Calibration of a parsimonious distributed ecohydrological daily model in a data scarce basin using exclusively the spatio-temporal variation of NDVI

Guiomar Ruiz-Pérez[1], Julian Koch[2,3], Salvatore Manfreda[4], Kelly Caylor[5] and Félix Francés[6]

[1]Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden
[2]Department of Hydrology, Geological Survey of Denmark and Greenland, Copenhagen, Denmark
[3]Department of Geosciences and Natural Resources Management, University of Copenhagen, Denmark
[4]Department of European and Mediterranean Cultures, University of Basilicata, Matera, Italy
[5]Bren School of Environmental Science and Management & Department of Geography, UC Santa Barbara, California, USA
[6]Research Group of Hydrological and Environmental Modelling (GIHMA), Research Institute of Water and Environmental Engineering, Universitat Politècnica de València, Spain

Correspondence to: Guiomar Ruiz-Pérez (Guiomar.ruiz.perez@slu.se)

Abstract. Ecohydrological modelling studies in developing countries, such as sub-Saharan Africa often face the problem of extensive parametrical requirements and limited available data. Satellite remote sensing data may be able to fill this gap, but require novel methodologies to exploit its spatio-temporal information that could potentially be incorporated in model calibration and validation.

The present study tackles this problem suggesting an automatic calibration procedure, based on Empirical Orthogonal Functions techniques, for distributed ecohydrological daily models. The procedure is tested with the support of remote sensing data in a data-scarce environment – the Upper Ewaso river basin in Kenya. In the present application, the TETIS-VEG model is calibrated only using NDVI (Normalized Difference Vegetation Index) data derived from MODIS. The results demonstrate that: (1) satellite data of vegetation dynamics can be used to calibrate and validate ecohydrological models in water-controlled and data-scarce regions; (2) the model calibrated only using satellite data is able to reproduce both the spatio-temporal vegetation dynamics and the observed discharge at the outlet; and (3) the proposed automatic calibration methodology works satisfactorily and it allows for a straightforward incorporation of spatio-temporal data in the calibration/validation framework of a model.

1 Introduction

Drylands cover 30% of the Earth's land surface and 50% of Africa (Franz *et al.,* 2010). Projections of the IPCC (Intergovernmental Panel on Climate Change, 2007) indicate that the extent of these regions will likely increase in the coming decades. Dryland expansion would have a considerable additional impact on water resources, which should be taken into account by water management plans (Franz *et al.,* 2010).

In water-controlled ecosystems, the vegetation assumes a critical role influencing all components of the hydrological cycle (Rodriguez-Iturbe *et al.,* 2001; Manfreda and Caylor, 2013). For instance, actual evapotranspiration (aET) may account for more than 90% of the annual precipitation in water-controlled areas (Zhang *et al*., 2016; Jasechko *et al.,* 2013). Montaldo *et al.* (2005) affirmed that the use of constant LAI (Leaf Area Index) values, commonly used in hydrological applications, produces large errors in land surface flux predictions. Given the strong control exerted on aET by the vegetation, reliable estimates of spatio-temporal variations of vegetation patterns are vital to obtain trustworthy predictions of available water resources (Andersen, 2008). In this sense, ecohydrological modeling becomes essential in order to include the vegetation dynamics as an additional state variable (Rodriguez-Iturbe *et al.,* 2001).

Evidence of aET being prevalent driver of hydrological records of streamflow and water-table depth, i.e. available water resources, has been observed in many studies (Bond *et al*., 2002; Nyholm *et al*., 2003;Gribovszki *et al.,* 2008, Conradt *et al.,* 2013). Recently, Tsang *et al.* (2014) showed that adding a better evapotranspiration scheme in a widely used runoff model improves streamflow predictions. Conradt *et al.* (2013), who compared three different strategies for deriving sub-basin aET, affirmed that incorporating spatial variation of aET in a semi-distributed model increases its robustness. Contrarily, Stisen *et al.* (2011) and others stressed that those improvements are not necessarily seen in the outlet hydrograph. However, it could also be interpreted in the inverse sense; good performances in terms of the outlet hydrograph do not necessarily mean more reliable estimates of aET.

The stream flow record is traditionally the only observation used for the calibration of hydrological models, but several studies demonstrated the limited capabilities of such an approach when models are validated at interior points of a river basin. Discharge represents an integrated catchment response, and hence provides only limited insight on the lumped behavior of a catchment (Stisen *et al.,* 2011; Koch *et al.*, 2016a; Michaud and Sorooshian, 1994; Reed *et al.,* 2004 or Smith et al., 2013). In that sense, Conradt *et al.* (2013) provided several examples for large simulation errors within the model domain and they mentioned, among others, the outcomes given by Feyen *et al.* (2008), Merz *et al.* (2009) and Smith *et al.* (2012). Moreover, Wi *et al.* (2015) pointed out also that caution is needed when using an outlet calibration approach for streamflow predictions under future climate conditions. This leads to the idea of using spatial state variables with which to implement the new era of distributed (temporal and spatially) models emerges in order to balance the conceptual distributed nature of this kind of models (Stisen et *al.,* 2011).

Traditional observation which generally consist of point data with little spatial support are effectively strengthened by remote sensing data which offer the capacity to provide detailed spatial coverage and pattern information (Franssen *et al.,* 2008, McCabe *et al.,* 2008 and Stisen *et al.,* 2011). Additionally, satellite data has the great advantage to be also available in data scarce areas. In this sense, the application of remotely sensed data represents an excellent source that provides information with a fairly good spatial/temporal resolution (Yang *et al.,* 2012). In modeling, remote sensing data has been basically utilized in three different ways: (1) as forcing data (Xiao *et al.,* 2004; Yuan *et al.,* 2010; Samaniego *et al.,* 2011; and Stisen *et al.,* 2011), (2) as priori information of particular parameters (Winsemnius *et al.,* 2008; Stisen *et al.,* 2011) and (3) for model calibration and validation (see next section for an in-depth discussion of this point).

