The SAFRAN daily gridded precipitation product in Tunisia (1979-2015)

Haifa Feki[1,2]
Yves Tramblay[3]
Pere Quintana-Seguí[4]
José A. Guijarro[5]
Julie Carreau[3]

[1] Ecole supérieure des ingénieurs de Medjez el beb, Tunisia
[2] Laboratoire GREEN TEAM, Université de Carthage, Tunisia
[3] HydroSciences Montpellier (Univ. Montpellier, CNRS, IRD), Montpellier France
[4] Ebro Observatory, Ramon Llull University – CSIC, Roquetes (Tarragona), Spain
[5]Agencia Estatal de Meteorología (AEMET), Balearic Islands Office, Spain

1. INTRODUCTION

Middle East and North Africa (MENA) region is subject to water scarcity (Ragab and Prudhomme, 2002) and more than 60% of the population live in areas of high water stress compared to a global average of about 35%. Droogers et al. (2012) mentioned that, in present time, the average per capita water availability in MENA region is slightly above the physical water scarcity limit at about 1076 m^3/yr compared to the world average of about 8500 m^3/yr. In particular, water resources in Tunisia are identified by their scarcity, low quality, poor distribution and seasonal distribution (Ben Zaied and Bient, 2015).

The 4th assessment report of the IPCC (IPCC, 2012) projects strong changes in climate across the MENA region. Climate change is expected to increase water stress through various mechanisms including reduced precipitation, intensifying rainfall variability and rising temperature. However, the problem of water scarcity in Tunisia, according to Haddadin (2009), is not only solely based on the availability of the resource but also a man-made problem. Several studies (Ragab and Prudhomme, 2002, Schilling et al., 2012, Tramblay et al., 2018), mentioned that by 2050, North Africa is expected to have reduced rainfall amounts of 20 to 25% less the present mean value. A recent study et Zittis (2017) using various existing gridded datasets, showed that the long term trends in the Middle-East and North Africa (MENA) region is indicating an overall drying since the beginning of the twentieth century mainly, over the Maghreb region. They also noted that the different data sources have statistically significant differences in the distribution of monthly precipitation for about 50% of the domain.

Precipitation studies are mostly carried out based on gridded precipitation data such as the EOBS (Haylock, 2008), CRU (Harris et al., 2014) or GPCC (Schamm, 2014) datasets. This type of data is necessary for local climate studies, climate change monitoring at regional scale, validation of regional climate models (RCM) and impact models (Haylock, 2008). Several gridded precipitation products have been developed for countries such as SPAIN2 in Spain (Herrera et al., 2012), SAFRAN in France (Quintana-Seguí et al., 2008) and Spain (Quintana-Seguí et al., 2016). For the Euro-Mediterranean region, the EOBS dataset is probably the most employed and provides daily high-resolution (25 km to 10 km) gridded precipitation. Yet, these gridded datasets are widely used for climate studies but in regions with data scarcity they can introduce a significant uncertainty (Romera et al., 2015, Prein and Gobiet 2017, Zittis, 2017). As noted by several authors, the Euro-Mediterranean domain is covered by an uneven station density, and the use of this dataset could be problematic in particular when looking at extreme precipitation (Flaounas et al, 2012, Turco et al., 2013, Fantini et al., 2016). As for Tunisia, EOBS contains a small number of stations (Haylock et al., 2008). Beside gridded datasets based on interpolated precipitation, a growing number of satellite-based precipitation products are becoming available with almost a global coverage and relying on different sensors (Kummerow et al., 1998; Mehta and Yang, 2008, Dhiba et al., 2017). Several of these products have been successfully validated in the Mediterranean context (Nastos et al., 2013, Tramblay et al., 2016) and also merged with gauge and reanalysis datasets such as in the MSWEP product (Beck et al., 2017).

In Tunisia, rainfall is highly variable both temporally and geographically, while surface water is a very important resource for agricultural activities and consumption. In this context, there is a need for country-scale information systems to analyze and mitigate climate change impacts but also to develop regional or basin-scale surface modeling to improve resources management. Country-scale reliable precipitation data is currently lacking to better estimate the spatial variability of precipitation extremes (Dhib et al., 2017) or to validate the most recent climate models (Bargaoui et al., 2014, Fathalli et al. 2018). The goal of the present study is to develop a gridded data set of precipitation in Tunisia, making use of the whole rain gauge monitoring network of the country. Different interpolation methods are first compared and the SAFRAN reanalysis is implemented over the whole Tunisian territory. Then, the high-resolution gridded dataset produced is compared with a state-of-art daily precipitation product; the EOBS database (Haylock et al., 2008).

2. DATASETS

The full rain gauges database of the *Direction Générale des Resources en Eau* (DGRE) of Tunisia containing over 2000 stations has been processed. The cover has a higher density the North of Tunisia, where are located most the dams and reservoirs of the country. A global quality check has been performed; since several stations were lacking information's about their locations or had long periods of missing data. Only the 960 stations with at least 5 complete years between the years 1979 to 2015 have been considered for subsequent analysis. For the stations with no metadata, we used the historical publications of the DGRE, the *Annuaire Hydrologiques de Tunisie*, (see one example here for the year 2007/2008: *http://www.hydrosciences.fr/sierem/produits/biblio/annales/TN/2007-2008.pdf*) available for several years and containing the data, maps and station information.

