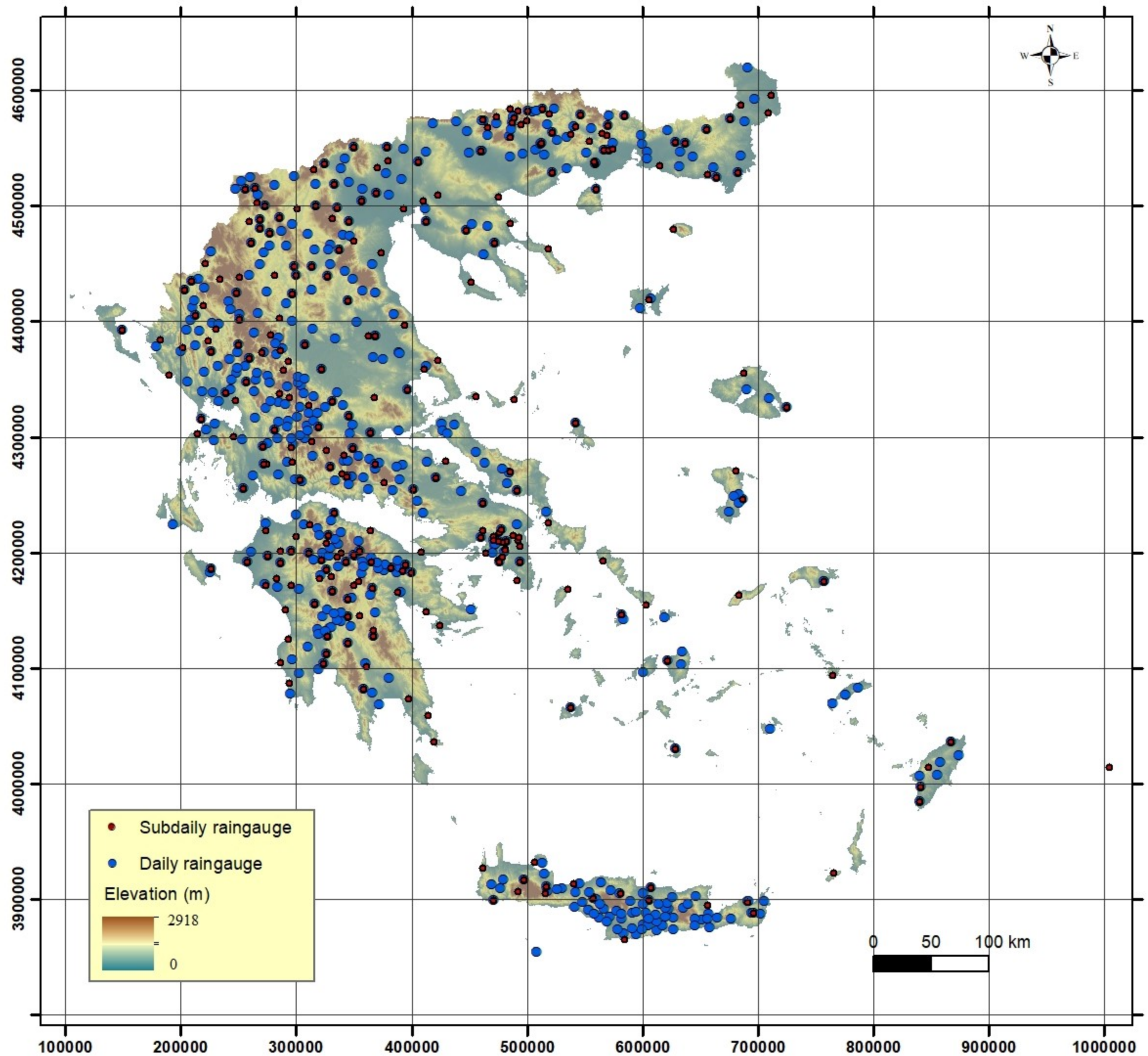
- The study area is the entire Greek territory, spanning 131957 km^2 and including a multitude of islands, and a diverse landscape with considerable differences in rainfall distribution

- Greece includes 4 climatic types: Arid, Temperate, Cold and Polar and 11 sub-types: BSh, BSk, Csa, Csb, Cfa, Cfb, Dsb, Dsc, Dfb, Dfc and ET, based on the Köppen-Geiger climate classification (Kottek et al., 2006)

Data

- Daily records: 61 rainfall series with over 60 years of data, 56 of which are sourced from the Hydroscope database (http://www.hydroscope.gr/) and 5 from the Greek Meteorological Service
- Monthly records within Greece: 31 rainfall series from the Global Historical Climatology Network (GHCN), each spanning more than 30 years
- Monthly records from neighboring countries: 36 stations from the GHCN, located in Turkey, North Macedonia, Bulgaria, and Albania, also with data spanning over 30 years

The elevation of each station to be incorporated as an additional explanatory variable, is derived from the Shuttle Radar Topography Mission (SRTM) dataset