Satellite imagery provides not only temporal information but also valuable information on spatial patterns which can facilitate a spatial pattern orientated model evaluation. As highlighted by Koch *et al.* (2015), spatial model evaluation is an active field of research not only in Hydrology but also in other disciplines, such as, Atmospheric Sciences (Brown *et al.,* 2011; Gilleland *et al.,* 2010). However, up to now, there exists no formal guideline on how to assess the goodness of fit of the simulated spatial

5   patterns and little information can be found about how to effectively utilize spatio-temporal data. Some authors such as Conradt *et al.,* (2013), Graf *et al.,* (2014) and Koch *et al.,* (2015 and 2016b) focused on developing and testing metrics to be employed when spatio-temporal data is involved. For example, Koch *et al.* (2015) compared Kappa statistics, Fuzzy theory, and EOF-analysis in an attempt towards a true spatial model evaluation of distributed models. But, besides these efforts, there are only a limited number of spatial validation studies that fully embrace the availability of satellite remote sensing data by means of

10   true spatial performance metrics (Koch *et al.,* 2016b). The EOF analysis is a versatile methodology to investigate the spatio-temporal patterns of fluxes and states in the soil-vegetation-atmosphere continuum (Fang *et al.,* 2015). As mentioned previously, Koch *et al.* (2015) carried out a validation of a distributed model using satellite based land surface temperature data by means of an EOF analysis. With other statistical purposes, the EOF analysis was used by Graf *et al.* (2014), Kim and Barros (2002) and Liu (2003). A fine scaled study was carried out by Drewry and Albertson (2006) who used the EOF analysis

15   to associate spatial pattern in the errors of a canopy-atmosphere model with errors in the parameters. But, to our knowledge, the EOF analysis has not been applied in model calibration yet.

In summary, the main objectives of this research are: (1) to incorporate spatio-temporal data into the calibration process by applying the EOF methodology as an objective function and (2) to exploit satellite data as a proxy of reliable estimates of vegetation dynamics for both the calibration and validation of an ecohydrological model. To address these key challenges, a

20   distributed parsimonious ecohydrological model was applied in a water-controlled basin located in Kenya.

2 Satellite data and model calibration/validation

The applicability of remote sensing to calibrate and/or validate a model by exploiting information on spatial patterns still remains a challenging task. In order to better understand this issue in more detail, a bibliographic survey of the ISI Web of Knowledge Science Citation Index database was undertaken using the following words combinations in the topic search: (1)

25   satellite calibration, (2) satellite implementation, (3) satellite ecohydrological modelling, and (4) remote sensing ecohydrology. From the total number of publications obtained by this search, only those that incorporated satellite data to specifically models calibration were selected.

At plot scale, Quevedo and Francés (2008) and Pasquato *et al.* (2015) calibrated and validated a parsimonious ecohydrological model using satellite information. More recently, Ruiz-Pérez *et al.* (2016) discussed the applicability of satellite data during

30   the calibration process comparing the results obtained by a parsimonious model calibrated only using satellite data against the results obtained by a complex model calibrated using field measurements at pixel scale. Similarly, Quevedo and Francés (2008) and Pasquato *et al.* (2015) used time-series of NDVI to validate a parsimonious ecohydrological model named HORAS. At

catchment scale, Immerzeel and Droogers (2008) used satellite-based evapotranspiration in combination with observed streamflow to calibrate the semi-distributed SWAT. Zhang *et al.* (2009) concluded that multi-objective calibration of SymHyd model against streamflow and satellite-based aET produced better daily and monthly runoff compared to calibration with streamflow alone. More recently, Rientjes *et al.* (2013) calibrated a semi-distributed hydrological model using streamflow data

5   and satellite-based aET. Regarding other satellite products, GRACE data, which can be used to detect variations in terrestrial moisture storage (e.g. Lettenmaier and Famiglietti, 2006), have been used to calibrate both global and regional-scale surface hydrology models, in combination with stream discharge data (e.g. Lo *et al.*, 2010). Zhang *et al.* (2011) calibrated the AWRA-L model with streamflow, NOAA-AVHRR LAI and TRMM-MI (Tropical Rainfall Measuring Mission- Microwave Imager) soil moisture using a multi-objective calibration framework. Only few studies carried out the calibration exclusively against

10   remote sensing data. For instance, Gutmann *et al.* (2010) calibrated landscapes hydraulic properties in the Noah land surface model using only MODIS surface temperatures from 14 different sites and using observed flux data for model verification. Also, Velpuri *et al.* (2012) modelled Lake Turkana water level only using satellite information. All the above mentioned studies shared the same conclusion: including remote sensing data into the model calibration/validation improves the overall performance.

15   In general, from the total of reviewed publications, calibration only using satellite data was performed in the 47% of cases while a combination of satellite data and field measurements (specially, streamflow at the outlet) was used in the remaining contributions. Similar results were obtained regarding to the validation: 35.3% of publications adopts only field measurements (specially, historical streamflow) employing satellite data exclusively for the model calibration, 47% using a combination of field measurements and satellite data, 11.8% using only satellite data and one publication without any specification. But, more

20   interesting is how the different calibrations were carried out. In most of the cited examples, a sort of multi-objective calibration was used adopting only some points/pixels to calibrate the entire catchment. In other cases, lumped or semi-distributed models were implemented instead of fully distributed ones, considering aggregated values of the satellite data. In other words, the spatial heterogeneity of the basin is neglected and the full potential of satellite imagery, namely the information on spatial patterns, is not fully exploited. Therefore, a method able to make use of the potential of the spatio-temporal information

25   contained in remote sensed data is highly desirable as well as a calibration scheme which relies solely on remote sensing data will be greatly beneficial in modelling of data scarce catchments (Kunnath-Poovakka *et al.,* 2016).

3 Study area and data

The Upper Ewaso Ngiro Basin is located in the Laikipia region of Kenya (Figure1). The basin is part of the Laikipia Plateau which lies between Mount Kenya (South East) and the Aberdare Mountains (South West). The basin has a drainage area of

30   15,200 km2, with the largest river being the Ewaso Ngiro. This region is characterized by distinct rainy and dry seasons. The first rainy season occurs from March to May, while the second rainy season occurs from October to December. Both air

temperature and precipitation patterns are heavily influenced by elevation. A full description of the precipitation patterns in the region can be found in Franz (2007).

Soil texture ranges from sandy clay to clay soils (according to the 1980 UNESCO Soil Map). Although the most characteristic landscape is savanna, higher elevations are dominated by forests and a large piece of land has been converted to cropland (Franz, 2007). The remaining surfaces of the study region are classified as grassland, shrubland, and wooded grassland (savanna ecosystems).