In addition to station data, we used the EOBS database (Haylock et al., 2008) to obtain different precipitation indices to be compared with the SAFRAN product. These indices are computed on an annual basis and include the highest 1-day precipitation amount (RX1day), the number of wet days (R1mm) and the total precipitation due to wet days (PRCPTOT).

3. METHODS

3.1 Quality check and homogeneity tests

The precipitation dataset was checked for quality and homogeneity by means of the R package Climatol v.3.1 (Guijarro, 2017 and 2018). As this software works better with whole years and the dataset comprised data until August 2015, quality controls were applied to the period 1979-2014. Due to the high variability of daily rainfall, especially in arid climates, it is not possible to check the homogeneity of the series at the daily time step. Therefore, the homogenization was performed on monthly totals calculated from the daily data. The procedure implemented in Climatol consists in applying the Standard Normal Homogeneity Test (Alexandersson, 1986) to the series of anomalies, i.e., differences between observed data and their estimates from nearby stations, both in normalized form by dividing each data by their corresponding average. This procedure led to the detection of 35 break-points (changes in the mean of the anomaly series) with the SNHT test in 30 series. For these stations, the dates of the break-points were then used to split the daily series into separate homogeneous sub-periods and adjusting them to the longest ones.

Outliers were also checked in the anomaly series, but beside a few obvious errors (i.e. negative precipitation or daily precipitation above 1000 mm/day, for example) no daily data were rejected because the few outliers found were considered feasible in the frame of an arid climate with frequent isolated precipitations of convective origin.

3.2 Interpolation of rain gauge precipitation

The interpolation of the rain gauge data is performed in the present study with different methods. Deterministic methods such as Inverse distance weighting (ID) and the nearest neighbor (NN) are considered. In addition, the geostatistical method of Ordinary Kriging (OK) is also considered (Goovaerts 2000). The variograms required for the OK method are fitted automatically with a spherical variogram model for each time step when rain is measured at least in 3 stations, otherwise ID interpolation is performed (Tramblay et al., 2016). The spherical variogram model is convenient for precipitation, which is not a spatially continuous field like temperature, since it provides a value of the de-correlation distance (Lebel et al., 1987).

In order to take the influence of elevation into account in the interpolation (Feki et al., 2012), the residual kriging (RK) method (Hengl et al., 2007) is implemented in addition to OK. It is mathematically equivalent to universal kriging or kriging with external drift methods, but RK allows the separate interpretation of the two interpolated components and use of a broader range of regression techniques. The RK approach combines a regression model between precipitation and altitude with the spatial interpolation of the residuals of the regression. At each time step, for each location *i* the estimate of precipitation *z* can be computed as:

$$z(i)=m(i)+e(i) \qquad (1)$$

Where *m(i)* is the regression model estimate and *e(i)* the spatially interpolated residual of the regression.

An ordinary least square (OLS) regression cannot be used here since the constraint of heteroscedasticity of the residuals would not be fulfilled: the elevation at each grid cell is very likely spatially correlated. Therefore generalized least squares (GLS) should be used instead of OLS models (Hengl et al., 2007), since they do not rely on such a constraint. Here a GLM model is considered for the relationship between rain gauge and elevation data (Tramblay et al., 2016). For each day, if the correlation between the precipitation measured at the rain gauges and satellite precipitation is significant at the 10% confidence level, a GLS model is built and the residuals of the model are estimated by simple kriging with known mean (0). The variogram for the residuals is estimated for each time step, in a similar way as for the OK interpolation method explained above.

3.2 The SAFRAN reanalysis

SAFRAN is a high-resolution atmospheric analysis system developed at Météo France (Durand et al., 1993), based on an optimal-interpolation algorithm. It had been initially designed to provide atmospheric forcing data in mountainous areas for avalanche hazard forecasting (Durand et al., 1999) and it was then extended to France (Quintana-Seguí et al.; 2008, Vidal et al. 2010) for hydro-meteorological applications (Habets et al., 2008). Later, SAFRAN was applied in Spain (Quintana-Segui et al., 2016), where it was shown that its precipitation fields are comparable to those of SPAIN02 (Quintana-Seguí et al., 2017), which is based on interpolation method that takes the relief into account.

SAFRAN performs its analysis in two steps. First, the analysis is performed on climatological homogeneous zones. These zones, which in this case have a mean area of ~700 km2, should have weak horizontal gradients of the studied variables. In our application they were manually drawn using the relief, a map of elementary catchments from the HydroShed database (https://hydrosheds.cr.usgs.gov), and our knowledge of the local climate as guiding information. The SAFRAN zones for Tunisia are shown in Figure 1. SAFRAN calculates one value of the studied variable, precipitation in this case, for each vertical level of 300 m. of the zone, using the available meteorological stations, mainly within the zone (but it can, if necessary, use information of neighboring stations). In a second step, the values are interpolated to a regular

5 km grid, considering the vertical gradients. As a result, within each zone, the values of precipitation for each grid point are different only if the altitudes of the grid points are different.