Elevation map of Greece along with the locations of the daily and sub-daily resolution rainfall stations used in the analysis (Iliopoulou et al., 2024)

Bilinear Surface Smoothing - BSS

- The general idea behind both methods is to compromise the trade-off between the objectives of minimizing the fitting error and the roughness of the fitted bilinear surface, therefore termed Bilinear Surface Smoothing (BSS) (Malamos & Koutsoyiannis, 2016a)
- The larger the weight of the first objective, the rougher the surface will appear, while the opposite is true for a larger weight of the second objective
- BSS suggests that fitting is based on minimizing the Generalized Cross Validation Error (GCV) between the set of the given data points and the corresponding estimates. (Craven & Wahba, 1978)
- The method was extended by the introduction of an additional explanatory variable available at a denser dataset compared to that of the main variable (BSSE)

Estimation functions

- For a point *z* with spatial coordinates (*x*, *y*) on a plane, we have:

Method	Estimation function
BSS	$\hat{z}(x,y) = d$
BSSE*	$\hat{z}(x,y) = d + t \cdot e$

- *d* and *e* are the values of two fitted bilinear surfaces at that point.
- These are not a global linear relationships but local as the quantities *d* and *e* change in space
- For each point *z* there corresponds a value *t*

Smoothing parameters

- In the case of the BSSE, there are six adjustable parameters
  - The numbers of intervals along the horizontal and vertical direction respectively i.e., mx, my
  - For surface *d*: the smoothing parameters $\tau_{\lambda x}$ and $\tau_{\lambda y}$
  - For surface *e*: The incorporation of the explanatory variable, requires two more adjustable parameters the smoothing parameters $\tau_{\mu x}$ and $\tau_{\mu y}$

- The values of all the smoothing parameters are restricted in the interval [0, 1) for both directions
- When the smoothing parameters are close to 1, the resulting bilinear surfaces exhibit greater smoothness, whereas for small values of these parameters interpolation among the known points is obtained

BSS and Credible intervals formulation

- The estimation function for BSSE can be written:

$$\hat{\boldsymbol{z}} = \boldsymbol{A}\,\boldsymbol{z}$$

- with: $\boldsymbol{A} = \boldsymbol{\Pi}\,\boldsymbol{T}\boldsymbol{\Pi}\,\boldsymbol{\Omega}^{-1}(\boldsymbol{\Pi}\boldsymbol{T}\boldsymbol{\Pi})^{\mathrm{T}}$,
- $\boldsymbol{\Omega} :=$

$$\begin{bmatrix} \boldsymbol{\Pi}^{\mathrm{T}}\boldsymbol{\Pi} + \lambda_x \boldsymbol{\Psi}_x^{\mathrm{T}}\boldsymbol{\Psi}_x + \lambda_y \boldsymbol{\Psi}_y^{\mathrm{T}}\boldsymbol{\Psi}_y & \boldsymbol{\Pi}^{\mathrm{T}}\boldsymbol{T}\boldsymbol{\Pi} \\ \boldsymbol{\Pi}^{\mathrm{T}}\boldsymbol{T}\boldsymbol{\Pi} & \boldsymbol{\Pi}^{\mathrm{T}}\boldsymbol{T}^{\mathrm{T}}\boldsymbol{T}\boldsymbol{\Pi} + \mu_x \boldsymbol{\Psi}_x^{\mathrm{T}}\boldsymbol{\Psi}_x + \mu_y \boldsymbol{\Psi}_y^{\mathrm{T}}\boldsymbol{\Psi}_y \end{bmatrix}$$

- Matrix $\boldsymbol{\Pi}$ contains the weights of each element of the bilinear surfaces, while $\boldsymbol{T}$ is a diagonal matrix with its elements being the values of the explanatory variable at the given data points.
- Matrices $\boldsymbol{\Psi}_x$ and $\boldsymbol{\Psi}_y$ represent the smoothness of the bilinear surfaces and assure a unique solution of the fitting problem.
- The elements of these two matrices are the differences of slopes between consecutive segments of the bilinear surface according to the x and y directions

Residuals and Residual Sum of Squares

- The vector of residuals ***r***, i.e., the difference between the observed values and the fitted values and in the BSS framework is defined as:

$$\boldsymbol{r} = \boldsymbol{z} - \hat{\boldsymbol{z}} = \boldsymbol{z} - \boldsymbol{A}\boldsymbol{z} = (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{z}$$

- The Residual Sum of Squares (RSS) measures the total squared difference between the observed values and the fitted surface thus providing the total discrepancy between them:

$$\mathrm{RSS} = \sum_{i=1}^{n} (z_i - \hat{z}_i) = \|\boldsymbol{z} - \boldsymbol{A}\boldsymbol{z}\|^2 = \|(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{z}\|^2$$

(Malamos et al., 2025)