As meteorological data, we used the weather stations of the Natural Resource Monitoring, Modeling and Management Project (NRM3) of Nanyuki, Kenya illustrated in Figure 1. Daily precipitation and temperature from 1959 to 2003 were validated by Franz *et al.* (2010). Considering the available hydrological information, we selected a sub-basin with an area of about 4,600 km2 for the present study (Figure 1). The selected catchment is equipped with a streamflow gauge at the outlet, operating from 1980 to 2002. The reference evapotranspiration (ET_0) was calculated using the Penman Monteith equation with the simplifications proposed by Allen *et al.* (2006). This approach is extremely useful to describe the spatial distribution of solar radiation and to derive the ET_0 maps during any day of the year (Manfreda et al., 2013).

Regarding the satellite data, we adopted the Normalized Difference Vegetation Index (NDVI) included in the MOD13Q1 and MYD13Q1 products provided by NASA (NASA Land Processes Distributed Active Archive Center (LP DAAC)). This satellite product is available from 2000 to present. The MOD13Q1 and MYD13Q1 data are provided every 16 days at 250-meter of spatial resolution. The used NDVI products (MOD13Q1 and MYD13Q1) are in level 3 that means they are not raw satellite data. NDVI indices are retrieved from daily, atmosphere-corrected, bidirectional surface reflectance.

At last, based on previous experiences (Ruiz-Pérez *et al.,* 2016 and Pasquato *et al.* 2015) and in similar climatic conditions, we declined to use other products such as LAI or ET derived from MODIS because this kind of products are produced by models. And, for example, Ruiz-Pérez *et al.* (2016) found large discrepancies between the LAI provided by satellite and the LAI measured in field. At this point, we had no information to determine the accuracy of these particular models and the spatial information used to implement them. In contrast, NDVI values are calculated by direct differences of spectrum bands, i.e. no models are involved and we therefore decided to use this latter product instead of satellite LAI and/or ET.

4 Model description: TETIS-VEG

The proposed model, called TETIS-VEG, is based on a distributed hydrological model called TETIS (Francés *et al.,* 2007) coupled with a dynamic vegetation model. Both models have simplicity in model structure in common. The used equations are as simple as possible in order to reduce the number of parameters (Table I). The sub-models are inter-connected. The transpiration calculated in the hydrological sub-model depends on the leaf area index (LAI) simulated by the dynamic vegetation model. At the same time, the simulated LAI depends on the water stress which is calculated using the hydrological sub-model. The hydrological sub-model can be used at different time scales (from few minutes up to daily time steps) while the vegetation dynamic sub-model has to be applied at daily scale. Hence, the TETIS-VEG model must be used at daily scale.

Both sub-models can be used at a broad range of spatial scales. In this research, the resolution of the implemented model was 90X90 meters.

4.1 The hydrological sub-model: TETIS

TETIS's conceptual scheme consists of a series of connected reservoirs, each one representing different water storages in the soil column: vegetation interception, first static soil layer (retained water by upper soil capillary forces, i.e., below field capacity plus water detention in surface puddles; evaporation and transpiration can occur), second static soil layer (retained water in deeper soil by capillary forces; only transpiration can occur), surface (for overland runoff), gravitational soil layer (upper soil water content above field capacity for interflow) and aquifer (for river baseflow). Vertical connections between reservoirs describe the precipitation, evapotranspiration, infiltration and percolation processes. The horizontal flows describe the three different hydrological responses that give the discharge at the catchment outlet: overland runoff, interflow and baseflow. A more detailed description of the TETIS model can be found in Francés *et al.* (2007) and GIMHA (2014).

The TETIS model uses a split-structure for the effective parameter value at each cell (Francés and Benito, 1995; Francés *et al.,* 2007). The effective parameter is calculated using a correction factor multiplied by the estimated value of the parameter in each cell using all the available information (land cover map, soil type map, DEM, depth of roots and soil layer, etc.) and expert's knowledge. Hence, the effective parameter in each grid cell is computed as the product of two terms: (1) a common correction factor for each type of parameter that takes into account the model, information and input errors and the temporal and spatial scale effects; and (2) the a priori estimated value at each cell. For a given parameter, the a priori and effective values are different from cell to cell while the correction factor is common for all cells (and different from map to map).

With the split-parameter structure, only nine correction factors are calibrated. Each one related to one of these estimated parameter maps: maximum static storage, reference evapotranspiration, infiltration capacity, hillslope velocity, percolation capacity, horizontal saturated conductivity for interflow, horizontal saturated conductivity for aquifer and river channel velocity.

4.2 The dynamic vegetation sub-model: LUE-model

The proposed dynamic vegetation sub-model is based on the concept of Light Use Efficiency (LUE) (Medlyn, 1998) and calculates the leaf biomass (Bl) according to the equation 1. The LUE is the proportionality between plant biomass production by terrestrial vegetation and absorbed photosyntetically active radiation (APAR) in optimal conditions. However, the LUE can be strongly affected by stress conditions. The key factors contributing to the variation of this efficiency are: soil moisture content, air temperature (Landsberg and Waring, 1997; Sims *et al.,* 2006), and nutrient levels (Gamon *et al.,* 1997; Ollinger *et al.,* 2008). Since this model is designed to be used in water-controlled areas, the nutrient levels are not considered.

$$\frac{dB_l}{dt} = (LUE * \varepsilon * PAR * fPAR - Re) * \varphi_l(B_l) - k_l * B_l \tag{1}$$

where Bl is the leaf biomass, LUE is the above-mentioned Light Use Efficiency, ε is the water stress factor, PAR is the Photosynthetically Active Radiation, fPAR is the fraction of Photosynthetically Active Radiation, Re is the respiration, $\varphi_l(B_l)$ is the fractional leaf allocation and kl is the leaf natural decay factor to reproduce the senescence. The water stress factor depends on the amount of water contained in the two static reservoirs and it is calculated according to Porporato *et al.* (2001).

5 Basically, the stress factor is equal to 1 (maximum stress) if the water storage is less than the water storage at wilting point; it is equal to 0 (minimum stress) if the water storage is higher than the water storage at critical point (plants start the stomatal closure); and, it varies from 0 to 1 using a potential function which depends on the wilting point, the critical point and an exponent set equal 2. Then, this stress multiplies the LUE index, reducing the efficiency when its value is lower than 1 (non-optimal conditions).

10 The LAI is simulated through the product between the leaf biomass, the specific leaf area (SLA) and the vegetation fractional cover. Later, the LAI is used to calculate the transpiration in the hydrological sub-model according to the Eq. (2).

$$T_i = (ET_0 - EI) * \min(1, LAI) * \zeta * Z_i * fc \tag{2}$$

where Ti is the transpiration from the i soil layer, ET_0 is the reference evapotranspiration, EI is the evaporation of the intercepted water, LAI is simulated by the model, Zi is the percentage of roots in the i soil layer and fc is the coverage factor.