3.3 Validation framework

To validate the different interpolation methods and the SAFRAN reanalysis, two different validation samples are randomly generated with the following constraints:

1. A minimum of 50% of complete years during the period 1979-2015
2. A minimum distance between two stations of two times the average distance between stations

Two samples of 103 stations each have been generated with these criterions. Then, for each station, the relative bias and the Spearman correlation coefficient between daily rainfall amounts, daily rainfall occurrence from the interpolated data and the observed data at the validation stations are computed.

4. RESULTS

4.1 Comparison of interpolation methods on the two validation samples

The quality check and the homogeneity test performed on the whole station database led to select 960 stations with at least 5 years of data between 1979 and 2015, shown in Figure 2. These 960 stations have been used in the different interpolation method and to build the SAFRAN reanalysis. From Figure 3, it can be seen that the efficiency of the different methods is not dependent on the validation sample, with similar results obtained with the two samples. On average, the OK method seems to perform slightly better, in terms of relative bias and correlation than the other spatial interpolations methods. However, the different interpolation methods provide very similar estimates, due to the high spatial density of observed precipitation notably for the northern part of Tunisia. It is worth to observe also that the use of elevation as a covariate in the RK approach does not improve the efficiency of this method by comparisons with the other interpolation schemes. On average, the SAFRAN products provide the lowest bias and the highest correlations on daily amounts and occurrence. When considering the spatial distribution of the results, it can be seen than SAFRAN and the NN method provide the most consistent estimations across Tunisia, while the other methods have a very strong bias in the southern stations (Figure 4). For correlation (Figure 5), for some stations located in the North there is a high correlation close to one, but for the other regions the correlation patterns seems less organized than the bias observed in Figure 4. Overall, there is a clear North/South behavior, with degraded performances in southern stations, due to the lower density of stations but also more arid conditions than in the North. This is exemplified in Figure 6; south of 35°N the relative bias in validation for both samples is very high and for most stations is exceeding 100%. Therefore the use of spatially interpolated data in these regions is not recommended.

4.1 Comparison between SAFRAN and EOBS

The comparison has been performed between the Rx1day, R1mm and PRCPTOT indices computed from EOBS and SAFRAN annually, the inter-annual means of the indices between 1979 and 2015 are compared by computing a relative difference (Figure 7). It must be noted that the EOBS dataset contains only 13 stations for Tunisia (Tabarka, Bizerte, Tunis, Kelibia, Jenbdouba, Kairouan, Monastir, Gafsa, Sfax, Gabes, Djerba, Remada) as shown in Figure 2. This implies very smooth interpolation surfaces, not taking into account the regional differences due to orography or local climate characteristics. This is particularly true for PRCPTOT and R1mm indices, with the spatial gradient from Northwest to Southwest clearly underestimated. On average over the whole country, the relative bias of EOBS compared to SAFRAN is -47% for precipitation totals (PRPTOT) and -59% for the number of wet days (R1mm). For annual maximum precipitation, there is a more complex picture. First, the areas with the higher precipitation intensities do no show a clear spatial organization in both datasets indicating the strong spatial variability of extreme precipitation. Secondly, the areas with high values of Rx1day in the South must be interpreted with care since the density of stations is low in these areas and these patterns might just be caused by the interpolation scheme and do not necessarily represents reality. The relative bias of EOBS for Rx1day is -17% compared to SAFRAN.

5. CONCLUSIONS

We introduced a new high-resolution (5 km) gridded precipitation dataset for Tunisia. It is the first product of this kind, to our knowledge, that covers one of the countries of the Middle East and North African region, which could be useful for various purposes such as climate model evaluation, climate studies, hydrological modelling. A validation experiment has been conducted and it was found that the SAFRAN reanalysis outperforms other standard interpolation methods on two different validation samples. Yet a note of caution must be provided about the uncertainties in the South of Tunisia, due to aridity and the low density of stations that does not guarantee robust spatial interpolation estimates.

Acknowledgements

This work is a contribution to the HYdrological cycle in The Mediterranean EXperiment (HyMeX) program, through INSU-MISTRALS support. The SAFRAN database is available on the HyMex database at the address: DOIXXX

References

Bargaoui Z., Tramblay Y., Lawin E., Servat E., 2014. Seasonal precipitation variability in regional climate simulations over Northern basins of Tunisia. International Journal of Climatology 34, 235-248.

Beck H.E., van Dijk A.I.J.M., Levizzani V., Schellekens J., Miralles D.G., Martens B., de Roo A., 2017. MSWEP: 3-hourly 0.25_ global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data, Hydrol. Earth Syst. Sci., 21, 589–615.

Ben Zaied Y., Binet M-E., 2015. Modelling seasonality in residential water demand: the case of Tunisia, Applied Economics, 47:19, 1983-1996, DOI: 10.1080/00036846.2014.1002896

Dhiba S., Mannaerts C.M., Bargaoui Z., Retsios V., Maathuis B.H.P., 2017. Evaluating the MSG satellite Multi-Sensor Precipitation Estimate for extreme rainfall monitoring over northern Tunisia. Weather and Climate Extremes 16, 14–22.

Durand, Y., E. Brun, L. Mérindol, G. Guyomarc'h, B. Lesaffre, and E. Martin. "A Meteorological Estimation of Relevant Parameters for Snow Models." Ann. Glaciol. 18 (1993): 65–71.