Estimation of the Error Variance

- The error variance σ^2 represents the expected squared magnitude of the random errors r_i, i.e., how much the observed values deviate from the true underlying surface due to randomness or uncertainty. It quantifies the uncertainty or variability in the data that is not explained by the model
- The error variance σ^2 is estimated by scaling the RSS by the effective residual degrees of freedom
- In the BSS framework, this variance is estimated as:

$$\sigma^2 = \frac{\mathrm{RSS}}{v}, v = trace(\boldsymbol{I} - \boldsymbol{A})$$

- where v represents the equivalent degrees of freedom for error, which indicates the degrees of freedom that remain after fitting the bilinear surface to the data

Posterior Covariance Matrix of the Fitted Surface and Credible Interval Assessment

- The posterior covariance matrix $\boldsymbol{\Sigma}$ of the fitted surface values is given by:

$$\boldsymbol{\Sigma} = \sigma^2 \left(\boldsymbol{I} - \boldsymbol{A}\right)$$

- The diagonal elements of matrix $\boldsymbol{\Sigma}$ i.e. Σ_{ii}, represent the posterior variances of the fitted values at each observation point
- The credible intervals for the fitted bilinear surface $\hat{z}(x_i, y_i)$ at each observation point (x_i, y_i) are given by:

$$\hat{z}(x_i, y_i) \pm t_{\alpha/2,\nu} \cdot \sqrt{\Sigma_{ii}}$$

- where $\hat{z}(x_i, y_i)$ is the fitted value at point (x_i, y_i), $t_{\alpha/2,\nu}$ is the critical value from the Student's t-distribution with $\nu = \text{trace}(\boldsymbol{I} - \boldsymbol{A})$ degrees of freedom and confidence level $1 - \alpha$ and Σ_{ii} is the i_{th} diagonal element of the posterior covariance matrix $\boldsymbol{\Sigma}$, representing the posterior variance of the fitted value at point (x_i, y_i)
- This credible interval reflects uncertainty around each estimated point by means of the fitted surface, based on the observed data and prior information about the smoothness, as controlled by the smoothing parameters

Results

- The BSSE framework was applied to the regionalization of for every temporal scale of rainfall as estimated from the ground stations by employing the elevation as covariate