15 Therefore, the LUE-model has eight parameters to be calibrated: (1) Specific leaf storage (the maximum interception storage is calculated as the product between the specific leaf storage and the LAI simulated by the model), (2) the LUE index (explained above), (3) the coverage factor, (4) the distribution of roots between the first and the second static storage layers, (5) the maximum LAI sustainable by the system (the simulated LAI is limited by a maximum), (6) the light extinction coefficient k , (7) the SLA and, (8) the optimal temperature (the stress factor also depends on the temperature).

20 **5 Methodology**

One of the main objectives of this research was to explore the potential of satellite remotely sensed data for model calibration. The TETIS-VEG model was calibrated purely against MODIS NDVI according to the following three steps: (1) a manual calibration in order to obtain a first approximation of model parameters, (2) an automatic calibration based on the combined use of EOFs and a genetic algorithm in order to refine model parametrization and (3) a model validation carried out with both

25 remote sensed data and traditional data (such as streamflow measurements). Since the meteorological data (precipitation and temperature) were available from 1960 to 2003 and the MODIS NDVI was available from 2000 to present, we decided to use the year 2003 as the calibration period and the period from 2000 to 2002 for validation. In order to avoid the effect of the initial conditions, we used one year as warming up period (the year 2002 and 1999 for model calibration and validation respectively). For these purposes, we adopted the NDVI as a descriptor of the state of the vegetation assuming that LAI and NDVI are

30 intimately related. Studies on various vegetation types, e.g., agroecosystems (Cohen *et al.,* 2003), grasslands (Friedl *et al.,* 1994), shrublands (Law and Waring, 1994), conifer forests (Chen and Cihlar, 1996), and broadleaf forests (Frassnacht *et al.,*

1997) have led to the general conclusion that the spectral vegetation indices such as NDVI have considerable sensitivities to LAI. The relationship between LAI and NDVI can be considered linear for low values, while it becomes nonlinear for the higher values of NDVI due to the greenness saturation (e.g., Turner *et al.,* 1999). In this case study, the maximum LAI values are around 2.0 – 2.5, according to Franz (2007), that are lower than the greenness saturation threshold. Therefore, the relationship between the observed NDVI and the simulated LAI is expected to be linear.

5.1 Empirical Orthogonal Function method (EOF)

The EOF method is generally used to analyze the spatio-temporal variability of a single variable but, comparison between different variables can also be performed using coupled EOF techniques (Björnssson and Venegas, 1997). The method decomposes a dataset in a time series and spatial patterns. Furthermore, the method allows to estimate a measure of the "importance" of each spatial pattern. We refer to the spatial patterns as the EOFs (in literature, they are also referred to as principal components), and to the time variation as loadings (in literature, there are several terms: expansion coefficient time series, expansion coefficients, EOF time series, principal components time series, etc.).

The EOF method is essentially a linear algebra methodology based on matrix transformation. The first step is thus the conversion of the spatio-temporal data to be analyzed into a matrix. Basically, we construct a matrix (F) in which each column is the temporal variation of the data in a particular cell while each row represents the cells values during a particular time step. Usually, the second step is to compute the anomalies of the analyzed data which was not needed in this study because we used normalized data (for reasons that will be explained below).

The next step of the applied EOF method consists on the calculation of the spatial F's covariance matrix (R) according to Eq. (3). Then, the eigenvalue problem is solved Eq. (4).

$$R = F^T * F \tag{3}$$

$$R * C = C * \Lambda \tag{4}$$

Λ is a diagonal matrix containing the eigenvalues λi of R. The ci column vectors of C are the eigenvectors of R corresponding to the i-respective eigenvalues. Each of these eigenvectors can be regarded as a map which denote the EOFs (or principal spatial patterns). In what follows, we always assume that the eigenvectors are ordered according to the value of the eigenvalues. Thus, EOF1 is the eigenvector associated with the biggest eigenvalue. The fraction of the total variance in R explained by EOFi is found by dividing the λi by the sum of all the other eigenvalues. The time evolution of an EOF$_i$ ($\vec{a_i}$) is calculated according to Eq. (5). The components of this time vectors are referred to as loadings in this paper.

$$\vec{a_i} = F * EOF_i \tag{5}$$

Using the spatial covariance calculated according to the Eq. (3), the EOF technique provides three different results: the main patterns or EOFs, their time evolution whose components are called loadings and the portion of spatial variance explained by each EOF which is calculated dividing each λ by the trace of Λ.

5.2 Manual calibration

5 The manual calibration was done with a dual purpose. First, we wanted to test the applicability of the proposed TETIS-VEG model in the study basin. Second, we wanted to obtain a first approximation for the parameters and, at the same time, constrain the automatic calibration. This manual calibration consisted of the usual ad hoc method (manual adjustment of parameter values) considering the Pearson correlation coefficient between the simulated LAI and the observed NDVI in a total of 32 different points inside the basin. These points were selected within homogeneous areas defined according to the main spatial

10 patterns of the observed NDVI (EOFs) and the available maps of land cover, soil texture, DEM, slope and soil depth.
In this case, the EOF analysis was used to identify the main spatial patterns of the observed NDVI. Once the main spatial patterns were identified, we combined our own human perception with the confusion matrices between the main spatial patterns and the spatial maps of model parameterization. Confusion matrices are widely applied for map comparison in distributed modelling comparing actual to predicted values for each specific category defined previously (García-Arias *et al.,* 2016;

15 Bennett *et al.*, 2013; Van Vliet *et al.,* 2013 among many others). Generally, the rows in the matrix represent the values predicted by the model, whereas the columns represent the actual values. By its nature, the confusion matrix is an overall measure for similarity between two categorized maps. However, the comparison of numerical maps is feasible if they are categorized previously. While most of the spatial maps of the basin characteristics (land use, soil type, etc.) were categorical, the main patterns obtained by the EOF analysis were numerical. To build the confusion matrices, the main patterns were therefore

20 discretized according to the number of river basin features (such as land cover map, soil type map, etc.) and based on the similitude between the corresponding histograms. Once the discretization was done, by a cell-by-cell comparison of the discretized NDVI main pattern maps obtained after the EOF analysis and the available spatial maps, the confusion matrices were built.
These confusion matrices allowed the calculation of the weighted kappa (k) coefficient (Cohen, 1968). This coefficient, whose

25 maximum value is 1, representing a perfect agreement, was employed to identify which spatial maps (land cover map, soil type map, DEM, etc) were linked with the main patterns of the observed NDVI. Then, they were used in order to select the most appropriate points for the manual calibration.

