Sub-daily stochastic weather generator based on reanalyses for water stress retrieval in central Tunisia

Nesrine Farhani, Julie Carreau, Zeineb Kassouk, Bernard Mougenot, Michel Le Page, Lili -Chabaane, Rim Zitouna, Gilles Boulet

Sub-daily stochastic weather generator based on reanalyses for water stress retrieval in central Tunisia

Nesrine Farhani[a,b,*], Julie Carreau[c], Zeineb Kassouk[a], Bernard Mougenot[b], Michel Le Page[b], Zohra Lili-Chabaane[a], Rim Zitouna-Chebbi[d], Gilles Boulet[b]

[a] *Université de Carthage, Institut National Agronomique de Tunisie, LR17AGR01-GREEN-TEAM, Tunis, Tunisia*
[b] *Centre d'Études Spatiales de la Biosphère, Université de Toulouse, CNRS, CNES, IRD, UPS, INRAE, Toulouse, France*
[c] *HydroSciences Montpellier, Univ. de Montpellier/CNRS/IRD, Montpellier, France*
[d] *LR VENC, INRGREF, University of Carthage, Rue Hedi Karray 2080, Ariana, Tunisia*

Abstract

In semi-arid areas, evapotranspiration that characterizes plant water use and water stress are needed to better manage water resources and agrosystem health. They both can be simulated by a dual source energy balance model that relies on hydro-meteorological variables and satellite data. Available hydro-meteorological observations may often be insufficient to account for the variability present in the study area. Our aim is to adapt a stochastic weather generator (SWG) driven by large-scale reanalysis data to semi-arid climates and to the sub-daily resolution. The SWG serves to perform consistent gap-filling and temporal extension of multiple hydro-meteorological variables. It is compared with two state-of-the-art bias correction methods applied to large-scale reanalysis data. The surrogate series that are either produced by the SWG and the bias correction methods with a cross-validation scheme or taken as the un-processed reanalysis data, are evaluated in terms of their ability to reproduce the statistical properties of the hydro-meteorological observations. They are also used to constrain a dual source energy balance model and compared in terms of estimated evapotranspiration and water stress.

Keywords: hydro-meteorological variables, evapotranspiration, semi-arid climate, gap-filling, bias correction methods, dual energy balance model, ERA5

1. Introduction

In arid and semi-arid areas, water is a major limitation factor for agricultural production. Indeed, these areas are characterized by short rainy seasons and strong irregularity of precipitation events in time and space [9]. This generates natural variations in the water cycle that affect the availability of water, irregularities in agricultural production [44] and constitutes the main driver of agricultural droughts. In these areas, the vegetation health status is generally representative of water availability [46]. Therefore, an important issue concerns the monitoring of droughts derived from the state of the vegetation over a long period in the past.

To achieve proper monitoring, a better understanding of the physical mechanisms leading to droughts,

*Corresponding author
Email address: farhani.nesrine@gmail.com (Nesrine Farhani)

in particular actual water use and water stress, is required. Plant water use can be obtained from the estimation of evapotranspiration which is the preponderant component of the terrestrial water balance and is crucial for scarce water resources management. On the other hand, detecting and quantifying drought events in the past allows to understand drought mechanisms and to predict drought occurrences. To this aim, water stress index or anomalies thereof can be computed. Water stress index reflects the state of the plant ranging from a no-stress (index equals to zero) to a stress condition (index equal to one) [26]. The probability that the water stress index exceeds a given threshold could be useful for drought monitoring. Anomalies of the water stress index give the status of the vegetation in comparison to the best and worst vegetation conditions for a particular monthly period over many years [7]. They help to make a relative temporal assessment of the severity of drought periods according to frequency, intensity, spatial extension and duration. Evapotranspiration and water stress indices may be estimated thanks to energy balance models.

Dual source energy balance models provide more robust estimates of evapotranspiration as well as water stress than most models when meteorological forcing and vegetation cover are accurately known [17]. This results from the fact that they account for the interactions between the soil and the vegetation that are two contrasting sources of turbulent and radiation fluxes. They combine medium to low resolution Remote Sensing (RS) data. RS data from the thermal infrared (TIR) domain is particularly informative for monitoring agrosystem health and adjusting irrigation requirements. Indeed, in water deficit conditions, plants reduce their transpiration rate to preserve the remaining water. This reduced transpiration triggers an increase in the leaf temperature that can be measured from thermal infrared sensors [29]. Most energy balance models that estimate evapotranspiration and water stress from TIR data solve the latent heat flux from a residual term of the surface energy budget. As a result, total fluxes are derived and not the soil and vegetation components of the fluxes. Dual source energy balance models are able to compute separate energy budgets for the soil and the vegetation and therefore retrieve both evaporation and transpiration. Those fluxes correspond respectively to the water loss from the soil surface and from the root zone [4]. Two sources of data are required in these models : observations of hydro-meteorological variables (air temperature, relative air humidity, global radiation and wind speed), usually measured at gauged stations, and satellite information (Normalized Difference Vegetation Index, Leaf Area Index, albedo and surface temperature). However, hydro-meteorological observations are often insufficient to account for the strong temporal and spatial variability of the water fluxes in semi arid areas due to the sparsity of gauged networks, the lack of long observation periods and the presence of numerous gaps.

Statistical downscaling methods applied to reanalysis data can serve to generate surrogate series of hydro-meteorological variables that either fill the gaps in the observation period or extend the observation period in the past. Reanalysis data combine observations and models thereby providing a multivariate,

spatially complete and coherent record, without gaps, of the global atmospheric circulation [18, 23]. In addition, reanalysis are available for a long period in the past (from 1950 till now). Nevertheless, their spatial resolution is too low (horizontal resolution of 31 km for ERA5 product [23]) and thus local-scale variability is not sufficiently accounted for [24]. To exploit such reanalysis data in energy balance models, statistical downscaling methods may be used. These methods have been developed to account for the scale mismatch between global circulation models' simulations, that are the major source of information concerning climate change, and impact studies [35, 8]. Indeed, hydro-meteorological series at finer time scales are useful to select agricultural practices in response to water availability. Statistical downscaling consists in developing quantitative relationships between large-scale atmospheric variables (predictors) and local surface variables (predictands) in order to generate high resolution time series [50, 49]. There are two main families of statistical downscaling approaches : *perfect-prog* methods and *model output statistics* methods.

Perfect-prog methods are downscaling methods that require temporal synchronicity between large- and local-scale data. A prominent class of perfect-prog methods are regression models that can represent linear or nonlinear relationships between predictands and the large scale atmospheric predictors [50]. Linear regression is the most basic and frequently used predictive model [22]. Nevertheless, it fails to reproduce extreme events and observed variance [50]. Weather classification methods or weather typing, another class of perfect-prog methods, consist in classifying large scale weather into dominant weather types or states according to their synoptic similarity [50]. Local weather is defined by local situations from the observation period with weather states matching the current one. The weather types are necessarily the same on all periods considered but their frequency might be different [49]. This approach is very useful to provide multi-site and multi-variate series. However, it may be unsatisfactory whenever observation series are too short or the number of classifying predictors is large [47]. Stochastic weather generators (SWGs) can be used as perfect-prog methods as well. They are a class of flexible statistical models based on probability distribution functions [1]. They seek to reproduce observed statistical properties and are better suited to account for the variability in the observations. SWGs may provide high resolution surrogate series by introducing large-scale information (such as reanalysis) as covariates that influence the parameters of the probability distribution functions. Indeed, they allow long simulations of all hydro-meteorological variables at sub-daily scales [50]. These simulated series can be used to constrain water and energy balance models [25]. The main limitation in SWGs is the difficulty to adjust the parameters in a physically realistic and consistent manner [27]. Nevertheless, SWGs are promising flexible downscaling methods that can encompass regression models and weather type methods.

Model Output Statistics (MOS) methods are downscaling methods that aims to link statistical properties and do not require temporal synchronicity. Bias correction methods, the most important type of MOS

methods, aim to transform the distribution of low spatial resolution hydro-meteorological variables in order to match the distribution of the corresponding high resolution hydro-meteorological variables [48]. Biases, which are systematic differences in distributional properties (mean, variance or quantile), are computed between the large- and local-scale variable of interest in a reference period and removed from the whole study period [13]. Bias correction methods do not seek to produce values of the local-scale variable that are perfectly synchronized with the observations. They rather aim at yielding values that should be closer to the observations in terms of distribution. Initial bias correction approaches were univariate [38] and performed bias correction variable per variable, i.e. not explicitly accounting for inter-variable dependency while most recent bias correction approaches are multivariate [13].

In this work, we choose to adapt the multi-variable stochastic weather generator (SWG) proposed in Chandler [15] to the sub-daily resolution to perform gap-filling and temporal extension in order to estimate the water stress in semi-arid areas. The SWG is based on generalized linear models (GLMs) for each hydro-meteorological variable with a suitable probability distribution (Normal, Gamma or Binomial). The inter-variable dependencies are taken into account by including a subset of hydro-meteorological variables (excluding the one being modelled) in the covariates of the GLMs. Such interactions between the variables must be preserved to maintain consistency and realism [25]. As large-scale covariates taken from reanalysis data are introduced in the GLMs for all hydro-meteorological variables, the SWG can be considered as a statistical downscaling approach. Additional covariates are used in order to reproduce deterministic effects (geographical information, seasonal and diurnal cycles) and temporal persistence [15]. We rely on a two-step backward selection procedure to determine a parsimonious set of relevant covariates. The proposed multi-variable sub-daily SWG is compared to two state-of-the art bias correction approaches applied to anomalies of hydro-meteorological variables over the diurnal cycles.

In our comparative analyses, we considered CDF-t, a univariate bias correction method, and MBCn, a multivariate bias correction approach. The CDF-t method allows non-linear corrections [38]. It relies on a transformation of the distribution function of the variable at low resolution based on empirical distribution functions. The CDF-t method is designed to perform bias correction on one hydro-meteorological variable at a time. It has shown to perform well at reproducing the statistical distributions of the local series. However, the spatial, temporal and inter-variable dependence structures of the corrected series may be misrepresented which may lead to unrealistic situations [48]. This can be a major limitation for hydrological applications in which consistency between hydro-meteorological variables is crucial. The N-dimensional probability density function transform (MBCn) is a multivariate bias correction algorithm that considers jointly multiple hydro-meteorological variables [13]. All the characteristics of the observed continuous multivariate distribution of the local variable are transferred to the corresponding multivariate distribution of the simulated variables. The MBCn algorithm looks iteratively for linear combinations of

the variables and performs bias correction with a univariate algorithm on the linear combinations rather than on each variable separately. As several hydro-meteorological variables are needed for the energy balance model, multivariate bias correction approaches that correct simultaneously all the variables in order to preserve the inter-variable dependence structure could be useful.

