Scenarios of hydrometeorological variables based on auxiliary data for water stress retrieval in central Tunisia

Nesrine Farhani [1,2] & Julie Carreau [3] & Gilles Boulet [1] & Zeineb Kassouk [2] & Bernard Mougenot [1] & Michel Le Page [1] & Zohra Lili Chabaane [2] & Rim Zitouna [4]

[1] *Université de Carthage / Institut National Agronomique de Tunisie/ LR17AGR01-GREEN-TEAM, Tunis, Tunisie*
[2] *Centre d'Études Spatiales de la Biosphère, Université de Toulouse, CNRS, CNES, IRD, UPS, Toulouse, France 10*
E-mail: farhani.nesrine@gmail.com
[3] *HydroSciences Montpellier (HSM), CNRS/IRD/UM1/UM2, Place Eugéne Bataillon, 34095 Montpellier, France*
E-mail: Julie.carreau@ird.fr

Résumé. L'estimation de l'évapotranspiration est nécessaire pour mieux gérer les ressources en eau. De plus, le stress hydrique dérivé des données de l'infrarouge thermique est particulièrement utile pour une bonne gestion de l'agrosystème et des besoins en irrigation. L'évapotranspiration et le stress hydrique peuvent être simulés par un modèle de bilan énergétique à double source en se basant sur des observations climatiques et des données satellitaires. Cependant, ces observations climatiques peuvent être insuffisantes pour tenir compte de la variabilité spatiale et temporelle de la zone d'intérêt dû à la rareté des stations agrométéorologiques, à des périodes d'observation trop courtes et à la présence des données manquantes. Notre objectif est d'adapter un générateur stochastique de conditions météorologiques qui se base sur des données de réanalyse afin de simuler des scénarios spatio-temporels de différentes variables climatiques. Le générateur stochastique sert à effectuer l'imputation des données manquantes et à effectuer une projection de la série climatique dans le passé. Nous comparons cette méthode avec des méthodes de correction de biais qui exploitent également les données de réanalyse. Nous considérons une méthode de correction de biais univariée et multivariée, CDFt et MBCn, respectivement. Ces approches statistiques sont utilisées pour générer des scénarios climatiques dans la région de Kairouan, située au centre de la Tunisie et caractérisée par un climat semi-aride. Les différents scénarios sont évalués et comparés en termes de leur capacité à reproduire les caractéristiques des observations climatiques et également en termes de leur capacité à reproduire la simulation de l'évapotranspiration et du stress hydrique.

Mots-clés. évapotranspiration, ré-analyses, imputation, correction de biais

Abstract. Characterization of plant water use, generally determined by estimating evapotranspiration, is needed to better manage water resources. Furthermore, water stress derived from remote sensing data in the thermal infrared domain is particularly informative for monitoring agrosystem health and adjusting irrigation requirements. Evapotranspiration and water stress can be simulated by a dual source energy balance model from climatic observations and satellite information. However, it may occur that the available climatic observations are insufficient to account for the spatial and temporal variability of the area of interest due to the sparsity of gauged networks, the lack of long observation periods and the presence of numerous gaps. We aim to adapt a stochastic weather generator that relies on low resolution ERA reanalysis data to provide spatio-temporal scenarios of multiple climatic variables. This stochastic generator serves to perform imputation of missing data and to drive simulation in order to extend the climatic time series in the past. We compare this method with two bias correction methods which exploit also ERA reanalyses. We consider a univariate and a multivariate bias correction method, CDFt and MBCn respectively, to assess to added-value of correcting simultaneously the climatic variables. The statistical approaches are applied to generate climatic scenarios in the Kairouan area in central Tunisia which is subject to semi-arid climate. The different scenarios are evaluated and compared in terms of their ability to reproduce several features of the climatic observations and also in terms of their ability to reproduce the simulation of evapotranspiration and water stress with a dual source energy balance model.

Keywords. evapotranspiration, reanalysis, imputation, bias correction

In arid and semi-arid areas, water is a major limitation factor for agricultural production. The vulnerability to climate change and, in particular, to drought periods is high. Indeed, these areas are characterized by a short rainy season and strong irregularity in time and space of precipitation events. This induces more frequent annual and intra-seasonal droughts. Natural variations in climate water cycle affect the availability of water and form the main driver of droughts. Agriculture and irrigation are particularly influenced by water deficit.

An important issue in climate and hydrology is to improve the monitoring of droughts and the prediction of their occurrence in the future. This requires a better understanding of the physical mechanisms that lead to this phenomenon. To analyze the actual water use over a long period in the past is a key component of better drought management. It allows to establish a temporal analysis and a monitoring of agricultural practices. Another key component is the quantification of droughts to help analyze drought occurrences. Stress indices can be established to quantify periods of droughts according to their frequency, intensity, spatial extension and duration. To this end, long series of drought observations are required.