5.3 Automatic calibration

The most innovative aspect of the proposed procedure was the direct use of the EOF analysis in the automatic calibration. As

30 proposed by Koch *et al.* (2015), we decided to build one integral matrix concatenating both the observed and predicted data: the matrix contained the normalized values of the NDVI provided by MODIS and the normalized values of the LAI simulated by the model. In this way, the upper part of this matrix contained the temporal variation of the normalized observed NDVI in

all cells as columns while the lower part contained the temporal variation of the normalized simulated LAI in all cells as columns. We decided to use the normalized values of the NDVI and LAI because, although they are correlated, they differ in range.

However, normalization implies that some spatial information is lost. In order to avoid these losses, we added two rows in the matrix F: the first containing the difference between the temporal mean of the observed NDVI at a particular cell and the general mean using the complete NDVI dataset; and the second with the same content but referred to the simulated LAI. In this way, we included the spatial gradient of the observed NDVI and the spatial gradient of the simulated LAI. These two rows represents two additional maps included in the evaluation of the model performance. If they were similar, it would mean that the spatial gradient remains and is properly reproduced.

The number of pixels was 1,034,706. For the calibration period (year 2003), there were 44 NDVI maps (one each 8 days more or less). Hence, the built integral matrix's size was 90 rows (44 + 44 + 2 additional rows) X 1,034,706 columns. After the construction of this matrix, the EOF analysis was applied obtaining: the EOF maps for the matrix containing both NDVI and LAI, the portion of variance explained by each EOF map and the loadings of each EOF map. The combined EOF analysis yielded orthogonal EOF maps that explained the combined intervariability and intravariability of both data sets. For each time step, the loadings express how much the respective LAI and NDVI map contribute to the direction of the corresponding EOF. Hence, if the observed NDVI and the simulated LAI were completely correlated, the temporal evolution of the EOF maps for both, NDVI and LAI, would be essentially equal.

The automatic model calibration was carried out trying to minimize the differences between the loadings of simulated and observed data. The used objective function was based on that idea and it also took into account the portion of variance explained by each EOF in order to consider that the variance contribution decreases consecutively for the EOFs. The adopted error measure is described in following equation:

$$Error = \sum_{i=1}^{k} w_i * \sum_{i=1}^{t} |load\_sim_{i,j} - load\_obs_{i,j}| \tag{6}$$

where Error is the objective function to minimize, w_i is the portion of variance explained by the EOF_i, $load\_sim_{i,j}$ is the loading of the EOF_i at time step j for the simulated data (in this particular case, the normalized LAI) and $load\_obs_{i,j}$ is the loading of the EOF_i at time step j for the observed data (in this particular case, the observed NDVI).

The calibration was performed using a genetic algorithm called Pyevolve (Perone, 2017). This algorithm needs a seed (initial values of the parameters) and a searching boundary of the parameters to be calibrated. We used the results obtained after the manual calibration explained above as seed and made sure that the searching boundaries were wide enough (Table1).

In order to explore the outcomes of the proposed calibration framework, we additionally calculated both the temporal Pearson correlation coefficient between the NDVI provided by MODIS and the LAI simulated by the TETIS-VEG model in each cell

and the spatial Pearson correlation at each time step. For the spatial and temporal correlation coefficients, we used the original values of both datasets (NDVI and LAI), not the normalized values as used by the EOF analysis. It is important to mention that the Pearson correlation coefficient between two datasets X and Y is positive if X and Y tend to be simultaneously greater than, or simultaneously less than, their respective means. Hence, the mean should be representative. For this reason, in the case of the spatial correlation coefficient, we distinguished between the main land covers whose means can be significantly different: tree, shrubs and grass.

5.4 Validation

The period selected for the model validation was of three years from 2000 to 2002. As during the calibration period (year 2003), there were data of precipitation, temperature and, also, NDVI provided by MODIS. To validate the model, we used the same performances indexes applied during the automatic calibration process. Keeping the parameter values obtained by the automatic calibration, we built the matrix concatenating the normalized value of the observed NDVI and the normalized value of the simulated LAI with two additional rows used to incorporate the spatial gradient of both datasets as explained above. We also plotted these two maps and compared them as we did during the model calibration. Using the EOF techniques, we obtained the coupled EOF maps and their associated loadings and portion of variance explained by them. As during the calibration, we compared the deviation of the loadings for each EOF map and we calculated the Error function defined in Eq. (6). We also calculated the temporal and the spatial Pearson correlation coefficient.

In addition to this, we explored the reliability of the calibrated model in reproducing streamflow. In fact, during the validation period, the observed discharge at the outlet point was available unlike during the calibration period. The reliability of the hydrological sub-model in reproducing the streamflow was an extremely challenging task considering that the entire modelling structure had been calibrated only using vegetation data from remote sensing along with physical information about the basin. Furthermore, we included the Nash and Sutcliffe efficiency index (NS-Nash and Sutcliffe, 1970) and the bias (or volume) error (E) value between the observed and simulated discharges at the basin outlet in the model validation. We also decided to strengthen our discharge analysis by using the concept of flow duration curves (FDCs). FDCs are simple and powerful tools, commonly used in hydrology to describe the runoff regime in a river basin that can be representative of the model ability in reproducing the different components of the streamflow (e.g., Manfreda *et al.,* 2005). In fact, FDCs represent the relationship between magnitude and frequency of streamflows, providing thus an important synthesis of the relevant hydrological processes occurring at the basin scale (Pumo *et al*., 2013). Actually, the shape of a flow-duration curve in its upper and lower regions is particularly significant in evaluating the stream and basin characteristics (Coopersmith *et al.,* 2012). The shape of the curve in the high-flow region indicates the type of flood regime the basin is likely to have, whereas, the shape of the low-flow region characterizes the ability of the basin to sustain low flows during dry seasons (Cheng *et al.,* 2012). Hence, the flow duration curve represents the full spectrum of variability in terms of their magnitudes (Wagener *et al.,* 2013).

6 Results

6.1 Manual calibration

The main objective of this a priori manual calibration was the identification of the most appropriate points where the model could be tested. To accomplish that, we identified the spatial main patterns of the observed NDVI and, then, we compared the EOFs with the spatial features of the river basin (such as: land cover map, DEM, soil type map, etc).