The three proposed statistical downscaling methods (multi-variable sub-daily SWG, univariate and multivariate bias correction methods) are used to simulate surrogate series on periods for which no observations are available at three stations in the Merguellil catchment in central Tunisia. The simulations can be carried over either a long period in order to provide temporal projections, in the past or in the future, of the observed series or over a short number of time steps to perform gap filling. The three surrogate series generated at the stations by the downscaling methods together with a fourth surrogate series taken as the un-processed large-scale reanalysis variables are evaluated and compared in terms two sets of criteria. The first set pertains to features of the hydro-meteorological observations to assess whether the distributions of the intensities and the strength of the inter-variable dependencies are well reproduced. The second set of criteria pertains to features of evapotranspiration and water stress that are estimated by a dual source energy balance model constrained with the surrogate series.

The two main objectives of this paper are (1) simulation of evapotranspiration and water stress over a long period in the past using surrogate series and (2) comparison between the three proposed downscaling methods in terms their ability to mimic hydro-meteorological observations and their impact on the performance of water stress and evapotranspiration retrieval.

2. Multi-variable sub-daily SWG

We developed a stochastic weather generator (SWG) to simulate surrogate series of hydro-meteorological variables that reproduce the climatic variability of the study area. The aim of this work is to define a model for each hydro-meteorological variable based on its physical understanding and preserving its stochastic behavior. Surrogate hydro-meteorological series should be generated through multivariate models and at sub-daily temporal resolution. The proposed SWG builds on the approach proposed by [15] in the `Glimclim` package available in `R` programming environment. The `Glimclim` package is designed for the modeling and simulation of univariate or multivariate daily weather sequences at a single or at multiple sites. It relies on stochastic regression with Generalized Linear Models (GLMs), as described in section 2.1. The dependence between multiple hydro-meteorological variables is modeled by decomposing the multivariate density with the product rule, see (4). Several covariates may be considered in the GLMs to account for temporal and spatial variability. Large-scale atmospheric variables may also be included in the covariates. We have made several developments, explained in the following sub-sections,

and independent stand-alone `R` library `MetGen` which is available upon request, to adapt the generator to our purposes and in particular, to extend it to the sub-daily resolution.

2.1. Stochastic regression with generalized linear models

Each hydro-meteorological variable is modeled by a regression model that is used stochastically for simulation according to its own probability distribution, whose parameters are driven by a large set of covariates, plus a random noise [30]. Indeed, noise can improve the ability of models to simulate climatic variables regimes and seasonal anomalies [42]. Classical linear regression makes the assumption that a hydro-meteorological variable Y given a covariate vector $\mathbf{x} \in \mathbb{R}^d$ is normally distributed :

$$Y|\mathbf{x} \sim \mathcal{N}(\beta\mathbf{x}, \sigma^2), \tag{1}$$

where $\beta \in \mathbb{R}^d$ is a vector of regression parameters and $\sigma > 0$ is the standard deviation. Stochastic regression consists, for a given $\mathbf{x}$, in simulating from Eq. (1) instead of estimating Y by the conditional expectation, i.e. by $\mathbb{E}[Y|\mathbf{x}] = \beta\mathbf{x}$. The covariates $\mathbf{x}$ influence the location parameter of the Normal distribution of which simulation is performed.

Generalized Linear Models (GLMs) are based on a simple transformation of linear regression [37]. They have been applied in several works to model hydrometerological variables in different contexts [16, 2, 6]. GLMs allow the use of different types of probability distributions belonging to the exponential family. Besides the Normal distribution, the stochastic generator relies on the Gamma distribution which is useful to model hydro-meteorological variables that take only positive values such as precipitation. In the generator framework, the associated regression model is

$$Y|\mathbf{x} \sim Gamma(\mu(\mathbf{x}), \nu) \quad \mu(\mathbf{x}) = \exp(\beta\mathbf{x}) \tag{2}$$

where $\mu(\mathbf{x}) > 0$ is the mean parameter of the Gamma distribution and $\nu > 0$ is the shape parameter. The stochastic generator also relies on the regression model associated to the Binomial distribution to model the probability of occurrence of precipitation as a discrete 2-category variable as follows :

$$\mathbb{P}(Y = 1|\mathbf{x}) = (1 + e^{-\beta\mathbf{x}})^{-1} \in [0, 1]. \tag{3}$$

In the SWG, a probability distribution from (1)-(3) must be chosen for each hydro-meteorological variable Y (the local-scale variables in the downscaling framework). Next, suitable covariates $\mathbf{x}$ have to be selected. These will allow the distributions of the hydro-meteorological variables Y to vary in time and in space. The covariate set $\mathbf{x}$ may contain: (1) one or more of the other hydro-meteorological variables in order to maintain inter-variable dependencies, see sub-section 2.2, (2) large-scale variables

(from reanalysis data in our case), (3) seasonal and (4) diurnal cycles, (5) geographical information and (6) memory effects, see sub-section 2.3.

2.2. Inter-variable dependencies

In order to model the inter-variable dependencies among p hydro-meteorological variables $Y_1, Y_2, \ldots, Y_p$, the stochastic generator relies on a decomposition of the p-dimensional multivariate distribution given by the product rule :

$$\mathbb{P}(Y_1, Y_2, \ldots, Y_p) = \mathbb{P}(Y_1) \prod_{i=2}^{p} \mathbb{P}(Y_i | Y_{i-1}, \ldots, Y_1). \tag{4}$$

In the SWG, modeling the multivariate density $\mathbb{P}(Y_1, Y_2, \ldots, Y_p)$ boils down to modeling a series of conditional (except the first one which is unconditional) univariate densities $\mathbb{P}(Y_1)$, $\mathbb{P}(Y_2|Y_1)$, $\ldots$, $\mathbb{P}(Y_p|Y_{p-1}, \ldots, Y_1)$. The sets of conditioning variables in the conditional univariate densities are included in the covariates $\mathbf{x}$ of the corresponding regression model, as described in sub-section 2.1. This ensures mutual consistency in the multivariate generator by incorporating inter-variable dependence.

To determine the order of the decomposition in (4) and to reduce the number of conditioning variables, we rely on the dependence graph from Fig. 1. It is strongly inspired from the proposal made in the HydEF project [15] to apply `Glimclim` in the UK. According to Fig. 1, atmospheric pressure is our independent variable Y_1 in (4), the wind speed depends on the atmospheric pressure, the air temperature depends on the wind speed and necessarily depends also on the atmospheric pressure but only indirectly, relative humidity depends on the wind speed and the air temperature, the global radiation depends on the air temperature and the precipitation depends on the global radiation.

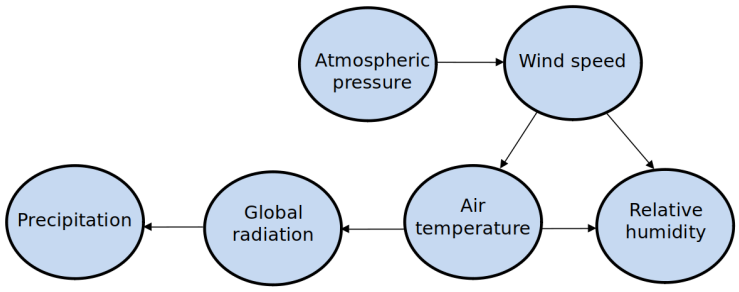


Figure 1: *Inter-variable dependency graph associated with the product rule in* (4) *to model the dependence structure of the hydro-meteolorogical variables in the SWG.*

Therefore, in our proposed SWG, regression models, selected among (1)-(3), are fitted for each hydro-meteorological variable independently by maximizing the log-likelihood (with the `glm` function in the base package of `R`). Preliminary univariate analyses are essential to ensure the selection of appropriate probability distribution and covariates that will result in a good fit of each variable. The simulation of the multi-variable surrogate series from the fitted SWG proceeds following the order dictated by the

dependence graph in Fig. 1 : atmospheric pressure is simulated first, driven by the covariates presented in sub-section 2.3, wind speed is then simulated including among the covariates the series simulated for atmospheric pressure, and so on and so forth.

2.3. Covariates

Each hydro-meteorological variable is modeled separately by a regression model, as described in sub-section 2.1, for which specific covariates are selected. In addition to the conditioning variables that serve to preserve inter-variable dependency as explained in sub-section 2.2, several other covariates are considered.

Important covariates are the large-scale variables provided by the ERA5 reanalysis [21]. As it relies on a consistent reprocessing of meteorological observations with data assimilation techniques, the ERA5 variable yields relevant meteorological information at an hourly time step on the 31 km $\times$ 31 km grid cell encompassing the gauged stations of interest. The stochastic generator aims at downscaling the information provided by the reanalysis, that is to simulate surrogate hydro-meteorological series at the spatial and temporal resolution of the gauged station. In addition to the large-scale variables, to account for temporal and spatial effects along with persistence in the surrogate series simulated by the stochastic generator, other covariates can be selected among the ones listed in Table 1.

Table 1: *Potential covariates to account for additional temporal and spatial variability and persistence.*

Annual cycle effects	$\cos\left(2\pi d/k\right), \sin\left(2\pi d/k\right)$ with $k \in (365, 183, 91, 30)$ and where $1 \leq d \leq 365$ is the day of the year
Diurnal cycle effects	$\cos\left(2\pi h/k\right), \sin\left(2\pi h/k\right)$ with $k \in (24, 12, 6)$ and where $1 \leq h \leq 24$ is the hour of the day
Spatial effects	x- and y-geographical coordinates
Memory effects	**Var.lag** : lagged values from the same Variable **SA.lag** : lagged Spatial Averages **MA.lag** : lagged Moving Averages at each station **SMA.lag** : lagged Spatial Moving Averages *with lag taking values from 1 to 8 time steps* *and MA with bandwidths from half a day to 3 days*

2.4. Covariate selection

As GLMs are relatively simple parametric models, the complexity of the variability of the hydro-meteorological variables can be reproduced thanks to an adequate choice of covariates. Indeed, since hydro-meteorological variables display high spatio-temporal variability in semi-arid climates, many additional covariates among the ones listed in Table 1 might be needed. A covariate selection procedure adapted for our study area, see Fig. 2, is thus necessary. To this end, we propose a *backward* procedure that starts with a large initial set of covariates containing a large number of cycles and memory effects from Table 1. This initial set includes four pairs of sine and cosine to account for annual cycle effects,

three pairs of sine and cosine to account for diurnal effects, and lags of up to eight time steps for each of the four types of memory effects, with a moving average bandwidths ranging from half a day to three days. Pruning this initial large covariate set is performed based on a two-step procedure.

First, LASSO regression [20] is applied to perform a preliminary screening. LASSO regression solves a regularized least squares problem which balances model complexity and model goodness-of-fit. It is recognized for its potential to perform simultaneously variable selection and parameter estimation [32]. Although it might only be an approximate regression model in non-gaussian cases, LASSO is straightforward to apply and it yields parsimonious solutions, i.e. with few covariates.