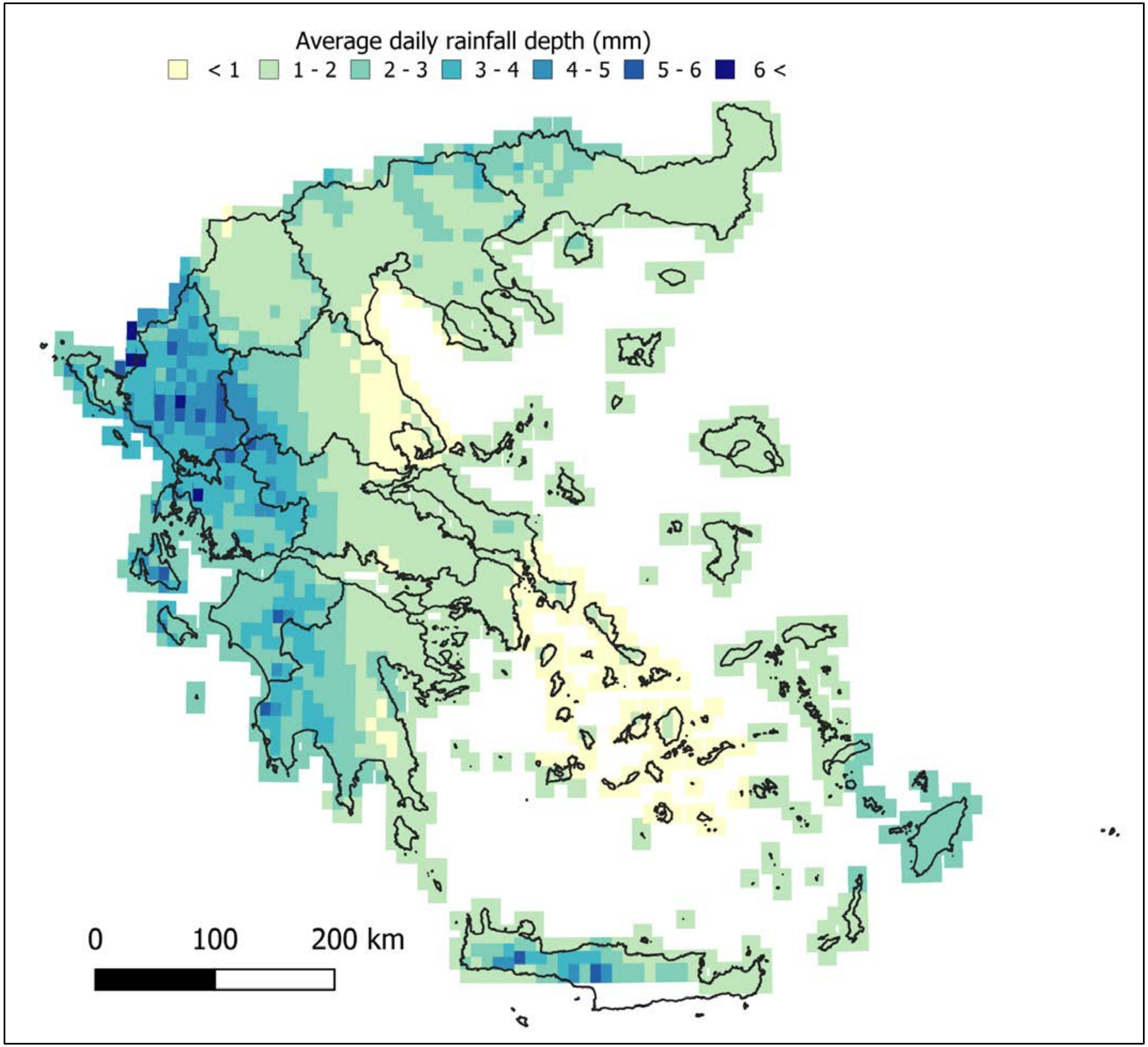
BSSE optimal parameter values

Timescale	number of segments, mx	number of segments, my	$\tau_{\lambda x}$	$\tau_{\lambda y}$	$\tau_{\mu x}$	$\tau_{\mu y}$
0.5 h	9	140	0.884742	0.024031	0.001	0.001
1 h	9	198	0.016242	0.009531	0.001	0.001
3 h	9	254	0.001	0.001	0.001	0.001
6 h	9	348	0.001	0.001	0.001	0.001
12 h	9	448	0.001	0.001	0.001	0.001
24 h	14	981	0.001	0.001	0.001	0.001
48 h	14	387	0.001	0.001	0.001	0.001
Daily Average	6	147	0.99	0.001	0.652857	0.018715

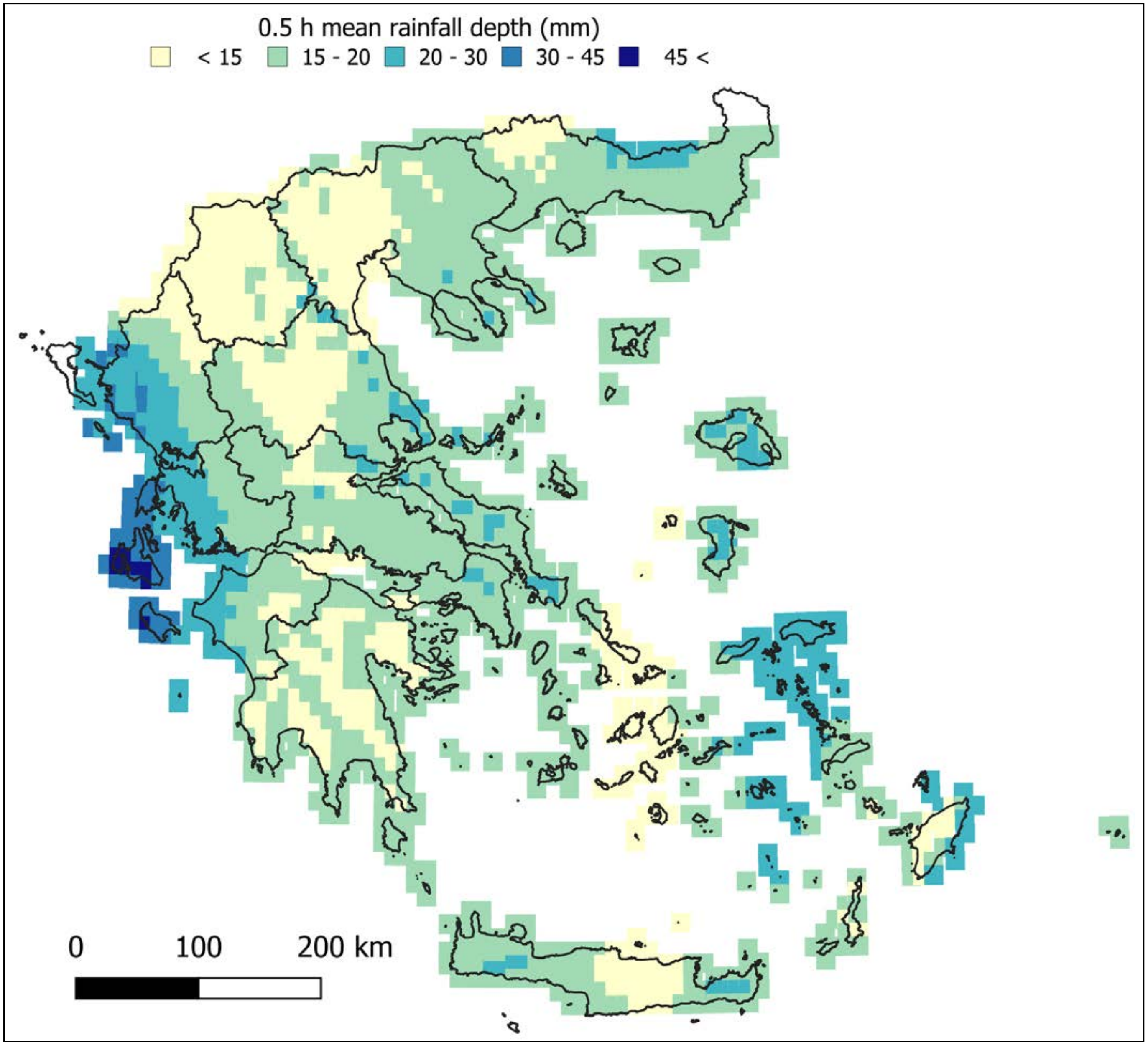
Performance assessment statistics for the regionalization for average daily rainfall