Using our own perception, we identified a certain relationship between the EOF_1 (which explained the 61.5% of the observed NDVI spatial variance) and the land-use map. This potential relationship was supported by the K coefficient (described in the methodology section) that assumed a value of 0.34. This is not a high value but it showed the existence of a relationship between the two maps. Regarding the EOF_2 (which explained the 10.5% of the observed NDVI spatial variance), no connections with the basin physical characteristics were found. It might contain a mix of several drivers and, therefore, it can't be directly linked to a single one. Contrarily, the EOF_3 showed a good agreement with the soil texture map (the K coefficient was 0.32). Therefore, we can state that the observed patterns of NDVI are strongly influenced by the spatial distribution of land cover and soil texture. In the following, we combined these two maps, extracted all possible combinations and selected randomly two points of each of these combinations obtaining 32 points covering all the catchment area.

After conducting the manual calibration, the Pearson correlation coefficient between the observed NDVI and the simulated LAI was positive in 25 points of the 32 considered points. All points with negative correlations had in common the fact that they were located near to de Mount Kenya or Aberdare mountains (Figure 2).

Finally, Table I shows the obtained set of parameters. This set was used as seed during the automatic calibration. It must be underlined that all parameters had values consistent with the reviewed literature (references embedded in Table I).

6.2 Automatic calibration

The proposed automatic calibration is based on the assumption that the closer the loadings of the simulated values are to the loadings of the observed values, the higher the similarity between the spatial patterns is. Calibration produced a good agreement between the observed and simulated loadings of the EOF_1 (upper part of the Figure 3, first graphic) while small deviation between the observed and simulated loadings related to the EOF_2 and the EOF_3. The loadings of the remaining EOFs were completely scattered mainly due to their corresponding low contribution (low weight) in the objective function of the automatic calibration process (Eq. 6). It is needed to remark that the EOF_1 explained more than 60% of the dataset spatial variance while the EOF_2 and the EOF_3 explained around 10% each. The remaining EOFs explained less than 3% each, but in any case they were considered during the calibration process (weighted by the portion of variance explained by each one).

We also used three additional metrics to evaluate the model performance: (1) the temporal Pearson correlation coefficient evaluated in each cell, (2) the spatial Pearson correlation distinguishing between trees, shrubs and grasses computed at any time and (3) comparison of the spatial gradient maps. First, the temporal Pearson correlation coefficient between the observed NDVI and the simulated LAI was higher than 0.4 (left panel of Figure 4) in most of the catchment. The weakest correlations

were obtained in the two highest areas of the basin near to the Mount Kenya and Aberdare Mountains with zero to negative values.

The spatial Pearson correlation coefficients were calculated excluding the regions with negative temporal Pearson correlation coefficient. Although slightly worse than the results in terms of temporal correlation, the mean spatial correlations were higher than 0.45 for all main land covers: trees (mean=0.58), shrubs (mean=0.49) and grasses (mean=0.55) (Figure 5, upper panel). The best scores were obtained in cells classified as trees. In fact, the median was almost 0.60 and the variance was not high (standard deviation= 0.16). Contrarily, the cells classified as grasses obtained the worst results with the lowest median and the highest variance (standard deviation= 0.18).

Figure 6 (upper panels) shows the comparison between the maps which represent, in each cell, the difference between the temporal mean and the general mean of the observed NDVI and the simulated LAI respectively. No great differences were found by comparing both maps indicating the good spatial performance of the ecohydrological model, at least from the vegetation point of view.

6.3 Validation

Similarly to the calibration process, the EOF_1 explained more than 60% of the spatial variance while the EOF_2 and the EOF_3 explained around 10% for the validation period. The remaining EOF maps are not presented because none of them explained more than 3%. The simulated and observed loadings of the EOF_1 were almost equal while the obtained results in relation to the EOF_2 and the EOF_3 were slightly worse (lower part of the Figure 3). However, it is important to stress that both showed the same clear temporal dynamics. Indeed, the resulted Error for the validation period was 4.03, just slightly worse than the Error for the calibration period. It must be considered that the Error value was calculated considering all EOFs (Eq. 6).

The temporal Pearson correlation map between simulated LAI and NDVI showed the same pattern observed in the calibration period: the two areas located near to the Mount Kenya and the Aberdare Mountains had a temporal correlation coefficient equal to zero or negative. However, in more than 80% of the catchment, this coefficient was between 0.3 and 0.9 (right panel of Figure 4).

Regarding the spatial Pearson correlation coefficient between simulated LAI and NDVI in the three main land cover, the results were not as good as the results obtained in terms of temporal correlation. Nevertheless, there were no negative spatial correlation coefficients at any time step. In the case of shrubs and grasses, the mean and median were almost 0.4 while the corresponding ones for the trees were around 0.35 (Figure 5, lower panel). The variance obtained during the validation period was narrower than the obtained during the calibration period for the three land covers: trees, shrubs and grasses. The spatial pattern of LAI was, as for the calibration period, well captured by the model (see the lower panels in Figure 6). The cells with high differences between their own temporal mean and the general mean were consistent in both maps.

Finally, since there was observed discharge at the basin outlet during the years 2000, 2001 and 2002, it was possible to compare the discharge simulated by the model with the observations. The volume error (E) was equal to -0.40 while the NS index was equal to 0.32. E is strongly affected by the results obtained at the beginning of the validation period, probably due to the

absence of information regarding the initial conditions. Although we used a year as warming-up period, the simulations improved only after 2001. In fact, having calculated the performances indexes in each year, the E decreased from -0.88 in 2000 to only -0.17 during the year 2002 (Figure 7). Regarding to the NS index, the worst result was also obtained in the first year and it improved from a negative value in 2000 to 0.35 during the year 2002, as one should expect considering the visual

5   comparison in Figure 7. This trend is emphasized in the plot of the FDCs (Figure 8) where the underestimation in the first two years is clearly highlighted. The first panel compares the FDC of observations and simulations within the whole period while the following three panels compared the corresponding FDCs within the 2000, 2001 and 2002. In these plots, the simulation seems to interpret closely hydrological response in the year 2002.

7 Discussion

10   From the *a priori* manual calibration step up to model validation, it was possible to identify a behavioral pattern which would be also observed during the following automatic calibration and validation steps: the EOF_1 explains more than 60% of the spatial variance, the EOF_2 around 10%, the EOF_3 around 5% while the remaining EOFs could be considered negligible. The fact that the EOF_1 and EOF_3 of the observed NDVI was related to the land cover and soil type maps respectively was consistent with our expectations that NDVI is a suitable proxy of vegetation dynamics.