Second, a series of diagnostic tools are mobilized to identify the most significant covariates among the ones retained by LASSO. Conventional goodness-of-fit statistics such as the AIC (Akaike Information Criterion) [3] and BIC (Bayesian Information Criterion) [45] penalize models with too many covariates in order to avoid unnecessary and redundant information [43]. In practice, AIC and BIC indicate whether the removal of a given covariate is detrimental to the fit. Other diagnostic tools involve maximum likelihood ratio test [12], residual plots with various temporal frequencies (annual, monthly, daily) and covariate coefficient p-values. Nevertheless, the final verdict concerning covariate selection remains validation analyses (see Section 4 and 5), i.e. whether the stochastic generator with the selected covariates can reproduce satisfactorily the observation properties on data not used for calibration.

3. Water stress estimation in central Tunisia

3.1. Merguellil catchment

The study site called the Merguellil catchment lies in a semi-arid region located in central Tunisia, see Fig. 2. It is characterized by a relatively mountainous upstream area (1200 km^2) and by a downstream alluvial plain (676 km^2). The upstream area presents a hilly topography (altitude between 200 and 1200 m with a median elevation of 500 m) [34]. However, in the plain, the landscape is mainly flat, and the vegetation is typical of semi-arid regions: rainfed agriculture (olive tree and cereals) and summer vegetables (melons, peppers and tomatos). Farms in the downstream are composed mainly of small cultivated areas [39]. The upstream and downstream areas are separated by the El Haouareb dam (Fig. 2), which was built in 1989 to protect the village from inundations and to store irrigation water for the plain [10].

The study area is influenced both by the Mediterranean climate (dry subhumid) and the pre-Saharan's climate (arid). It is characterised by the inter-annual irregularity of precipitation, with an average of annual rainfall of about 300 mm per year, and by a high evaporative demand of about 1600 mm per year. There is no balance between water supply and water demand. Indeed, the demand for water is rising

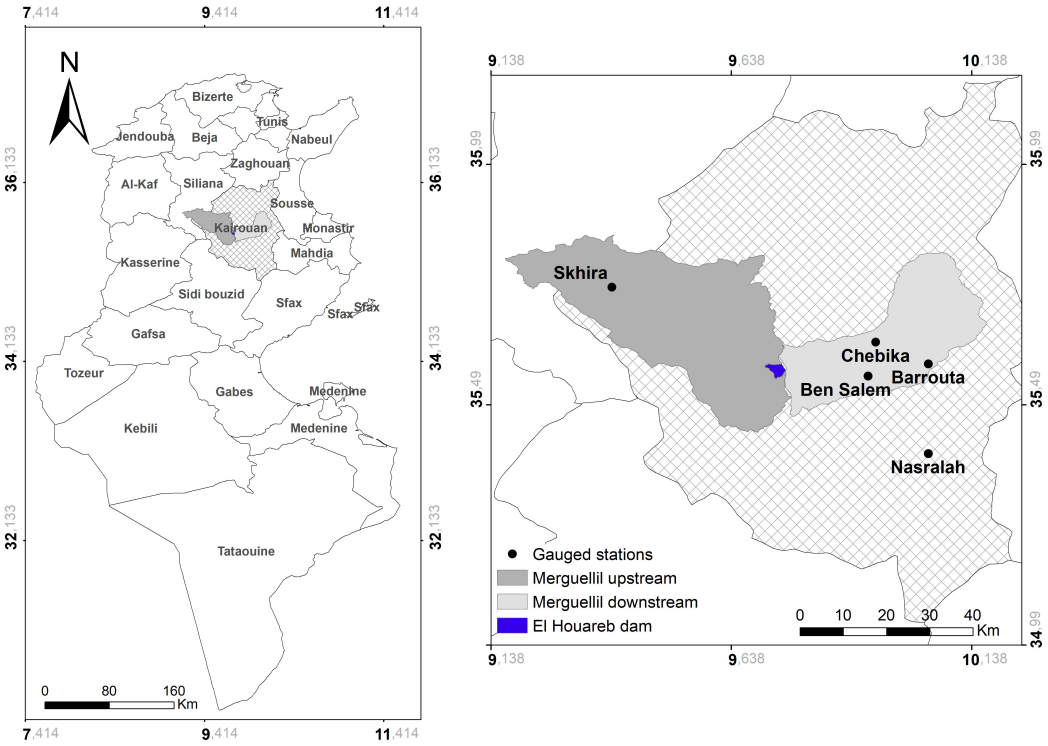


Figure 2: *Localisation of gauged stations : Merguellil catchment in central Tunisia.*

steadily, due to the increase in population and industrial development, but most importantly due to the intensification of agriculture, which is the main water consumer (around 80 %) [33].

The gauged network in the Merguellil catchment has five stations (Fig. 2). There is a station called Skhira located upstream in the mountainous area. Three stations, Ben Salem, Chebika and Barrouta, are very close to each other downstream of the El Houareb dam. The fifth station, Nasralah, is close by but lies outside the downstream plain. Six hydro-meteorological variables, namely atmospheric pressure, wind speed, air temperature, relative humidity, global radiation and precipitation, are measured and collected at these five stations at a half hourly time step since 2012 for the earliest ones.

We apply the sub-daily stochastic weather generator described in Section 2 on the observation series from the three gauged stations located in the Merguellil downstream plain, namely Ben Salem, Chebika and Barrouta (see Fig. 2). These stations present geographical and meteorological proximity thus sharing the same local climate. Our aim is to combine the information from the three gauged stations to obtain a single multi-variable series that is representative of the local climate of the Merguellil plain. Although the generator is calibrated on the three stations and can simulate series at each one of them, a single series is retained for the subsequent analysis, see Sections 4 and 5. The retained series corresponds to the Ben Salem station which is our reference station as it complies more closely with the prescribed standards of meteorological station according to WMO guidelines [41].

3.2. Dual source energy balance model

Dual Source energy balance algorithms provide separate estimates of the two main components of evapotranspiration, the evaporation and the transpiration, from remotely sensed data [31]. A separate estimate of transpiration is needed to assess plant water status and plant water use for sustainable management of crops (for drought monitoring or irrigation scheduling for instance). This is particularly the case for arid areas, which are characterized by sparse crop canopy, and for which the relative contribution of evaporation (E) and transpiration (T) can vary throughout time space [36]. In most cases, the soil surface is rather dry while the root zone holds a significant amount of moisture.

The challenge is to compute the energy partitioning between the soil and the plant rather than the whole agrosystem complex. In this case, the soil vegetation system is considered as a two layer model [40]. The crucial elements of these two-component models are the radiometric temperature and aerodynamic temperature [40]. Indeed, the contribution of canopy and soil to fluxes depend on differences in temperatures between each component as well as with the atmosphere above the canopy [40]. Besides, an estimate of the fractional vegetation cover and of the view angle, is needed to derive the link between the radiometric temperature, e.g. available from satellite platforms, and the source temperatures of the soil and the vegetation (Ts and Tv, respectively).

In this work, we use the dual-source model Soil Plant Atmosphere and Remote Evapotranspiration (SPARSE) [11] which is based on the same rationale as TSEB (Two-Source Energy Balance model) [40]. SPARSE derives, from the remotely sensed surface temperature Tsurf, separate estimates of the instantaneous fluxes of the soil (subscript s) and vegetation (subscript v) components of the total fluxes of the energy budget at the satellite overpass time: net radiation (RN), soil (G), sensible (H) and latent heat (LE) expressed in $\mathbb{W}/\mathbb{m}^2$.

The SPARSE model can be run under two modes: a retrieval mode to simulate actual evaporation and transpiration from Tsurf, and a prescribed mode which simulates evaporation and transpiration rates for known stress levels (for instance, the two extremes of the water status spectrum: fully stressed or maximum moisture i.e. potential conditions). The prescribed mode provides an estimate of the potential latent flux for the soil and the vegetation (LEspot and LEvpot respectively). In retrieval conditions, the respective stress levels (between non evaporating/transpiring and potential rates) correspond to two unknown which are solved from the single piece of information (Tsurf) with the following simplification to solve the underdetermined problem : first, the vegetation is assumed to be unstressed. Tsurf is used to estimate LEs. If the vegetation is suffering from water stress, the resulting LEs will decrease to unrealistic levels (negative values). In that case, we assume that the soil surface is stressed and LEs is set to a minimum value close to zero. Then the energy budget equation is solved for LEv. If LEv is also negative, fully stressed conditions is imposed for both soil and surface components [44].

Water stress index (SI) is then defined from the actual and potential evapotranspiration rates simulated from retrieval and prescribed mode respectively at the time of the satellite overpass. This index can be defined to describe the water status for only the soil or vegetation rates (using LEs or LEv) or even for the soil-vegetation composite using the sum, as follows :

$$SI = 1 - \frac{LEv + LEs}{LEvpot + LEspot}. \tag{5}$$

If the actual evapotranspiration value is close to the potential value, the stress index takes low values close to zero that reflect unstressed conditions. However, if the actual evapotranspiration is low comparing to its potential value, the stress index values may reach 1, which represents fully stressed conditions. On the other hand, daily evapotranspiration is derived from an extrapolation algorithm in order to reconstruct its sub-daily variations by assuming the self preservation of the evaporative fraction [19]. Variables are not computed during days with no RS observations (cloudy days).

3.3. Surrogate series' evaluation procedure

The stochastic weather generator (SWG), described in Section 2, can be used either in *a gap filling mode* in which missing values during the observation period are imputed, i.e. plausible values are simulated by the generator to replace them or in *a projection mode* in which the generator simulates values on a period with no observations to extend the observations temporally. In the gap filling mode, the generator simulates surrogate series on short periods and is strongly constrained by the observations while in the projection mode, the generator runs freely for a long period.

For the validation in the gap filling mode, we perform a random removal of observations in order to introduce artificially missing values (see Fig 3). The SWG is calibrated on the non-missing values and gap-filled values are compared with observations that were set aside. Around 15 % of the observations are removed over a common period for the three gauged stations lying in the Merguellil downstream plain (see Fig. 2) denoted as S1, S2 and S3 in Fig 3. In the projection mode, the evaluation procedure is based on a cross-validation setup. The observation period covers five years (2012-2016). Each year is kept aside in turn for validation while the four remaining years serve to calibrate the SWG. This allows to produce hydro-meteorological surrogate series over the complete observation period and to assess the SWG's performance when it runs freely over a year.

On one hand, performance criteria pertain to the hydro-meteorological surrogate series themselves in order to evaluate how well they reproduce the observations. On the other hand, they are related to the variables (ET and SI) simulated by SPARSE model mentioned in sub-section 3.2, when constrained either by the observed series or by the surrogate series.