Durand, Y., G. Giraud, E. Brun, L. Mérindol, and E. Martin. "A Computer-Based System Simulating Snowpack Structures as a Tool for Regional Avalanche Forecasting." J. Glaciol. 45, no. 151 (1999): 469–84.

Droogers P., Immerzeel W.W., Terink W., Hoogeveen J., Bierkens M.F.P., van Beek L.P.H., Debele B., 2012. Water resources trends in Middle East and North Africa towards 2050, Hydrol. Earth Syst. Sci., 16, 3101–3114.

Fantini A, Raffaele F, Torma C, Bacer S, Coppola E, Giorgi F, Verdecchia M, 2016. Assessment of multiple daily precipitation statistics in ERA-interim driven Med-CORDEX and EURO-CORDEX experiments against high resolution observations. Clim Dyn. https://doi.org/10.1007/s00382-016-3453-4.

Fathalli, B., Pohl, B., Castel, T., Safi M.J., 2018. Errors and uncertainties in regional climate simulations of rainfall variability over Tunisia: a multi-model and multi-member approach. Clim Dyn, https://doi.org/10.1007/s00382-018-4150-2.

Feki H, Slimani M, Cudennec C, 2012. Incorporating elevation in rainfall interpolation in Tunisia using geostatistical methods. Hydrol Sci J 57:1294–1314.

Flaounas, E., Drobinski, P., Borga, M., Calvet, J., Delrieu, G., Morin, E., Tartari, G., Toffolon, R., 2012. Assessment of gridded observations used for climate model validation in the Mediterranean region: the HyMeX and MED-CORDEX framework, Environ. Res. Lett., 7, 024017, doi:10.1088/1748-9326/7/2/024017.

Goovaerts P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology 228: 113–129.

Guijarro, J.A., 2017. Daily series homogenization and gridding with Climatol v.3. 9th Seminar for homogenization and quality control in climatological databases and 4th conference on spatial interpolation techniques in climatology and meteorology. (Budapest, April 3-7), WMO WCDMP-No.85, pp.175-182.

Guijarro, J.A., 2018. Homogenization of climate series with Climatol. Available at http://www.climatol.eu/homog_climatol-en.pdf (last accessed : January 2019).

Habets, F., Boone, A., Champeaux, J. L., Etchevers, P., Franchistéguy, L., Leblois, E., Ledoux, E., Le Moigne, P., Martin, E., Morel, S., Noilhan, J., Seguí, P. Q., Rousset-Regimbeau, F., Viennot, P., 2008. The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, J. Geophys. Res.-Atmos., 113, D06113, doi:10.1029/2007JD008548.

Haddadin M.J., 2001. Water Scarcity Impacts and Potential Conflicts in the MENA Region, Water International, 26:4, 460-470, DOI: 10.1080/02508060108686947.

Harris I., Jones P.D., Osborn T.J., Listera D.H., 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset Int. J. Climatol. 34: 623–642.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, New M., 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, J. Geophys. Res., 113, D20119, doi:10.1029/2008JD010201.

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007. About regression–kriging: from equations to case studies. Comput. Geosci. 33, 1301–1315.

Herrera, S., Gutiérrez, J.M., Ancell, R., Pons, M. R., Frías, M.D., Fernández, J. (2012). Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02), Int. J. Climatol. 32, 74–85.

Kummerow, C.D., Barnes, W., Kozu, T., Shiue, J., Simpson, J., 1998. The Tropical Rainfall Measuring Mission (TRMM) sensor package. J. Atmos. Ocean. Technol. 15, 809–817.

Lebel T, Bastin G, Obled C, Creutin JD., 1987. On the accuracy of areal rainfall estimation. Water Resources Research 23: 2123–2134.

Mehta, A.V., Yang, S., 2008. Precipitation climatology over Mediterranean Basin from ten years of TRMM measurements. Adv. Geosci. 17, 87–91.

Nastos P.T., Kapsomenakis J., Douvis K.C., 2013. Analysis of precipitation extremes based on satellite and high-resolution gridded data set over Mediterranean basin, Atmospheric Research 131, 46–59.

Prein, A. F. and Gobiet, A. (2017). Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. Int. J. Climatol., 37, 305-327.

Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., Morel, S. 2008. Analysis of near-surface atmospheric variables: Validation of the SAFRAN analysis over France, J. Appl. Meteorol. Clim., 47, 92–107.

Quintana-Seguí, P., Peral, M. C., Turco, M., Llasat, M.-C., Martin, E., 2016: Meteorological Analysis Systems in North-East Spain: Validation of SAFRAN and SPAN. J. Environ. Inform., 27, 116–130.

Quintana-Seguí, P., Turco M., Herrera S., Miguez-Macho, G., 2017. Validation of a New SAFRAN-Based Gridded Precipitation Product for Spain and Comparisons to Spain02 and ERA-Interim. Hydrology and Earth System Sciences 21, 2187–2201. https://doi.org/10.5194/hess-21-2187-2017.

Ragab R., Prudhomme C., 2002. Climate Change and Water Resources Management in Arid and Semi-arid Regions: Prospective and Challenges for the 21st Century, Biosystems Engineering 81 (1), 3-34.