15   After the automatic calibration, the model matched the observed loadings of EOF_1 and its accuracy was slightly poorer for the second and third EOFs. Thus EOF_1 captured the predominant pattern that was found in both, the observed NDVI and the simulated LAI data. Furthermore, on one hand, the temporal variation of the EOF_1 loadings seemed to be related to the two typical growing seasons in the catchment: the first one during March-May and the second one during October-December (Franz *et al.,* 2010) (Figure 3). On the other hand, the loadings of EOF_2 and EOF_3 were not strongly connected with any feature.

20   The loadings of the remaining EOFs were scattered which implies that mainly measurement and model noise were covered by these EOFs.

The weakness of the proposed calibration methodology is that, although the associated weights to the loading deviation in Eq. (6) are needed, they are also misleading some spatial information. New ways to weight the loading deviations must emerge in future researches as proposed by Koch *et al*. (2015). In fact, due to the portion of variance explained by the EOF_1, this first

25   main pattern controlled the calibration process. In future applications, the proposed error index may be improved by focusing or excluding specific EOFs. A popular method for deciding which EOF to keep and which to discard is to use 'selection rules'. Basically, there are three classes of selection rules depending on whether they focus on the amount of variance explained by each EOF, the loadings or the EOF maps (Preisendorfer, 1988). Other option could be to rotate the EOFs as proposed by Bonaccorso *et al.* (2003). Basically, as each rotated EOF will not explain the same variance of the unrotated one, this approach

30   would be an option to use different combinations of EOFs which explain different amount of variance in order to reduce the influence of the EOF_1.

The automatic calibration process worked satisfactorily as shown by the additional metrics: temporal Pearson correlation coefficient, spatial Pearson correlation coefficient in the main land covers and the comparison between the gradient maps. In terms of spatial Pearson correlation coefficient, the weakest values were obtained in the higher portion of the basin near to the Mount Kenya and Aberdare Mountains, while the remaining cells within the study area showed a good agreement between

5   observed NDVI and simulated LAI. This same behavior was also observed when calibrating manually. Two reasons could explain such results. First, the observed NDVI in some cells of those areas had a really bad quality testified by the unrealistic oscillations of the NDVI from 0.8 to 0.1 (even zero) in just one week. These unrealistic oscillations could be produced by the presence of clouds over the area near to the mountains. The second reason is related to the conceptual limitation of the proposed model. The TETIS-VEG was designed to be used only in water-controlled areas. Franz (2007) analyzed the correlation between

10   the fractional woody cover and the mean annual precipitation within the catchment and they were strongly correlated. However, two different slopes were observed. The transition point, which indicates when water availability had a smaller influence on the fractional woody cover, occurred approximately around 800mm/year. Physically, the transition point is believed to be a good approximation of the transition from a water-controlled ecosystem to a nutrient-controlled ecosystem. This approach allowed to define the higher areas within the study catchment (where Mount Kenya and Aberdare Mountains are included) as

15   nitrogen limited ecosystems instead of water-controlled.

With the exemption of these two areas, there existed a strong correlation between NDVI and LAI, i.e. the model captured the temporal dynamic of LAI. However, this did not necessarily mean that the magnitude of LAI was reasonable. This last point was proven by calculating the good performance in terms of spatial Pearson correlation and the comparison between the gradient maps. No differences and good agreements were observed along the main land covers: trees, shrubs and grasses.

20   Finally, there were four parameters in the automated calibration which changed substantially (in relative terms) in comparison to the values obtained during the manual calibration: the correction factor of the maximum static storage, the correction factor of the reference evapotranspiration, the factor related to the distribution of roots between the first and second static storage layers and the maximum LAI sustainable by the system (Table I). These parameters affect directly the transpiration process and the amount of available water to be consumed by the plants. In any case, all obtained values were consistent with the

25   reviewed literature (embedded in Table I). All of them are completely included in the searching boundary used during the automatic calibration.

The validation process confirmed: (1) the model was able to capture completely the EOF_1 while the model performance worsened in the following two EOFs, (2) the simulated LAI and the observed NDVI were temporally correlated in most of the catchment and (3) the spatial distribution of LAI was consistent as shown by the comparison between the gradient maps and

30   the value of the spatial Pearson correlation coefficient at any time.

An additional interesting outcome provided by the validation was the comparison between simulated and observed hydrograph at the outlet. The simulated stream flow was promising, but not completely convincing. This limitation is obvious since the model parameters were calibrated on NDVI data, i.e. the model was calibrated on vegetation dynamics. Therefore the direct

comparison between hydrographs should not be too exigent when that no information about the parameters involved in the river flow routing or aquifer discharge was included in the calibration.

We strengthened our discharge analysis by using the concept of FDCs. By graphical comparison (Figure 8), it could be observed that the model is able to reproduce the shape of the observed FDC, while some discrepancies were found in terms of magnitude. However, its performance improved considerably year over year. Since the FDC shape is an important synthesis of the relevant hydrological processes occurring at the basin scale, this result pointed out the capability of the proposed model calibration methodology to reproduce the main hydrological behavior of the study basin.

8 Conclusions

The main two objectives of this research were: (1) to explore if it is possible to calibrate and validate an ecohydrological model only using satellite information, and (2) to incorporate spatio-temporal data about a model state variable into the automatic calibration process. In order to tackle these questions, a parsimonious distributed ecohydrological model was calibrated using exclusively NDVI data provided by MODIS. A methodology based on the EOF analysis was proposed to carry out the model manual and automatic calibration. Finally, the results were validated using satellite data referring to different periods and, also, the observed discharge at the basin outlet which was not used for calibration.

In general, the proposed model was able to reproduce properly the vegetation dynamics and the observed streamflow. The results highlighted the high usefulness of satellite data. It was possible to implement the hydrological and the vegetation components of TETIS-VEG daily model only using NDVI data and also to validate the model validation with satisfactory results. Such outcomes are promising because it demonstrates that satellite data could be exploited in order to predict river discharge in ungauged basins. More specifically, we do expect this result given the key role played by vegetation in water-controlled areas such as the upper Ewaso river basin in Kenya where having an appropriate description of vegetation and transpiration is critical for a correct description of the water balance at the local and basin scale.