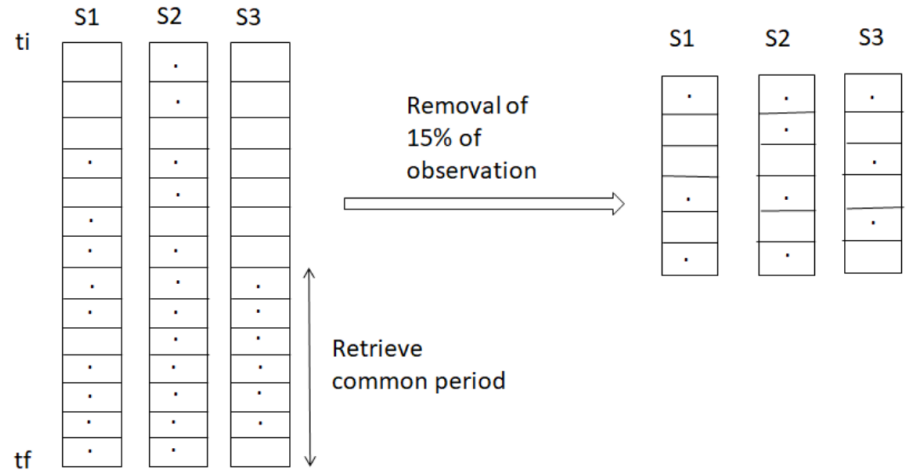


Figure 3: *Validation scheme of the gap filling mode : selection of a common observation period for the three stations in the Merguellil plain (denoted as S1, S2 and S3) and random removal of observations over that period. The evaluation is performed on the removed observations.*

4. Multi-variable sub-daily SWG in the Merguellil

This section is devoted to the evaluation of the feasibility of the application of the sub-daily stochastic weather generator (SWG) described in Section 2 to the three gauged stations situated in the Merguellil downstream plain, see Fig. 2. Model selection, see sub-section 4.1, consists in the selection of the appropriate probability distribution for each hydro-meteorological variable, see (1)-(3), the choice of covariates deduced from ERA5 reanalyses and the selection of other spatio-temporal effects in the covariates, see Table 1. We had to revise many times model selection until validation results, see sub-sections 4.2 and 4.3, were deemed satisfactory. A single gap filled series, corresponding to the Ben Salem station, is produced and retained to apply the bias correction methods in Section 5.

4.1. Model selection

The selection of the probability distributions presented in Table 2 was performed by visual inspection of the histogram of the hydro-meteorological variable considered. The Gaussian distribution, in three instances combined with preliminary transformations as indicated in Table 2, is used for all the variables except for precipitation occurrence and intensity. In contrast to precipitation occurrence which is simulated stochastically, the global radiation occurrence is deterministic. Indeed, the sunrise and the sunset are determined based on the coordinates of the station and the day of the year (see `R` package `insol`). Nocturnal time steps, where global radiation is set to zero, are thus identified.

Covariates deduced from ERA5 reanalysis are selected for each hydro-meterological variable as follows, see Table 3. For all variables but the atmospheric pressure for which mean sea level pressure is used, the large-scale version of the hydro-meteorological variable is taken. In most cases, this corresponds to raw reanalysis products, as indicated in Table 3. In contrast, the large-scale covariate for the wind speed

Table 2: *Preliminary transformations and selected probability distributions.*

Hydro-meteo. variable	Transformation	Probability distr.
Air Pressure **Pr**	✗	Gaussian, see (1)
Wind speed **WS**	$\ln(\exp(WS) - 1)$	
Air temperature **AirT**	✗	
Relative humidity **Rh**	$\tan(\pi(Rh - 0.5))$	
Global radiation intensity **GR**	$\ln[\max(\ln(GR)) - \ln(GR)]$	
Precipitation occurrence **Precip occ**	✗	Bernoulli , see (3)
Precipitation intensity **Precip int**	✗	Gamma, see (2)

(WS) was derived by applying a drawdown of the 10 m vertical and horizontal wind components from ERA5 products to 2 m [5], to match the gauged station's measurements, and by taking the Euclidean norm of the two wind components. The large-scale covariate for the relative humidity was also derived based on 2 m temperature and 2 m dewpoint temperature ERA5 products, according to the procedures defined in Allen *et al.* [5]. For a single hydro-meteorological variable, namely global radiation, we selected a second large-scale covariate, total cloud cover, which was found useful to improve the fit. This large scale variable is related to the presence of clouds that reduces the proportion of global radiation.

To match the half-hourly temporal resolution of the hydro-meteorological observations, the large-scale covariates, that are available at the hourly time-step (i.e. the temporal resolution of the ERA5 products) have each hourly values sampled twice.

Table 3: *Large-scale covariates deduced from ERA5 reanalyses for each hydro-meteorological variable (see Table 2 for the abbreviations).*

Hydro-meteo. variable	Large-scale covariate
Pr	mean sea level pressure (raw)
WS	2 m wind speed (derived)
AirT	2 m temperature (raw)
Rh	relative humidity (derived)
GR	surface solar radiation downwards (raw)
	total cloud cover (raw)
Precip	total precipitation (raw)

Based on the covariate selection procedure outlined in sub-section 2.4, in addition to the large-scale covariates in Table 3 and the variables to account for inter-variable dependencies (see sub-section 2.2), the final set of covariates selected among the potential ones listed in Table 1 are summarized in Table 4. The selected oscillation periods for the sine and cosine dedicated to annual and diurnal cycle effects are indicated in days and in hours respectively. Regarding the memory effects, the length of the longest lag is specified. For example, for the air pressure, we considered a spatial average (SA) lagged of one time step and the values of the variable (Var) lagged up to three time steps. In two instances, additional covariates were introduced to improve the fit of the SWG. These additional covariates are : *wind breaker*, a binary covariate pointing out the presence of a wind breaker at the station Chebika, *seasons*, a binary covariate

indicating the rainy season and *stations*, a three category covariate related to the three gauged stations.

4.2. Annual and diurnal cycle validation

Annual and diurnal cycles are the primary features the stochastic weather generator (SWG) should be able to reproduce in order to be useful for water stress estimation. These are computed by averaging values of each time step (half hour) in the day for diurnal cycles and in each month for annual cycles over the study period. We compare observations' cycles with cycles computed from the un-processed large-scale variables (see Table 3) and from the series generated by the SWG in projection mode, i.e. when surrogate series are obtained by a cross-validation procedure for the whole observation period (2012-2016) by leaving aside each year in turn for validation. Further evaluation in projection mode is deferred to Section 5 in which the SWG is compared with two state-of-the-art bias correction methods. The latter methods being applied to anomalies of the diurnal cycles computed for three seasons (see Section 5), cycles are well reproduced by construction.

In Fig. 4, we see that the cycles computed from large-scale variables (see Table 3), accurately reproduce observations' cycles for some hydro-meteorological variables. This is the case for air temperature, relative humidity and global radiation. However, the cycles computed from the large-scale variables for atmospheric pressure and the precipitation are over-estimated. In contrast, the wind speed cycles are under-estimated by the large-scale variable : in the diurnal cycle, the peak in the afternoon is not high enough and the annual cycle fails to reach observed high values. Despite these initial deviations in the large-scale variables behavior, the SWG is able to correct these and to reproduce correctly annual and diurnal cycles for most variables (precipitation being more challenging).

4.3. Gap filling mode validation

The multi-variable sub-daily SWG is then evaluated in gap filling mode. To this end, around 15% of the observations are removed at random over a common period, see sub-section 3.3. In Fig. 5, quantile-quantile plots for these 15% observations randomly removed and gap-filled serve to evaluate how well imputed values match the observations in terms of distribution. Results are presented for the Ben Salem station, our reference station as explained in sub-section 3.1 for which air pressure is not available hence there is no quantile-quantile plot for this hydro-meteorological variable. The other two stations displayed similar results (not shown). In addition, inter-variable dependencies, as measured by Kendall's rank coefficients [28], were well preserved (results not shown). Overall, the gap filling mode was found to perform well. The SWG was run again in gap filling mode on the original observation series in order to produce continuous series without missing values for all stations over the observation period.

Table 4: *Selected covariates for each hydro-meteorological variable (see Table 2 for the abbreviations) from the potential covariates in Table 1. The selected oscillation periods, k, is specified in days (in hours) for the annual (diurnal) cycles, for the memory effect covariates, the length of the longest lag (in time steps) is given and the "others" column contains additional covariates that were found useful.*

Hydro-meteo. variable	Geographical coordinates	Annual cycles	Diurnal cycles	Memory effects	others
Pr	✓	✗	✗	SA.lag1, Var.lag3	✗
Ws	✓	$k \in (365)$	$k \in (24, 12, 6)$	SA.lag4, MA.lag8, SMA.lag1, Var.lag4	wind breaker
AirT	✓	$k \in (365, 183)$	$k \in (24, 12, 6)$	SA.lag3, MA.lag8, SMA.lag8, Var.lag3	✗
Rh	✓	✗	✗	SA.lag3, MA.lag3, SMA.lag7, Var.lag1	✗
Gr	✓	$k \in (365, 183)$	$k \in (24, 12)$	✗	✗
Precip occ	✗	$k \in (365, 183, 91, 30)$	$k \in (24, 12, 6)$	SA.lag4, Var.lag1	✗
Precip int	✗	✗	✗	SA.lag1, Var.lag1	seasons, stations

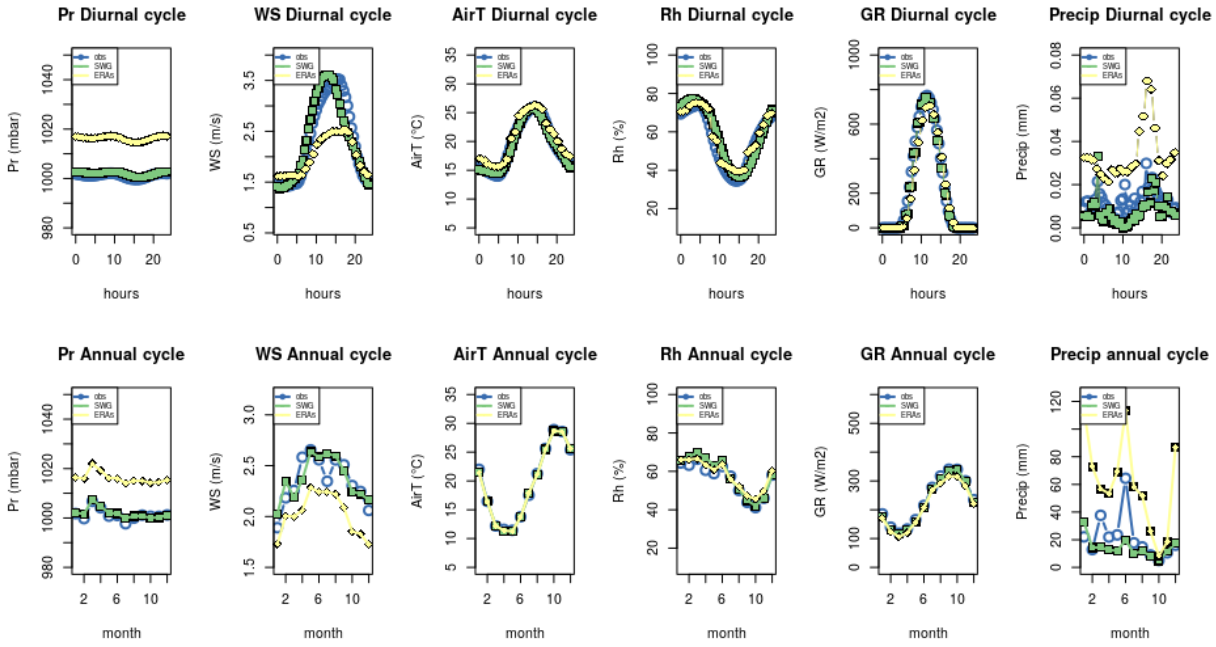


Figure 4: *Annual and diurnal cycles for all hydro-meteorological variables (see Table 2) at the Ben Salem gauged station in the Merguellil downstream plain (see Fig. 2).*