Romera R, Sánchez E, Domínguez M, Gaertner MÁ, Gallardo C, 2015. Evaluation of present-climate precipitation in 25 km resolution regional climate model simulations over Northwest Africa. Clim Res 66(2):125–139.

Schamm K., Ziese M., Becker A., Finger P., Meyer-Christoffer A., Schneider U., Schröder M., Stender P., 2014. Global gridded precipitation over land: a description of the new GPCC First Guess Daily product, Earth Syst. Sci. Data, 6, 49–60.

Schilling, J., Freier, K. P., Hertig, E., Scheffran, J., 2012. Climate change, vulnerability and adaptation in North Africa with focus on Morocco. Agriculture, Ecosyst. Environ., 156, 12–26.

Tramblay Y., Thiemig V., Dezetter A., Hanich L., 2016. Evaluation of satellite-based rainfall products for hydrological modelling in Morocco, Hydrological Sciences Journal 61(14), 2509-2519.

Tramblay Y., Jarlan L., Hanich L., Somot S., 2018. Future scenarios of surface water resources availability in North African dams. Water Resources Management 32(4), 1291-1306.

Turco, M., Zollo, A. L., Ronchi, C., De Luigi, C., Mercogliano, P., 2013. Assessing gridded observations for daily precipitation extremes in the Alps with a focus on northwest Italy, Nat. Hazards Earth Syst. Sci., 13, 1457-1468

Vidal J.-P., Martin E., Franchistéguy L., Habets F., Soubeyroux J.-M., Blanchard M., Baillon M., 2010. Multilevel and multiscale drought reanalysis over France with the Safran-Isba-Modcou hydrometeorological suite. Hydrol. Earth Syst. Sci. 14, 459–478.

Vidal JP, Martin E., Franchistéguy L., Baillon M., Soubeyroux J.M., 2010. A 50-Year High-Resolution Atmospheric Reanalysis over France with the Safran System. International Journal of Climatology 30, 1627–1644. https://doi.org/10.1002/joc.2003.

Zittis G., 2017. Observed rainfall trends and precipitation uncertainty in the vicinity of the Mediterranean, Middle East and North Africa, Theoretical and Applied Climatology. https://doi.org/10.1007/s00704-017-2333-0.