Surface statistics	All data	LOOCV
MBE (mm)	0.00	-0.01
MAE (mm)	0.22	0.36
RMSE (mm)	0.31	0.49
NSE	0.90	0.74
r^2	0.90	0.75
NRMSE (%)	6.95	10.99
σ (mm)	0.89	-0.01

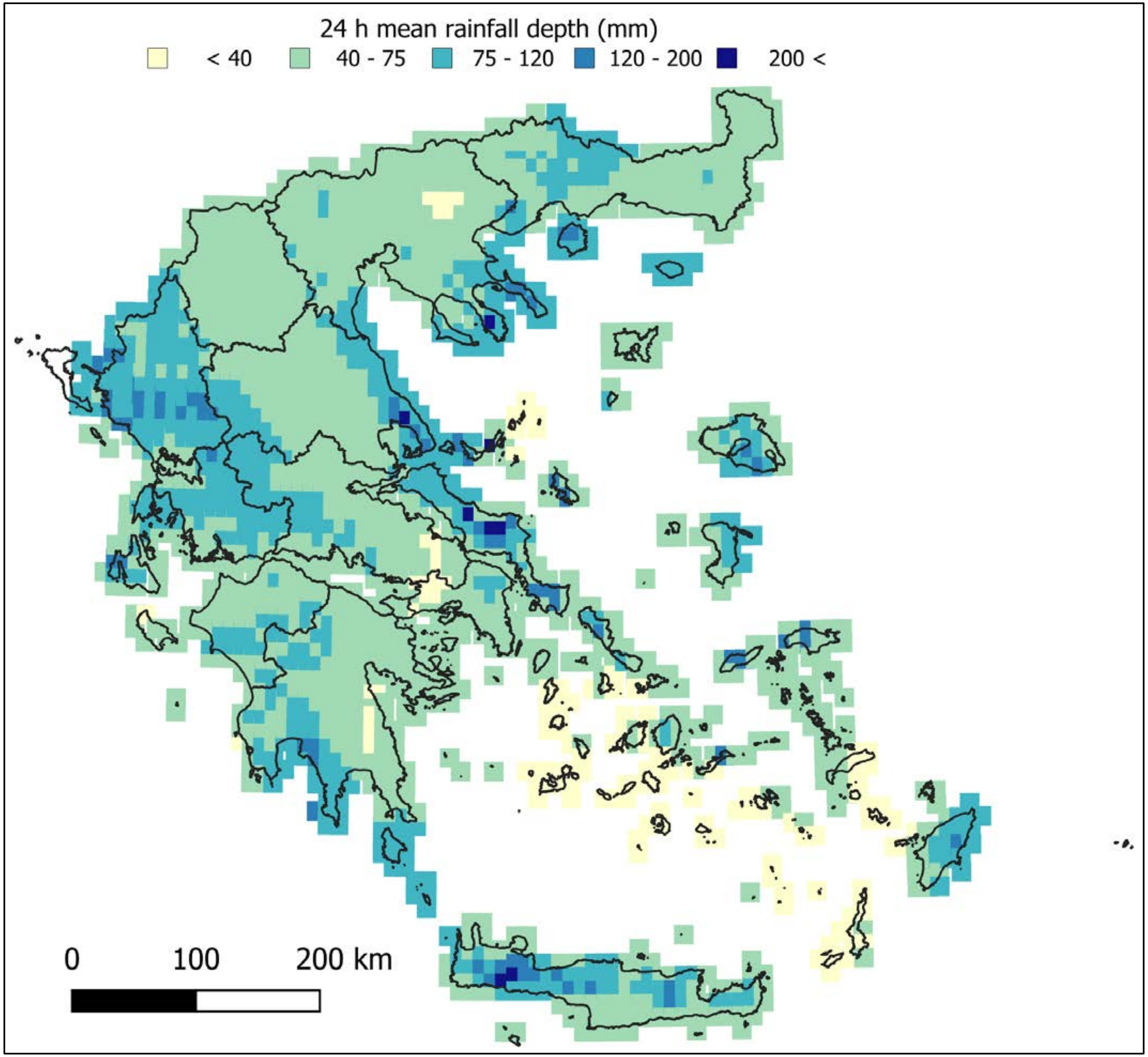
Daily average rainfall



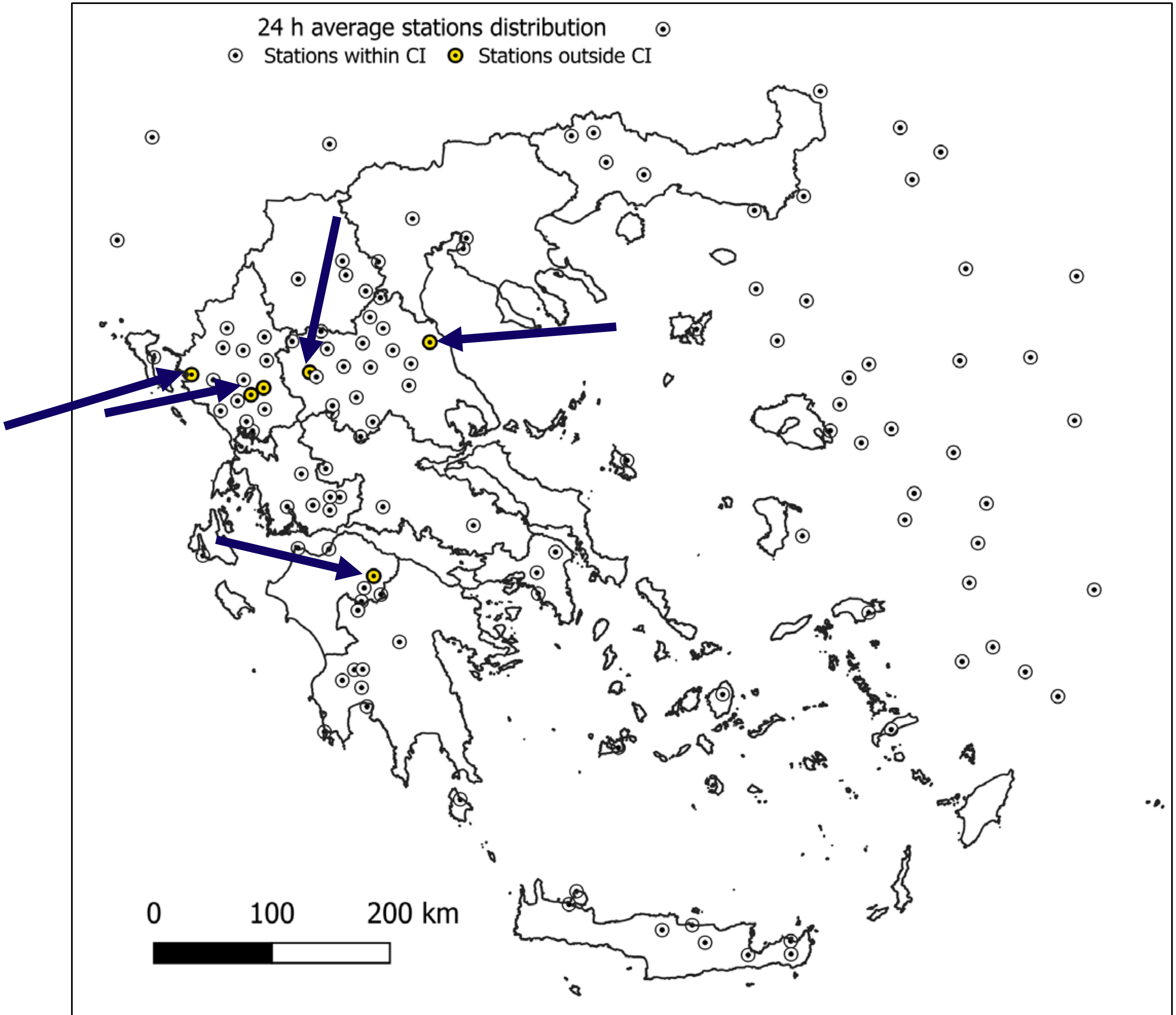
0.5 h maxima



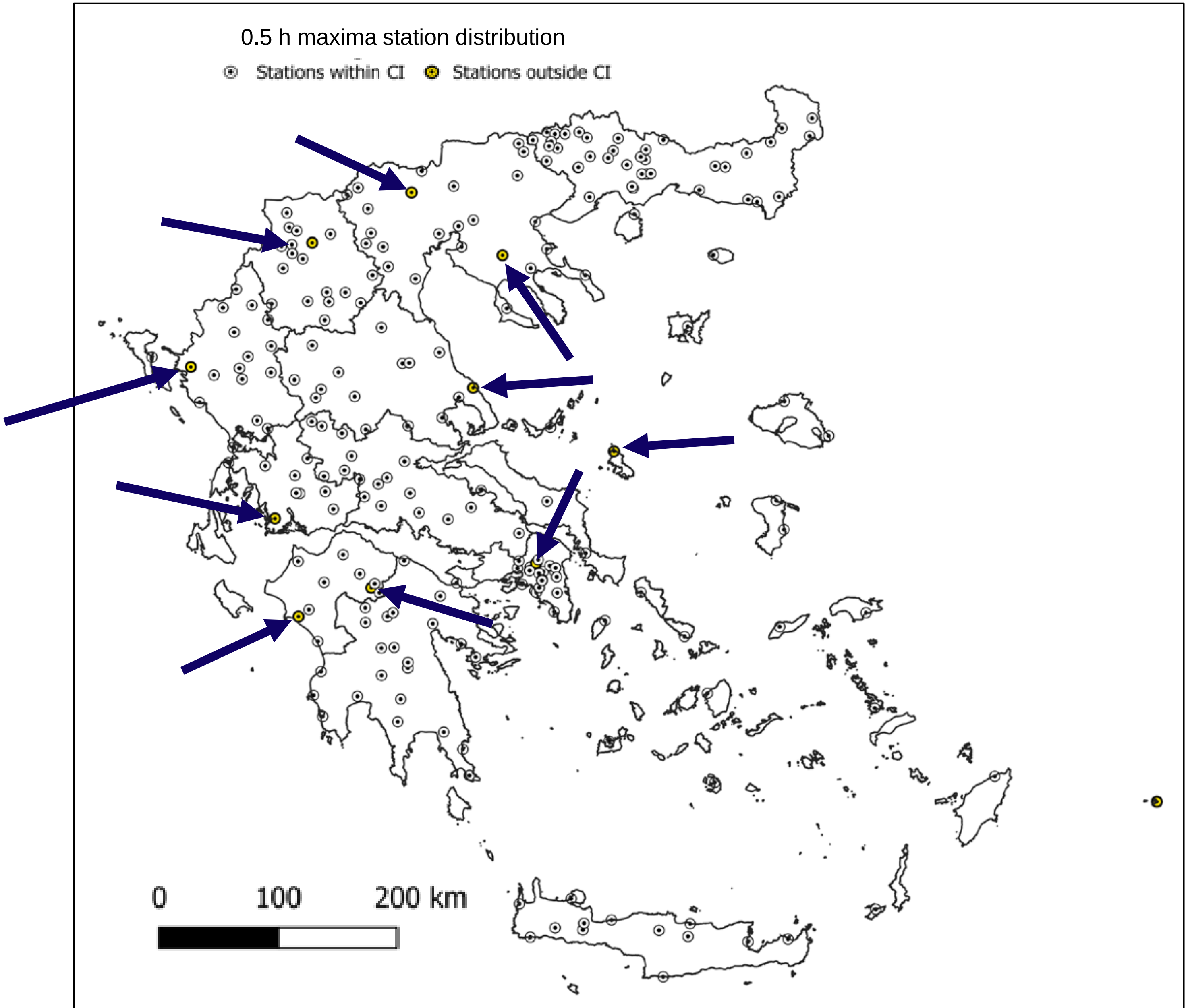
24 h maxima