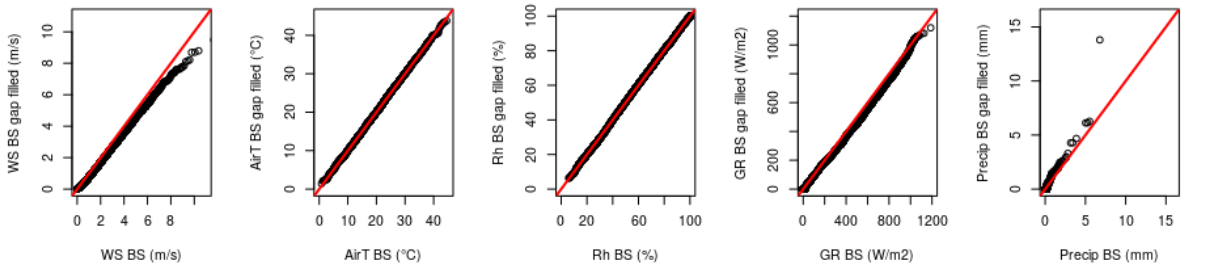


Figure 5: *Quantile-quantile plots for each hydro-meteorological variables (see Table 2) simulated in gap filling mode (15% values removed at random) for the Ben Salem station, see Fig. 2.*

5. Evaluation and comparison in projection mode

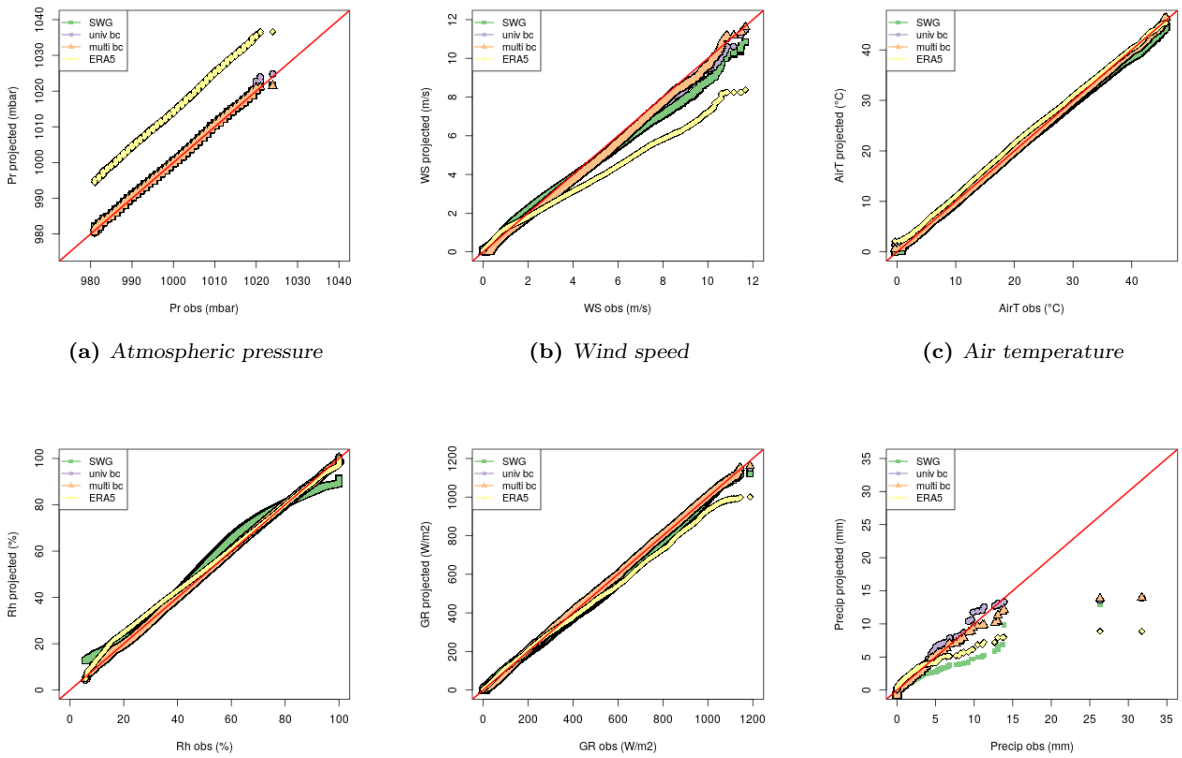
The multi-variable sub-daily SWG proposed in Section 2 is further evaluated and compared, in projection mode, with two state-of-the-art bias correction methods, namely a univariate approach called CDFt [38] and a multivariate approach called MBCn [13]. Bias correction methods seek to correct systematic differences in distributional properties (mean, variance or quantile) between a hydro-meteorological variable measured at a station (local-scale) and its large-scale counterpart. These methods, like the SWG, do not attempt to reproduce the chronological order of the local observation series. The local-scale meteorological observations are taken from the gap filled Ben Salem series produced by the SWG, see Section 4, and the large-scale variables are the same as listed in Table 3.

To remove systematic fluctuations from the hydro-meteorological variables, which are assumed to be constant over the period considered, the bias correction methods are applied on anomalies, defined as deviations from the diurnal cycles, of the large- and the local-scale variables. Working with anomalies allows to focus on random fluctuations around the diurnal cycles. These are computed for three seasons : summer (June to August), winter (November to March) and mid-season (the remaining months). To match the temporal resolution of the reanalysis product, the anomalies from the hydro-meteorological observation series are sub-sampled at the hourly time step. Conversely, to match the temporal resolution of the original observation series, the corrected large-scale anomalies are interpolated linearly to the half-hourly time step.

5.1. Evaluation in terms of hydro-meteorological variables

In what follows, *surrogate series* include the hydro-meteorological series generated in projection mode (i.e. in the cross-validation setup, leaving aside one year each time) either by the SWG or the two bias correction methods. Surrogate series also designate the un-processed large-scale variables obtained from the reanalysis products, i.e. without downscaling (see Table 3). First, we evaluate whether the distribution of each hydro-meteorological variable is well reproduced. To this end, we rely on quantile-quantile plots (Fig. 6). The large-scale variables deduced from the ERA5 reanalysis reproduce well the distributions of air temperature (Fig. 6c) and relative humidity (Fig. 6d) as can be seen from the good alignment along the first bisector, the red line in the quantile-quantile plots. The series generated from the bias correction methods are also accurate for these two hydro-meteorological variables however the series simulated by the SWG are slightly distorted for the relative humidity, see Fig. 6d. In contrast, some under-estimation occurs (quantile-quantile plots under the first bisector) for the higher values of global radiation (Fig. 6e), wind speed (Fig. 6b) and precipitation (Fig. 6f) for the large-scale variables. Bias correction methods are able to correct this under-estimation. The SWG also overcomes this initial bias of the large-scale variable but not always as well, see Fig. 6b. In Fig. 6a, we observe that the mean

sea level pressure variable from the ERA5 product has a systematic positive bias (atmospheric pressure is on average higher at sea level than at the stations' level). The SWG along with the two bias correction methods are able to correct this positive bias.



(a) *Atmospheric pressure* **(b)** *Wind speed* **(c)** *Air temperature*

(d) *Relative humidity* **(e)** *Global radiation* **(f)** *Precipitation*

Figure 6: *Quantile-quantile plots for each hydro-meteorological variable to compare surrogate series (un-processed large-scale variables, SWG series and the two bias corrected series) on the y-axis with the observations on the x-axis.*

Second, we assess whether the inter-variable dependencies are well reproduced. This is achieved in Fig. 7 by comparing Kendall's rank correlation coefficients from the observed series with those of the surrogate series (SWG and bias corrected series and un-processed large-scale variables listed in Table 3). Positive values indicate that both variables tend to increase or decrease simultaneously while negative values indicate that they tend to vary in an opposite manner. A value near zero signals a lack of dependence. In Fig. 7, there is an overall good alignment along the first bisector (the red line) showing that the inter-variable dependence strength is relatively well preserved in all cases. Nevertheless, compared with the two bias correction methods and the large-scale variables, the inter-variable dependence strength in the series produced by the SWG are more tightly aligned with the first bisector.

5.2. Evaluation in terms of water stress estimation

In this section, the surrogate series are compared in terms of their ability to constrain the SPARSE model (see sub-section 3.2) in order to retrieve daily evapotranspiration (ETd) and water stress index

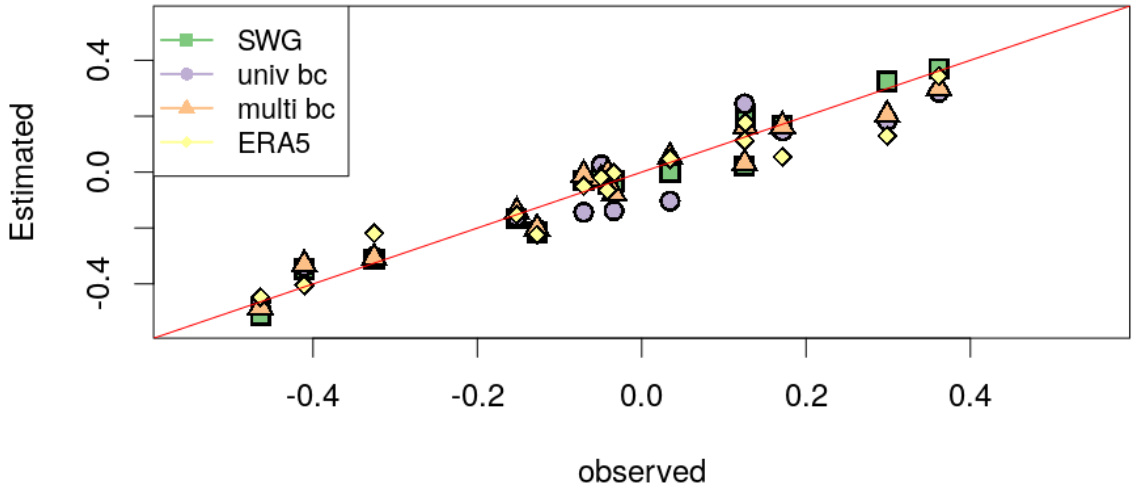
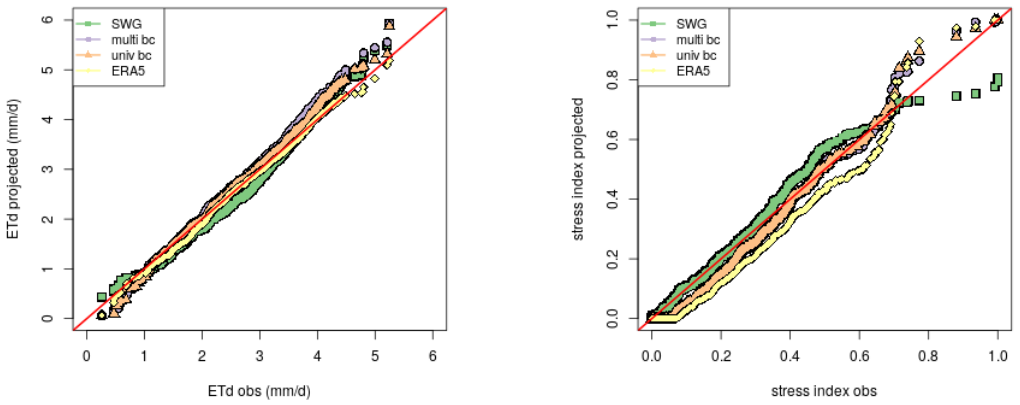