FIGURES

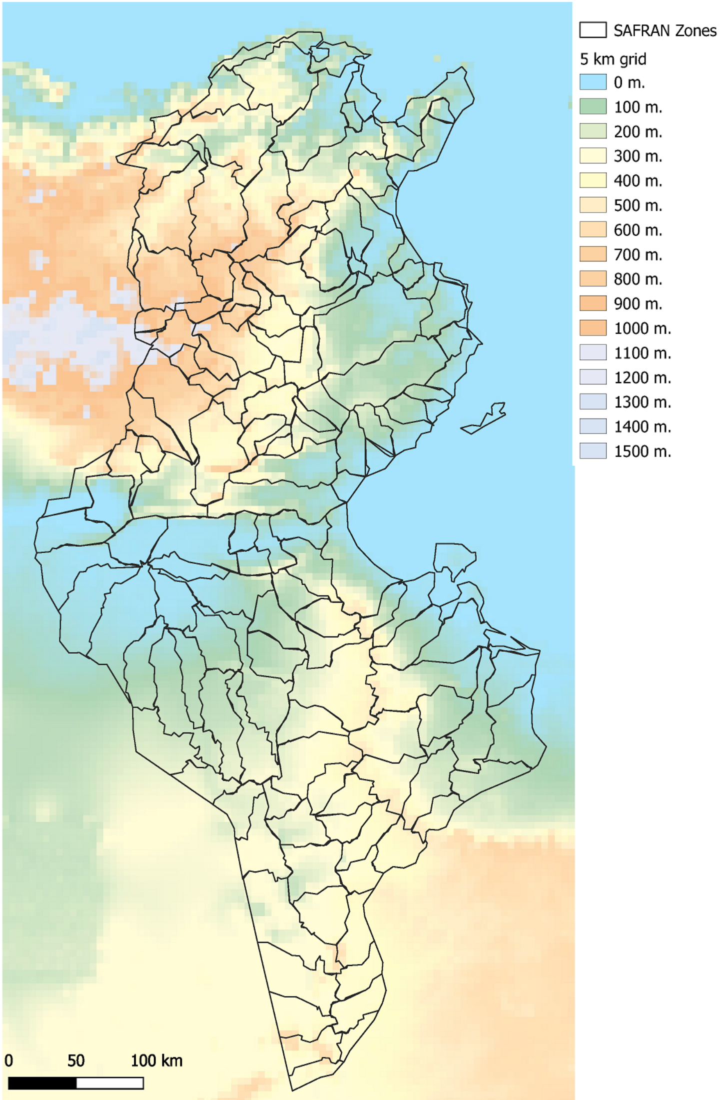

Figure 1: Map of the SAFRAN zones in Tunisia with elevation

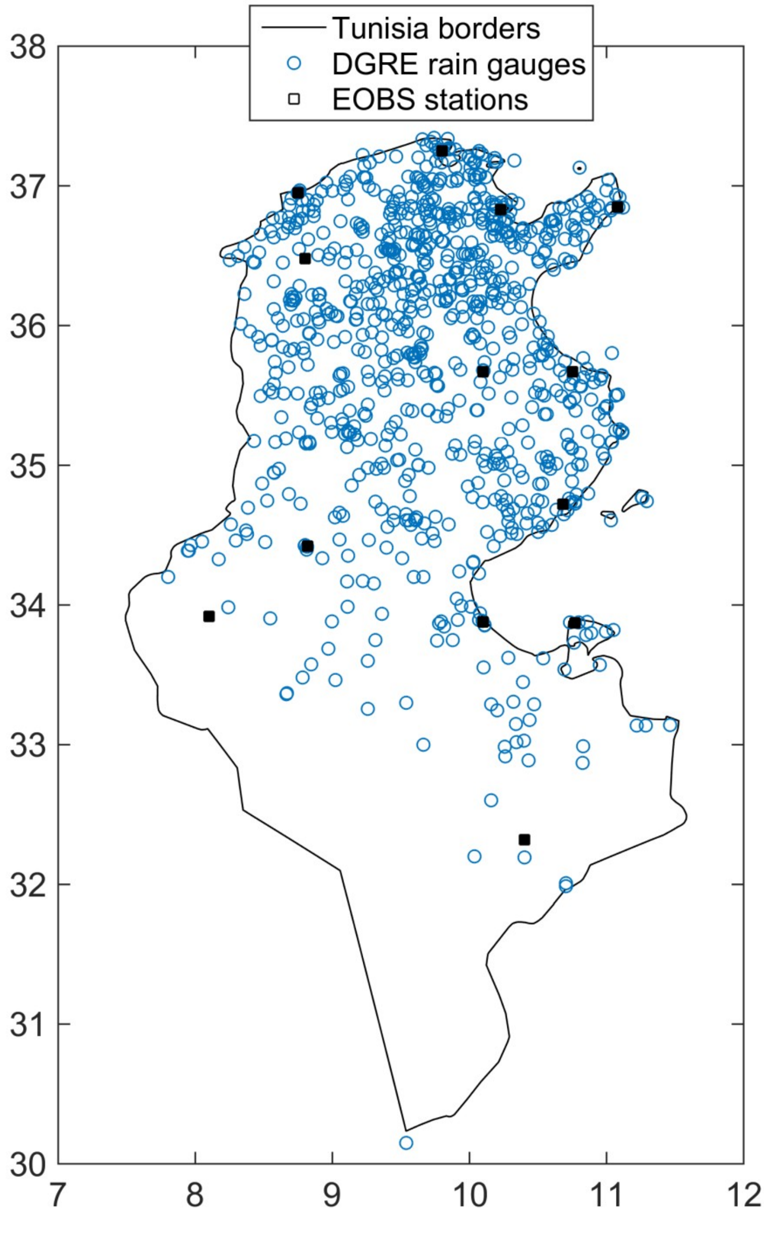

Figure 2: Location of the 960 rain gauges available in Tunisia and stations included in the EOBS dataset

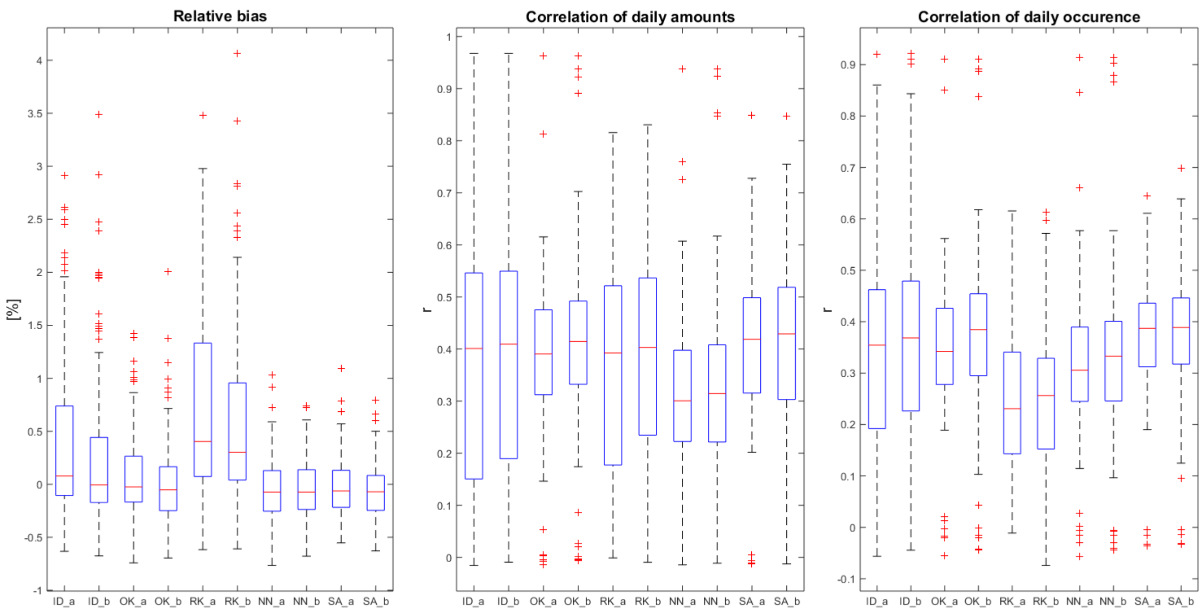

Figure 3: Box-plots of the relative bias, correlation of daily amounts and correlation of daily occurrence for the two validation samples (a and b). ID: Inverse-distance, OK: Ordinary Kriging, RK: Residual kriging, NN: Nearest neighbors, SA: SAFRAN.