24 h mean rainfall depth (mm)
< 40
40 - 75
75 - 120
120 - 200
200 <
0
100
200 km

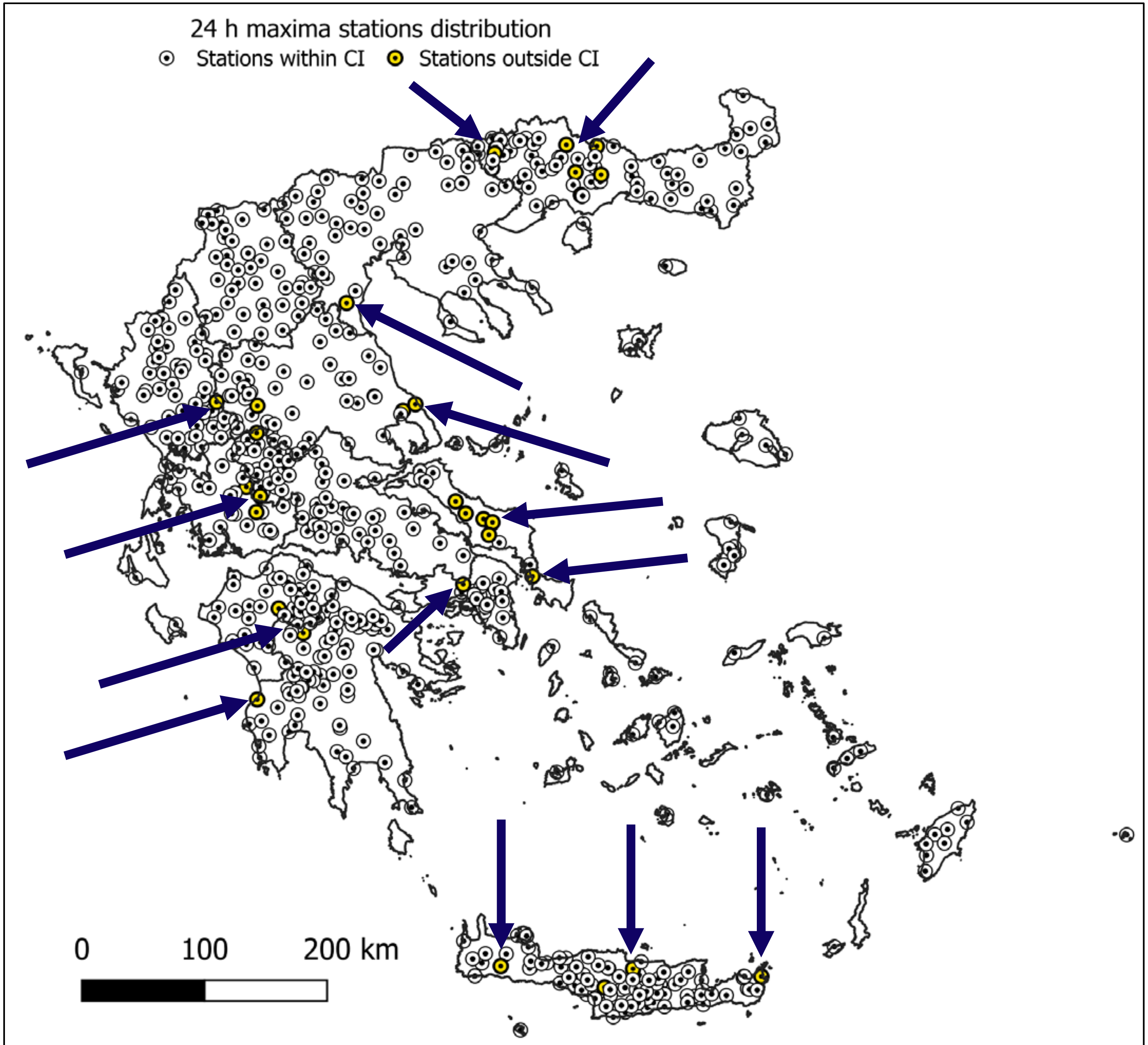
Outliers detection – Daily average rainfall