Figure 7: *Kendall's rank coefficients for each pair of hydro-meteorological observed series on the x-axis and estimated on the y-axis by the surrogate series (un-processed large-scale variables, SWG series and the two bias corrected series).*

(SI) as similar as possible as when hydro-meteorological observations are used to constrain the model. In Fig. 8, the comparison is first carried out in terms of distribution with quantile-quantile plots. On the x-axis, the daily evapotranspiration (Fig. 8a) and the water stress index (Fig. 8b) is simulated by the SPARSE model when constrained by the observations. On the y-axis of both panels in Fig. 8, the simulation is constrained by the surrogate series, i.e. either the un-processed large-scale variable (ERA5), see Table 3, or one of the series in projection mode (SWG, univariate and multivariate bias correction methods). Regarding daily evapotranspiration in Fig. 8a, the surrogate series yielded comparable results. In contrast, for the water stress index in Fig. 8b, the highest values are not well reproduced by the different surrogate series. This is especially striking for SI values that are under-estimated (quantile-quantile plots under the first bisector) when the SPARSE model is constrained by the un-processed large-scale variables (ERA5). However, low extreme values are better reproduced when contrained with the SWG.

In order to translate differences in distribution as visualized by discrepancies from the first bisector in the quantile-quantile plot from Fig. 8b into a more hydrologically interpretable analysis, we propose a further comparison based on the probability that the water stress index exceeds a given threshold, so-called *exceedance probability*. In Fig. 9, threshold values are represented on the x-axis, ranging from 0 to 1. The exceedance probabilities are on the y-axis, in black, as estimated when constraining the SPARSE model by hydro-meteorological observations, with an empirical 95% of confidence band in gray. The exceedance probabilities estimated when constraining the SPARSE model with a surrogate series, either the un-processed large-scale variables (ERA5) from Table 3 or one of the downscaled series (SWG,



(a) *Daily evapotranspiration* **(b)** *Water stress index*

Figure 8: *Quantile-quantile plots of the simulations from the dual source energy balance model, see sub-section 3.2, when constrained by observed series on the x-axis compared to surrogate series on the y-axis, i.e. either the unprocessed large-scale variables (ERA5) or the downscaled series (SWG, univariate and multivariate bias correction methods in projection mode).*

univariate or multivariate bias correction methods), are as indicated in the color legend in Fig. 9. When the SPARSE model is constrained by the large-scale variables (ERA5) and by the series generated by the bias correction methods, the estimated exceedance probability falls outside the confidence interval for low threshold values between 0 and 0.3 for univariate and multivariate bias correction and between 0 and 0.5 for large scale variables. In contrast, when constrained with the projected series from the SWG, the estimated exceedance probability is almost always within the confidence interval.

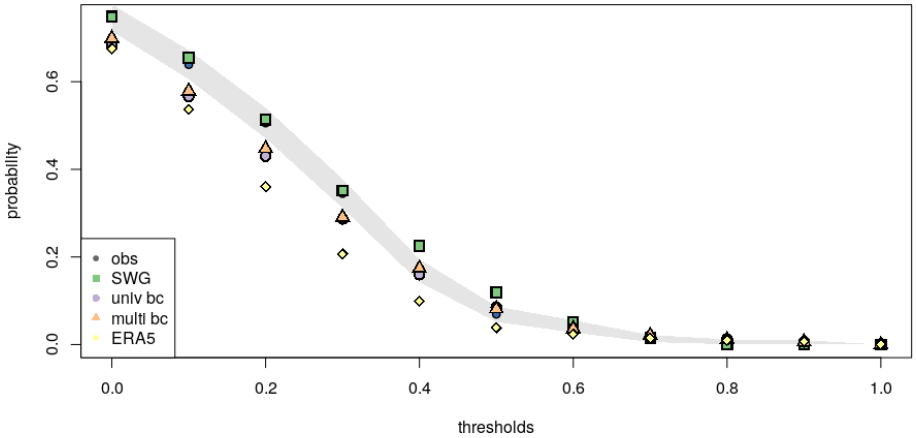
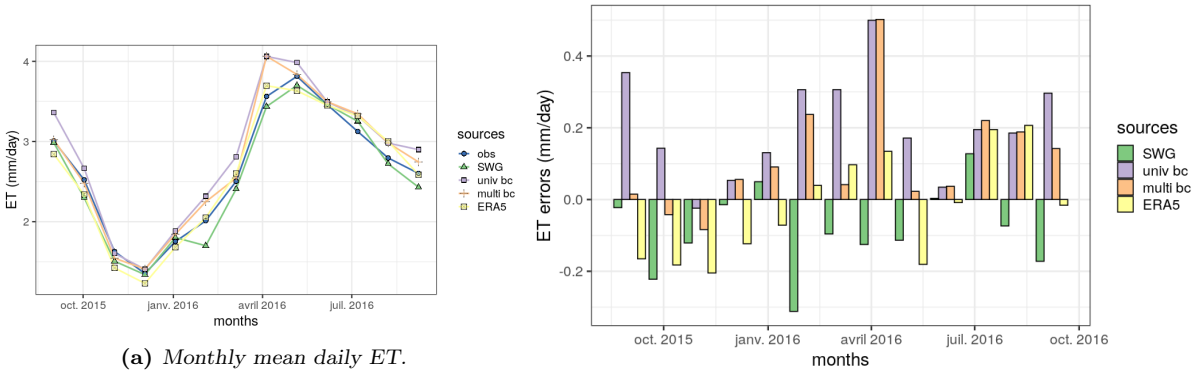


Figure 9: *Estimated exceedance probabilities for increasing threshold values for the water stress index as simulated by the energy balance model when constrained with the observed series (in black) along with a 95% empirical confidence band (in gray) and when constrained with the surrogate series (see the color legend). The SWG and the univariate and multivariate bias correction methods were applied in projection mode.*

In Fig. 10, a comparative analysis is carried out at the monthly scale. In Fig. 10a, monthly evapotranspiration average for the agricultural year 2015-2016 is shown when the energy balance model is constrained either by hydro-meteorological observations (in blue) or by one of the surrogate series (see the color legend). We observe that monthly ET simulations are very similar whatever series is used to

constrained the energy balance model. The same observation holds for all the years studied (results not shown). Further analyses is gained by looking into monthly ET errors in Fig. 10b that are computed as the differences between monthly ET average constrained by meteorological observed series and monthly ET average constrained by a surrogate series. Positive (negative) errors indicate over (under) -estimation of the observed monthly ET average while values near zero mean a good reconstruction of the observed monthly ET average. Errors shown in Fig. 10b vary strongly throughout the year. In summer, errors derived from the simulations constrained by series provided by SWG are rather low ($<$ 0.2 mm/day). Errors derived from the two bias correction methods are larger, noticeably in april, where we observe errors exceeding 0.5 mm/day. Relying on the SWG simulated series to constrain the SPARSE model could be very helpful to detect minimal stress events.



(a) *Monthly mean daily ET.*

(b) *Monthly mean daily ET errors.*

Figure 10: *Monthly average daily evapotranspiration of available data over the 2015-2016 period when the energy balance model is constrained with either meteorological observations or one of the surrogate series.*

A final analysis is carried out at the monthly scale. Fig 11 presents the temporal variation of the stress index simulated from different surrogate series for data available in August 2016. We observe that the stress index derived from the series produced by the bias correction methods and the un-processed large scale variables tend to be underestimated. In contrast, SI derived from the SWG is rather close to the SI obtained when simulated by the hydro-meteorological observations.

6. Summary and outlook

In semi-arid areas, actual water use deduced from evapotranspiration and water stress studies are very useful to gain greater understanding into the mechanisms leading to droughts. Dual source energy balance models such as the SPARSE model [11] can serve to retrieve estimates of evapotranspiration and water stress indices. These models rely on satellite information and hydro-meteorological observation series that contain almost always missing data for various reasons. In addition, the observation period might be too short for the purposes of drought studies. We propose to rely on a multi-variable stochastic weather

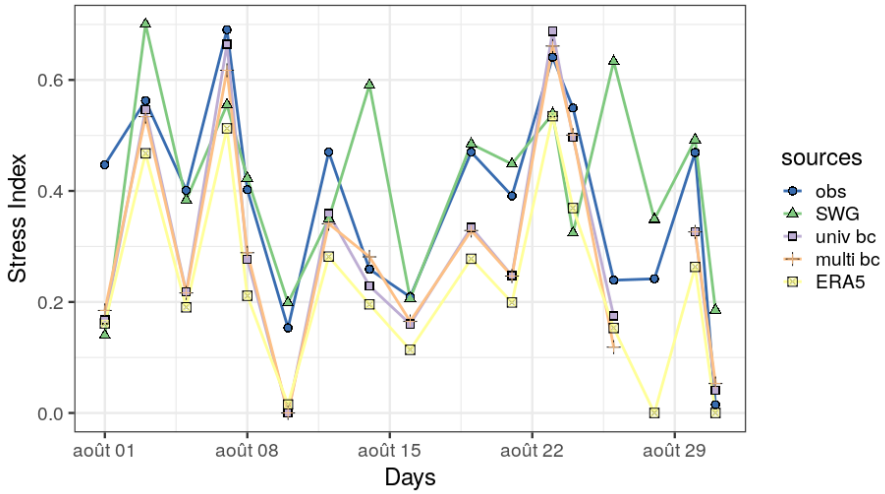


Figure 11: *Temporal variation of water stress index (August 2016) simulated when constraining the energy balance model either with hydro-meteorological observations or with one of the surrogate series (see color legend).*

generator (SWG) to fill in missing data and to provide long time series when necessary. The SWG can simulate hydro-meteorological series over any time interval based on previous local hydro-meteorological observations and large-scale variables (reanalysis).