In this work, we are interested in the vegetation stress. Indeed, the state of vegetation is generally representative of environmental water stress especially in arid and semi-arid areas (Sheffield & Wood, 2012). Plant water use is generally computed based on evapotranspiration estimation which is the preponderant component of the terrestrial water balance and is a key factor for scarce water resources management. The quantitative state of drought is defined by a stress index. Owing to the temporal and spatial scales of climate variability, evapotranspiration and water stress index must be monitored at subdaily to daily scales. Therefore, we choose to compute evapotranspiration using energy balance methods that combine from medium to low resolution remote sensing (RS) data. RS data in the thermal infrared domain is particularly informative for monitoring agrosystem health and adjusting irrigation requirements. In water deficit condition, plants reduce their transpiration rates to preserve the remaining water. This reduced evaporative part in the leaf surfaces induces a detectable thermal signal of elevated canopy temperatures that can be measured from thermal infrared sensors (Jones *et al.* , 2009). Thus, surface energy balance uses surface temperature to solve the coupled equations of sensible, latent and heat energy (Sheffield & Wood, 2012).

We rely on a dual source energy balance model that allows retrieval of separate estimates of evaporation and transpiration. High temporal and spatial resolution data are required. For this reason, we need long time series of hydrometeorological variables (air temperature, relative air humidity, global radiation and wind speed) and satellite information (NDVI, LAI, albedo and surface temperature obtained from the TERRA and AQUA sensors of the MODIS satellite). However, the available observations are often insufficient to account for the spatial and temporal variability of the area of interest due to the sparsity of gauged networks, the lack of long observation periods and the presence of numerous gaps.

To perform imputation of missing data and to extend the observation period in the past, we employ a stochastic weather generator that relies on low resolution reanalysis data to generate scenarios of hydrometeorological variables at sub-daily temporal resolution. The stochastic weather generator is based on generalized linear models (GLM) for each hydrometorological variable with a suitable probability distribution (Normal, Gamma or Binomial). In addition to the low resolution hydrometeorological variables provided by the reanalysis data, other covariates are used to introduce inter-variable dependencies, deterministic effects (geographical information, seasonal and diurnal cycles) and memory effects, i.e. lagged values (Chandler, 2015). The memory effects are computed either directly from the variable of interest, from spatial averages (the average of the values at all the sites at the given time step), from moving averages with a window of one day (48 steps of time) or less, or from a combination of a spatial and moving averages. These effects allow to take into consideration the past behavior of the variables.

We compare the proposed stochastic weather generator with two bias correction methods which also exploit low resolution reanalysis data. Indeed, the reanalysis data provide a multivariate, spatially complete and coherent record of the global atmospheric circulation (Dee *et al.* , 2011). In addition, they are available for a long period in the past (from 1979 till now). Nevertheless, its spatial resolution is low from 31 km (ERA 5) to 75 km (ERA interim) thus local variability is not accounted for (Hooker *et al.* , 2018). To use such reanalysis data in energy balance models, bias correction methods can be applied to account for the difference in spatial resolution between reanalyses and observations from the gauged network. By working on anomalies of diurnal cycles, existing bias correction methods can be adapted to sub-daily temporal resolution. We consider a univariate and a multivariate bias correction method to assess the added-value of correcting simultaneously the hydrometeorological variables. We selected, for the univariate method, the CDFt cumulative distribution function approach developed by Michelangeli *et al.* (2009). For the multivariate method, we use the MBCn approach from Cannon (2018). We rely on a split sample method where a statistical relationship is calibrated between the reanalyses and the observations on a reference period (2012-2016) and applied on a study period (2000-2016).

The statistical approaches are applied to generate hydrometeorological scenarios in the Kairouan area in central Tunisia which is subject to semi-arid climate. The different scenarios are evaluated and compared in terms of their ability to reproduce several features of the hydrometeorological observations and also in terms of their ability to reproduce the simulation of evapotranspiration and water stress with the dual source energy balance model.

References

Cannon, A. J. 2018. Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate dynamics*, **50**(1-2), 31–49.

Chandler, R. 2015. A multisite, multivariate daily weather generator based on Generalized Linear Models. *User guide : R package.*

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., & Vitart, F. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, **137**(656), 553–597.

Hooker, J., Duveiller, G., & Cescatti, A. 2018. A global dataset of air temperature derived from satellite remote sensing and weather stations. *Scientific data*, **5**, 180246.

Jones, H. G., Serraj, R., Loveys, B. R., Xiong, L., Wheaton, A., & Price, A. H. 2009. Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. *Functional Plant Biology*, **36**(11), 978–989.

Michelangeli, P.-A., Vrac, M., & Loukos, H. 2009. Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophysical Research Letters*, **36**(11).

Sheffield, J., & Wood, E. F. 2012. *Drought: past problems and future scenarios*. Routledge.