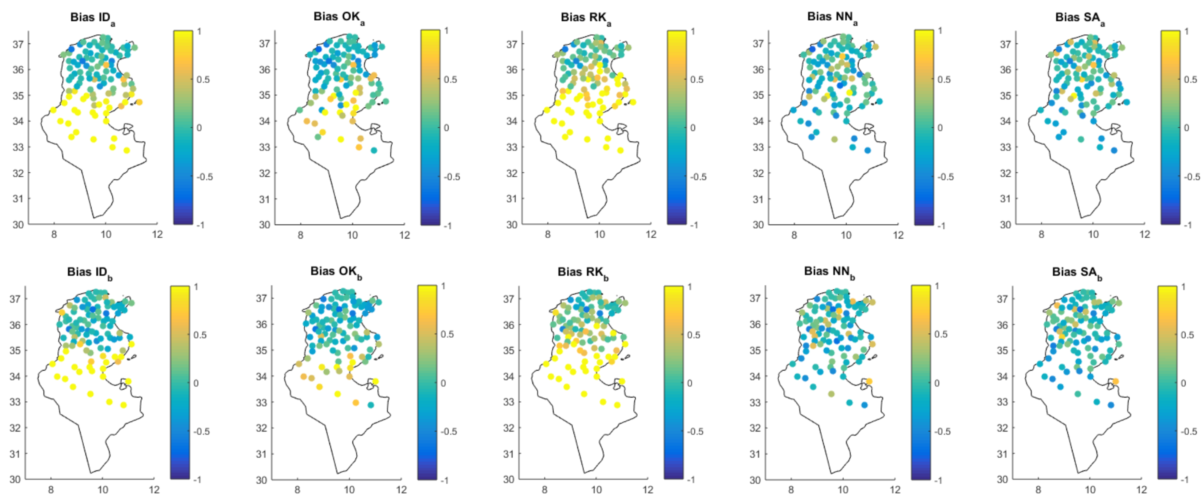

Figure 4: Maps of the relative bias between observed and interpolated precipitation for the two validation samples

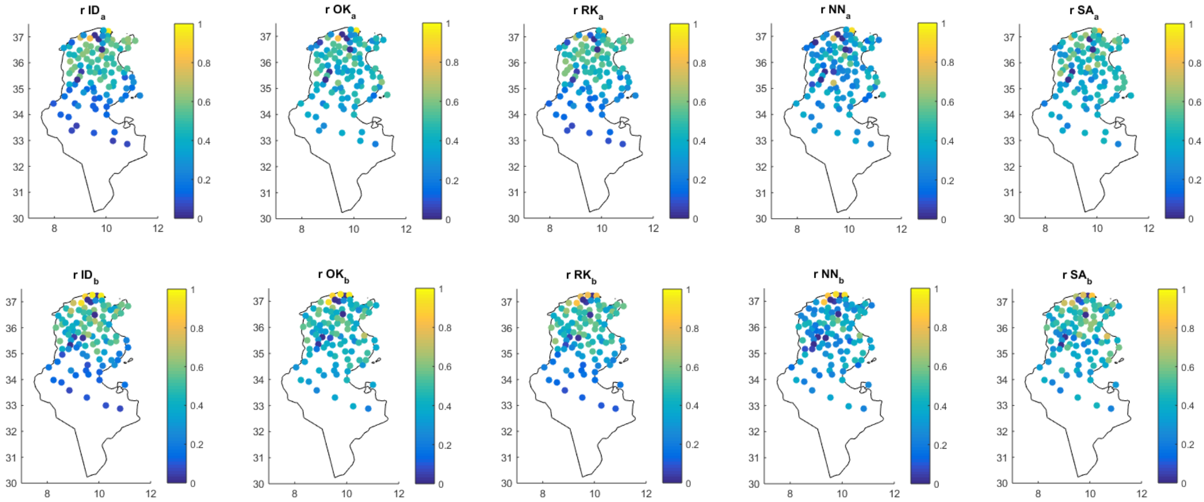

Figure 5: Maps of the daily correlation between observed and interpolated precipitation for the two validation samples