As far as we know, not many multi-variable SWG have been proposed in the literature. We chose to adapt the proposal of Chandler [16] as it relies on simple statistical models and mechanisms. The adaptation of the SWG required to develop our own `R` code, which will be made available soon as a package `MetGen`, in order to include modifications accounting for the presence of a diurnal cycle. Although the generalized linear models used to model each hydro-meteorological variable are implemented in the base package of `R`, a large amount of time was dedicated to perform model selection and validation. Model selection consists in the identification of an adequate dependence graph to account for inter-variable dependence, the choice of the probability distribution potentially combined with a transformation for each hydro-meteorological variable, the choice of the large-scale variables deduced from ERA5 data and the development of a selection procedure to include additional covariates accounting for spatial, temporal and memory effects. In particular, for the large-scale covariate choice for atmospheric pressure at the stations, we retained mean sea level pressure despite the difference in altitude. This is justified by the fact that all statistical downscaling methods - SWG and bias correction - are able to make the appropriate corrections (see Fig. 6 and Fig. 4). Moreover, the selection of a second large-scale covariate, total cloud cover, in addition to the large-scale global radiation, was decisive to improve the fit for the local-scale global radiation. Model validation was performed both in gap filling and projection mode. In both cases, the simulation is carried out for each hydro-meteorological variable following the order prescribed by the dependence graph, see Fig. 1. Simulation has to proceed one time step at a time in order to update the memory effects such as lagged moving averages.

The proposed SWG is compared with two state-of-the-art bias correction methods, CDF-t and MBCn, that are alternatives to statistically downscale large-scale information provided by the reanalysis variables. These bias correction methods are applied to anomalies over the diurnal cycles in order to remove systematic components of variability which are thus assumed to be constant over the study period. As, in our application, the study period has five years, this assumption is reasonable. In contrast to the SWG, the bias correction methods are not stochastic and therefore yield a single surrogate value at each time step. The comparison was carried out in projection mode, i.e. when the downscaling approaches, SWG and bias correction methods, are used to provide temporal extension of the existing observation series. Projection mode is more challenging than gap filling mode since the surrogate series are generated over a long period instead of just filling gaps in the observations series. A cross-validation procedure which sets one year aside in turn for validation implements the projection mode.

We consider a wide range of performance criteria to carry out the comparison between the observed series and surrogate series, un-processed large-scale variables (without downscaling) or provided by the SWG, the univariate or the multivariate bias correction methods. The first set of criteria gathers conventional criteria that assess by direct comparison how accurately the surrogate series reproduce the observed series. The second set of criteria is indirect in the sense that it focuses on the estimated evapotranspiration and water stress index provided by the SPARSE model when it is constrained either by the observed series or the surrogate series. Our main findings are as follows, concerning each set of criteria.

Hydro-meteorological variables are, in general, very well reproduced by the downscaled series (provided by the SWG, the univariate or multivariate bias correction methods), see Fig. 6, except for precipitation whose higher values are under-estimated by all downscaling approaches, see Fig. 6f. For the two bias correction methods, this may be caused by the temporal interpolation step that is performed after bias correction. Indeed, as ERA5 reanalysis data are available at hourly time intervals, corrected values are linearly interpolated to half-hourly time steps. Although this temporal disagregation step is sensible in most cases, it might be that a more refined strategy would be needed for high precipitation values. Concerning the SWG, the underestimation of high precipitation values in Fig. 6f is likely caused by the choice of the Gamma distribution, see (2). Indeed, the Gamma distribution is light-tailed and is prone to under-estimate extreme events. A more flexible distribution, such as the one proposed in Carreau & Vrac [14], could be used at the expense of a much more complex statistical model. Inter-variable dependence strength, as measured by Kendall's rank coefficients, is well reproduced by all surrogate series, see Fig. 7, even by the one produced by the univariate bias correction method which doesn't have any explicit mechanism to account for inter-variable dependencies. This might be explained by the fact that we worked on anomalies and the diurnal cycles of the observations are preserved by construction.

Regarding the second set of criteria, evapotranspiration and water stress index are globally more

similar in distribution to the observations when the SPARSE model is constrained by the series simulated from the SWG, see Fig. 8. This similarity in distribution translates into a more accurate estimation of the exceedance probability of the stress index, see Fig. 9. Lastly, monthly mean analyses of the estimated daily evapotranspiration show that the surrogate series provided by the SWG leads, overall, to lower errors (Fig. 10). As a result, the probability to miss drought events is lessen. This fact is emphasized in Fig. 11 where we see that the SWG better reproduce SI obtained when constrained by the observations. These analyses give confidence that the combination of the SWG and the SPARSE model yields a reliable tool to perform realistic water stress estimation and detection when hydro-meteorological information is lacking. Note that this tool relies on open source data (reanalysis and satellite data), that the SPARSE model is available on-line (http://tully.ups-tlse.fr/gilles.boulet/sparse) and that the SWG will be made available soon as an `R` package `MetGen` making the tool applicable, in principle, on any study area.

The main objective of this work is the adaptation and evaluation of a multi-variable sub-daily SWG geared towards an hydrological application. Detailed contributions are (1) the adaptation of the SWG to the sub-daily resolution, (2) its application in a semi-arid climate, (3) a comparison with other types of downscaling methods conventionally used with global climate model simulations and (4) the use of the surrogate series generated from all proposed downscaling methods to constrain the energy balance model in order to simulate evapotranspiration and water stress over a long period. Future work will be dedicated to develop a spatial extension of the SWG at high resolution in order to estimate water stress over the study area at fine spatial and temporal resolutions. Second, daily interpolation of impact variables will be performed for cloudy days. Lastly, the SWG-SPARSE tool will be tested for similar applications but in different climatic conditions (e.g. coastal area).

[1] Ailliot, P., Allard, D., Monbet, V., & Naveau, P. 2015. Stochastic weather generators: an overview of weather type models. *J. de la Socit Franaise de Statistique*, **156**(1), 101–113.

[2] Aissaoui Fqayeh, I., El Adlouni, S. E., Ouarda, T. B. M. J., & St-Hilaire, A. 2006. *Développement de l'estimateur GLM-ML pour le modèle log-normal non stationnaire et application à des précipitations extrêmes.* INRS, Centre Eau, Terre et Environnement.

[3] Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. *Pages 267–281 of:* B.N., Petrov, & F., Cski (eds), *Proc. 2nd Int. Symp. Inference Theory.*

[4] Allen, R. G., Pereira, L. S, Raes, D., Smith, M., *et al.* . 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome*, **300**(9), D05109.

[5] Allen, R. G., Pereira, L. S., Smith, M., Raes, D., & Wright, J.L. 2005. FAO-56 dual crop coefficient method for estimating evaporation from soil and application extensions. *Journal of irrigation and drainage engineering*, **131**(1), 2–13.

[6] Ambrosino, C., Chandler, R. E., & Todd, M. C. 2011. Southern African Monthly Rainfall Variability: An Analysis Based on Generalized Linear Models. *Journal of Climate*, **24**(17), 4600–4617.

[7] Amri, R., Zribi, M., Lili-Chabaane, Z., Duchemin, B., Gruhier, C., & Chehbouni, A. 2011. Analysis of vegetation behavior in a North African semi-arid region, using SPOT-VEGETATION NDVI data. *Remote Sensing*, **3**(12), 2568–2590.

[8] Ayar, P. V., Vrac, M., Bastin, S., Carreau, J., Déqué, M., & Gallardo, C. 2016. Intercomparison of statistical and dynamical downscaling models under the EURO-and MED-CORDEX initiative framework: present climate evaluations. *Climate Dynamics*, **46**(3-4), 1301–1329.

[9] Baccour, H., Feki, H., Slimani, M., & Cudennec, C. 2012. Interpolation de lévapotranspiration de référence en Tunisie par la méthode de krigeage ordinaire. *Science et changements planétaires/Sécheresse*, **23**(2), 121–132.

[10] Ben Ammar, S., Zouari, K., Leduc, C., & M'barek, J. 2006. Caractérisation isotopique de la relation barrage-nappe dans le bassin du Merguellil (Plaine de Kairouan, Tunisie centrale). *Hydrological sciences journal*, **51**(2), 272–284.

[11] Boulet, G., Mougenot, B., Lhomme, JP., Fanise, P., Lili-Chabaane, Z., Olioso, A., Bahir, M., Rivalland, V., Jarlan, L., Merlin, O., *et al.* . 2015. The SPARSE model for the prediction of water stress and evapotranspiration components from thermal infra-red data and its evaluation over irrigated and rainfed wheat. *Hydrology and Earth System Sciences Discussions*, 4653–4672.

[12] Buse, A. 1982. The Likelihood Ratio, Wald, and Lagrange Multiplier Tests: An Expository Note. *The American Statistician*, **36**(3a), 153–157.

[13] 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.

[14] Carreau, J., & Vrac, M. 2011. Stochastic downscaling of precipitation with neural network conditional mixture models. *Water Resources Research*, **47**(10).

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

[16] Chandler, R. E. 2005. On the use of generalized linear models for interpreting climate variability. *Environmetrics*, **16**(7), 699–715.

[17] Chirouze, J., Boulet, G., Jarlan, L., Fieuzal, R., Rodriguez, JC., Ezzahar, J., Raki, S. Er., Bigeard, G., Merlin, O., Garatuza-Payan, J., *et al.* . 2014. Intercomparison of four remote-sensing-based energy balance methods to retrieve surface evapotranspiration and water stress of irrigated fields in semi-arid climate. *Hydrology and Earth System Sciences Discussions*, 1165–1188.

[18] 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.

[19] Delogu, E., Boulet, G., Olioso, A., Coudert, B., Chirouze, J., Ceschia, E., Le Dantec, V., Marloie, O., Chehbouni, G., & Lagouarde, J. P. 2012. Reconstruction of temporal variations of evapotranspiration using instantaneous estimates at the time of satellite overpass. *Hydrology and earth system sciences*, **16**, 2995–3010.

[20] Friedman, J., Hastie, T., & Tibshirani, R. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, **33**, 1–22.

[21] Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., *et al.* . 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, **30**(14), 5419–5454.

[22] Gupta, S., Indumathy, K., & Singhal, G. 2016. Weather prediction using normal equation method and linear regression techniques. *International journal of computer science and information technologies*, **7(3)**(11).

[23] Hersbach, H., de Rosnay, P., Bell, B., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Alonso-Balmaseda, M., Balsamo, G., Bechtold, P., Berrisford, P., Bidlot, J.-R., de Boisséson, E., Bonavita, M., Browne, P., Buizza, R., Dahlgren, P., Dee, D., Dragani, R., Diamantakis, M., Flemming, J., Forbes, R., Geer, A. J., Haiden, T., Hólm, E., Haimberger, L., Hogan, R., Horányi, A., Janiskova, M., Laloyaux, P., Lopez, P., Munoz-Sabater, J., Peubey, C., Radu, R., Richardson, D., Thépaut, J. N., Vitart, F., Yang, X., Zsótér, E., & Zuo, H. 2018. *Operational global reanalysis: progress, future directions and synergies with NWP*. ERA Report Series.