Outliers detection – 0.5 h maxima



Outliers detection – 24 h maxima



Observations outside CI	0.5 h	1 h	3 h	6 h	12 h	24 h	48h	Daily
	5.2%	4.8%	4.5%	4.6%	5.7%	4.9%	6.1%	4.7%

- The identified stations were examined, revealing no unrealistic behavior, and their occurrence remained within the expected 5% threshold, specifically at 4.9% for the 24 h maxima and 4.7% for the average daily rainfall
- Most of these ‘outliers’ align with established patterns of either very high or low average maximum rainfall values
- A few stations were flagged due to limited data length and higher associated uncertainty
- The identified stations highlight areas where the gauge network is sparse, emphasizing the need for prioritized station installation
- Such efforts should focus on regions with complex rainfall patterns, where additional data is crucial for robust characterization

Conclusions (1)

- The use of credible intervals proved instrumental in assessing the uncertainty in the regionalization process

Conclusions (2)

- The credible intervals highlighted areas where the station network may be insufficient or where unusual rainfall patterns might exist

Conclusions (3)

- This information is invaluable for improving network design, as it can guide the placement of new stations or the adjustment of existing ones to better capture spatial variability, especially in regions prone to extreme rainfall events

Conclusions (4)

- By focusing on areas outside the 95% credible interval, the network can be optimized to provide more accurate rainfall estimates both for flood risk management and climatic studies

Conclusions (5)

- Overall, this work underscores the importance of selecting appropriate covariates and temporal scales when regionalizing rainfall patterns in complex terrains like Greece

References (1)

- Craven, P., & Wahba, G. (1978). Smoothing noisy data with spline functions. *Numerische Mathematik, 31*(4), 377–403. https://doi.org/10.1007/BF01404567
- Iliopoulou, T., Koutsoyiannis, D., Malamos, N., Koukouvinos, A., Dimitriadis, P., Mamassis, N., Tepetidis, N., & Markantonis, D. (2024). A stochastic framework for rainfall intensity–time scale–return period relationships. Part II: Point modelling and regionalization over Greece. Hydrological Sciences Journal, 1–21. https://doi.org/10.1080/02626667.2024.2345814
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift, 15*(3), 259–263. https://doi.org/10.1127/0941-2948/2006/0130

References (2)

- Malamos, N., & Koutsoyiannis, D. (2016a). Bilinear surface smoothing for spatial interpolation with optional incorporation of an explanatory variable. Part 1: Theory. *Hydrological Sciences Journal*,61(3),519–526. https://doi.org/10.1080/02626667.2015.1051980
- Malamos, N. & Koutsoyiannis, D. (2016b) Bilinear surface smoothing for spatial interpolation with optional incorporation of an explanatory variable. Part 2: Application to synthesized and rainfall data. *Hydrological Sciences Journal* 61(3), 527–540. doi:10.1080/02626667.2015.1080826
- Malamos, N., Iliopoulou, T., Dimitriadis, P., & Koutsoyiannis, D. (2025). Comparative Analysis of IMERG Satellite Rainfall and Elevation as Covariates for Regionalizing Average and Extreme Rainfall Patterns in Greece by Means of Bilinear Surface Smoothing.Geosciences,15(6),212. https://doi.org/10.3390/geosciences15060212