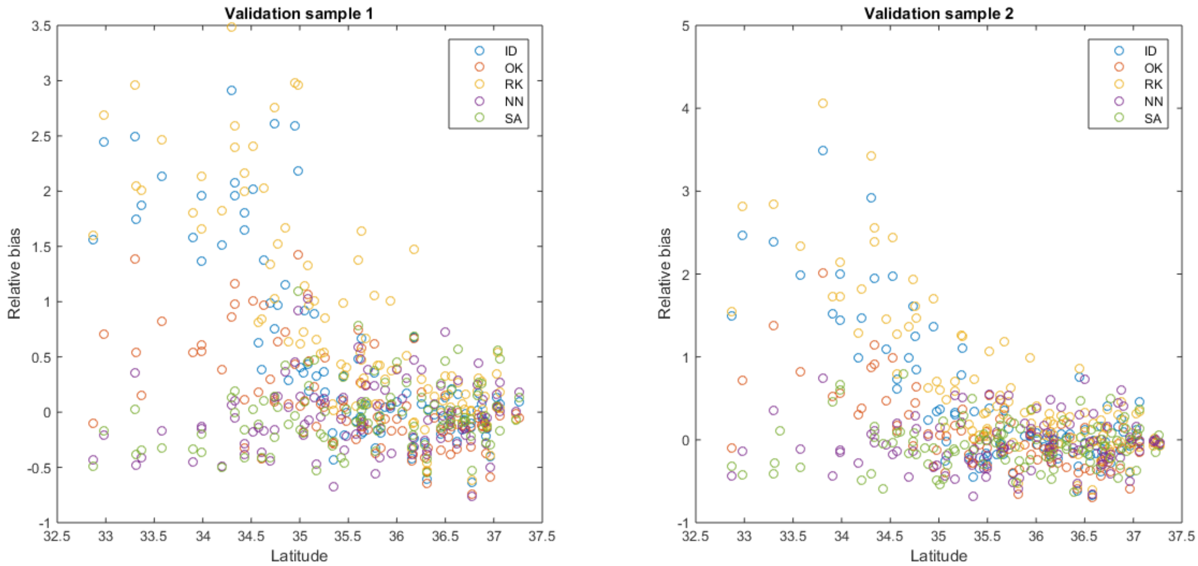

Figure 6: Relative bias for the two validation samples in relation with latitude

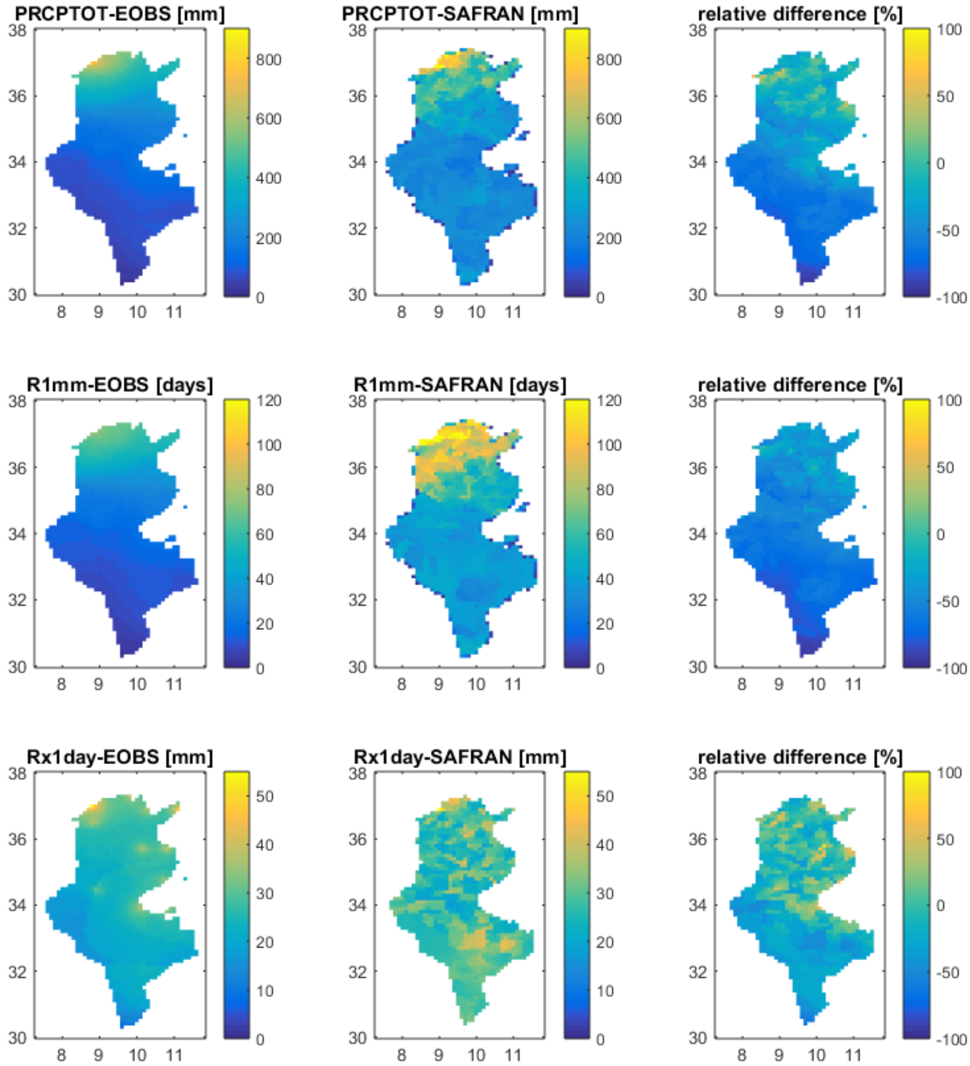

Figure 7: Comparison between the PRCPTOT, R1mm and Rx1day indices computed from EOBS or SAFRAN