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

[25] Ivanov, V. Y., Bras, R. L., & Curtis, D. C. 2007. A weather generator for hydrological, ecological, and agricultural applications. *Water resources research*, **43**(10).

[26] Jackson, R. D., Idso, SB., Reginato, RJ., & Pinter Jr, PJ. 1981. Canopy temperature as a crop water stress indicator. *Water resources research*, **17**(4), 1133–1138.

[27] Jeong, D. Il., St-Hilaire, A., Ouarda, T. BMJ., & Gachon, P. 2012. Multisite statistical downscaling model for daily precipitation combined by multivariate multiple linear regression and stochastic weather generator. *Climatic Change*, **114**(3-4), 567–591.

[28] Joe, H. 1997. *Multivariate models and multivariate dependence concepts*. CRC Press.

[29] 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.

[30] Josse, J., Pagès, J., & Husson, F. 2011. Multiple imputation in principal component analysis. *Advances in data analysis and classification*, **5**(3), 231–246.

[31] Kalma, J. D., McVicar, T. R., & McCabe, M. F. 2008. Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surveys in Geophysics*, **29**(4-5), 421–469.

[32] Kim, S., Pan, W., & Shen, X. 2014. Penalized regression approaches to testing for quantitative trait-rare variant association. *Frontiers in genetics*, **5**, 121.

[33] Leduc, C., Calvez, R., Beji, R., Nazoumou, Y., Lacombe, G., & Aouadi, C. 2004. Evolution de la ressource en eau dans la vallée du Merguellil (Tunisie centrale). *Pages 10–p of: Séminaire sur la modernisation de l'agriculture irriguée*. IAV Hassan II.

[34] Leduc, C., Ammar, S. Ben, Favreau, G., Beji, R., Virrion, R, Lacombe, G., Tarhouni, J, Aouadi, C, Chelli, B.Z., Jebnoun, N, *et al.* . 2007. Impacts of hydrological changes in the Mediterranean zone: environmental modifications and rural development in the Merguellil catchment, central Tunisia/ Un exemple d'évolution hydrologique en Méditerranée: impacts des modifications environnementales et du développement agricole dans le bassin-versant du Merguellil (Tunisie centrale). *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, **52**(6), 1162–1178.

[35] Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K.C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., & Thiele-Eich, I. 2010. Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, **48**(3).

[36] Massman, WJ. 1992. A surface energy balance method for partitioning evapotranspiration data into plant and soil components for a surface with partial canopy cover. *Water Resources Research*, **28**(6), 1723–1732.

[37] McCullagh, P., & Nelder, J. A. 1989. *Generalized linear models*. Monographs on statistics and applied probability. Chapman and Hall.

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

[39] Molle, F., & Wester, P. 2009. *River basin trajectories: societies, environments and development*. Vol. 8. IWMI.

[40] Norman, J. M, Kustas, W. P, & Humes, K. S. 1995. Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature. *Agricultural and Forest Meteorology*, **77**(3-4), 263–293.

[41] OMM. 2010. *Guide des instruments et des méthodes d'observation météorologiques*. OMM 8, Geneve.

[42] Palmer, T. N. 2001. A non linear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society*, **127**(572), 279–304.

[43] Roblès, B., Avila, M., Duculty, F., Vrignat, P., Begot, Stéphane, & Kratz, Frédéric. 2012. Mesures de pertinence par les critères du" maximum de vraisemblance" de" BIC" et" AIC" appliqués à l'évaluation des paramètres stochastiques de Modèles de Markov Cachés. *Pages 1–12 of: Journal national de la recherche en IUT*, vol. 1.

[44] Saadi, S., Boulet, G., Bahir, M., Brut, A., Delogu, E., Fanise, P., Mougenot, B., Simonneaux, V., & Chabaane, Z. 2018. Assessment of actual evapotranspiration over a semi arid heterogeneous land surface by means of coupled low-resolution remote sensing data with an energy balance model: comparison to extra-large aperture scintillometer measurements. *Hydrology and Earth System Sciences*, **22**(4), 2187–2209.

[45] Schwarz, G. 1978. Estimating the Dimension of a Model. *Ann. Statist.*, **6**(2), 461–464.

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

[47] Van Den Dool, H. M. 1989. A new look at weather forecasting through analogues. *Monthly weather review*, **117**(2230).

[48] Vrac, M., & Friederichs, P. 2015. Multivariateintervariable, spatial, and temporal bias correction. *Journal of Climate*, **28**(1), 218–237.

[49] Warner, T. T. 2010. *Numerical weather and climate prediction.* Cambridge University Press.

[50] Wilby, R. L., Charles, SP, Zorita, E., Timbal, B., Whetton, P., & Mearns, LO. 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods. *Supporting material of the Intergovernmental Panel on Climate Change, available from the DDC of IPCC TGCIA*, **27**.

Appendix A. Surface Energy Balance Model

SPARSE model is forced by a series of climatic observations composed of global solar radiation (GR), air temperature (AirT), relative humidity (Rh) and wind speed (WS). Moreover, the model requires a description of initial conditions and characteristics of the surface cover. Thus, we deploy remotely sensed data from the latest collection 6 of MODIS, available on (http://earthexplorer.usgs.gov).

We use the temporal 16-day composite series of MODIS NDVI (MOD13A2), daily Land Surface Temperature (LST), surface emissivity and viewing angle from (MOD11A1), and 8-day of albedo series (MCD43A3) having a spatial resolution of 500m. These data are acquired for our study period (2012-2016), at the resolution of the MODIS sensor at 1 Km. We extracted a sub image covering the whole plains. In addition, we performed a temporal interpolation of albedo and NDVI data to have daily information corresponding to the satellite overpass. Then, NDVI informations are used to compute remotely

sensed leaf area index LAI. Other parameters and constants are also necessary as inputs in SPARSE model, such us: vegetation height , the roughness length, Minimum stomatal resistance, Atmospheric forcing height etc. Inputs used and required in SPARSE model are well described in [44].

SPARSE outputs are simulated at meteorological time steps (30 minutes in our case). The model provides estimates of instantaneous surface fluxes by solving the energy budgets of the soil and the vegetation. So that, a system of three main equations should be solved iteratively ;

$$\begin{cases} Rns &= G + Hs + LEs \\ Rnv &= Hv + LEv \\ \sigma T^4 rad &= Ratm - Ras - Rav \end{cases}$$

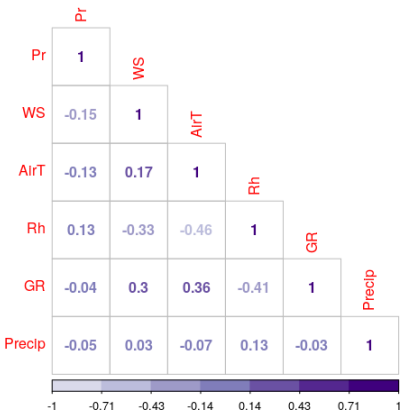
Ratm is the atmospheric radiation (W/m2), Ra is the net component longwave radiation (W/m2) and Trad is the radiative surface temperature (K) obtained from satellite acquisition; indexes s and v designate the soil and the vegetation components of the total fluxes, respectively. The first and the second equations represent the energy budget of the soil and the vegetation, and the third reports the link between the radiative surface temperature Trad and its two component skin temperature sources (Ts and Tv).

Appendix B. Inter-dependency evaluation

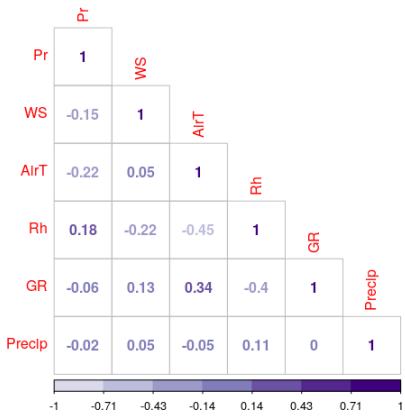
We use pair-plots with non-parametric rank correlations: Kendalls tau, a correlation analyse measure the strength of the relationship between two variables.

Kendall's coefficients are under-estimated in a number of cases by the large-scale variables. For instance, atmospheric pressure with air temperature (-0.22 instead of -0.13) or wind speed with air temperature (0.05 instead of 0.17) and global radiation (0.13 instead of 0.3), see Fig. B.12a and B.12b. Except for the first example, the SWG, see Fig. B.12c, yields closer Kendall's coefficients.

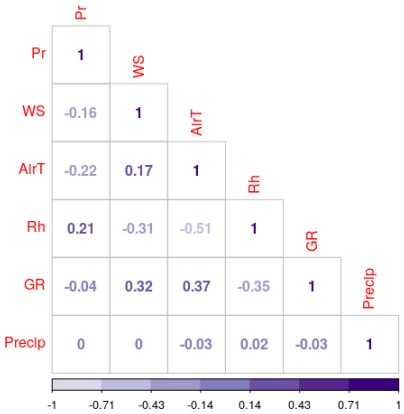
For example, in observed series, we observe that the rank correlation between the wind speed (WS) and the air temperature (AirT) is about 0.17 (figure B.12). It means a positive correlation between these variables. Using the weather generator and multivariate bias correction, we succeed to reproduce the strength of the relationship between these two variables, less reproduced using the univariate correction and inability to preserve this link using reanalysis (Kendalls tau is about 0.05). A negative non linear correlation is also observed between the air temperature and global radiation using different sources of climatic data.



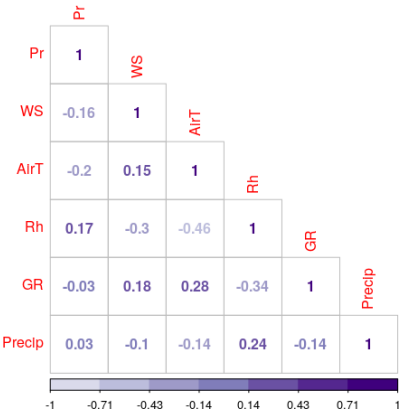
(a) *Observations*



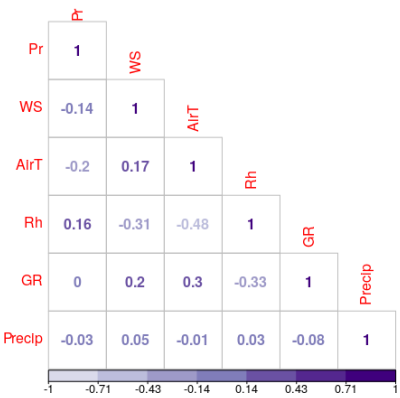
(b) *Large scale variables*



(c) *SWG*



(d) *univ BC*



(e) *multi BC*

Figure B.12: *Inter-variable dependency*