The proposed automatic calibration was completely designed to incorporate spatio-temporal data in order – taking thus the maximum advantage of the available satellite data. After calibrating, the simulated vegetation patterns display good agreement with measured NDVI in most of the basin except for some portions at higher altitudes. This non-satisfactory result may be due to the bad quality of the NDVI data and/or the limitation of the vegetation sub-model (that was specifically designed for semiarid regions).

Model limitation along with poor data quality and resolution affected negatively the overall model performances, but the proposed procedure allowed to exploit the amount of information available addressing the critical issue of identifying a procedure for the calibration of a distributed model. This allowed to obtain a correct description of vegetation dynamics in space and time also providing as a marginal benefit a fairly good stream flow prediction. In this context, it was mandatory to adopt a daily time step in order to have a coherence with satellite NDVI data and also removing the need for a runoff propagation module in our model.

Finally, we should consider that the potential of the present study is due to the large availability of remote sensing information (not only satellite) concerning spatial state variables and more information will be available in the future. Many efforts are being done to improve the quality and quantity of remote sensing data (drones, better devices, etc.). And, the scientific community must also be ready to exploit the enormous amount of information contained in this data (temporal, spatial and spatio-temporal). Therefore, we should identify the best way to use all of this new available information, not only for data assimilation, but also for model calibration and validation.

Acknowledgements

The research leading to these results has received funding from the Spanish Ministry of Economy and Competitiveness and FEDER funds, through the research projects ECOTETIS (CGL2011-28776-C02-014) and TETISMED (CGL2014-58127-C3-3-R). The collaboration between Universitat Politècnica de València, Università degli studi della Basilicata and Princeton University was funded by the Spanish Ministry of Economy and Competitiveness through the EEBB-I-15-10262 fellowship. The MODIS data were obtained through the online Data Pool at the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota (https://lpdaac.usgs.gov/get_data).

References

Allen R. G., Pruitt W.O., Wright J.L., Howell T.A., Ventura F., Snyder R., Itenfisu D., Steduto P., Berengena J., Yrisarry J.B., Smith M., Pereira L.s., Raes D., Perrier A., Alves I., Walter I., Elliott R., 2006. A recommendation on standardized surface resistance for hourly calculation of reference ET0 by the FAO56 Penman-Monteith method. *Agricultural Water Management*, 81 (1-2), 1-22. doi: 10.1016/j.agwat.2005.03.007

Andersen, F.H. 2008. *Hydrological modeling in a semi-arid area using remote sensing data*. Doctoral Thesis, Department of Geography and Geology, University of Copenhagen (Denmark).

Bjornsson, H., Venegas SA., 1997. *A manual for EOF and SVD analyses of climate data.* CCGCR Rep. 97–1, McGill University, Montréal, QC, Canada, 52 pp.

Bonaccorso, B., Bordi, I., Cancelliere, A., Rossi, G., Sutera, A., 2003. Spatial variability of drought: an analysis of the SPI in Sicily. *Water resources management*, 17(4), 273-296.

Bond, B. J., Jones, J. A., Moore, G., Phillips, N., Post, D., and McDonnell, J. J. 2002. The zone of vegetation influence on baseflow revealed by diel patterns of streamflow and vegetation water use in a headwater basin. *Hydrological Processes*, 16(8), 1671-1677.

Brown, B. G., Gilleland, E., Ebert, E.E. 2011. Forecasts of spatial fields, in Forecast Verification, edited, pp. 95–117, John Wiley.

Caylor, K.K., D'Odorico P., Rodriguez-Iturbe I., 2006. On the ecohydrological organization of spatially heterogeneous semi-arid landscapes. *Water Resources Research*, 42(7), W07424, doi:10.1029/ 2005WR004683.

Ceballos Y, Ruiz de la Torre J. 1979. *Árboles y arbustos de la España peninsular*. ETSI Montes Publications: Madrid.

Chen, J. M., & Cihlar, J.,1996. Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sensing of Environment*, 55, 153 – 162.

Cheng, L., Yaeger, M., Viglione, A., Coopersmith, E., Ye, S., and Sivapalan, M., 2012. Exploring the physical controls of regional patterns of flow duration curves – Part 1: Insights from statistical analyses, Hydrol. Earth Syst. Sci., 16, 4435-4446, doi:10.5194/hess-16-4435-2012.

Cohen, W. B., Maiersperger, T. K., Gower, S. T., & Turner, D. P., 2003. An improved strategy for regression of biophysical variables and Landsat ETM+data. *Remote Sensing of Environment*, 84, 561 – 571.

Cohen J. 1968. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin,* 70(4), 213–220. DOI: 10.1037/h0026256.

Conradt, T., Wechsung, F., and Bronstert, A., 2013. Three perceptions of the evapotranspiration landscape: comparing spatial patterns from a distributed hydrological model, remotely sensed surface temperatures, and sub-basin water balances, *Hydrol. Earth Syst. Sci*., 17, 2947-2966, doi:10.5194/hess-17-2947-2013.

Coopersmith, E., Yaeger, M. A., Ye, S., Cheng, L., & Sivapalan, M., 2012. Exploring the physical controls of regional patterns of flow duration curves–Part 3: A catchment classification system based on regime curve indicators. Hydrology and Earth System Sciences, 16(11), 4467-4482.

Drewry D.T., Albertson J.D., 2006. Diagnosing model error in canopy-atmosphere exchange using empirical orthogonal function analysis. Water Resources Research, 42, W06421, doi: 10.1029/2005WR004496

Fang Z., Bogena H., Kollet S., Koch J., Vereecken H., 2015. Spatio-temporal validation of long-term 3D hydrological simulations of a forested catchment using empirical orthogonal functions and wavelet coherence analysis. *Journal of Hydrology*, 529, 1754-1767

Feyen, L., Kalas, M., and Vrugt, J. A., 2008. Semi-distributed parameter optimization and uncertainty assessment for large-scale stream-flow simulation using global optimization, Hydrolog. Sci. J., 53, 293–308, doi:10.1623/hysj.53.2.293

Francés F., Benito J., 1995. La modelación distribuida con pocos parámetros de las crecidas. Ingeniería del Agua, 2, 7-24

Francés F., Vélez J. I., Vélez J. J., 2007. Split.-parameter structure for the automatic calibration of distributed hydrological models. *Journal of Hydrology*, 332:226-240

Franssen H.J.H, P. Brunner, P. Makobo, W. Kinzelbach, 2008. Equally likely inverse solutions to a groundwater flow problem including pattern information from remote sensing images. *Water Resour. Res*., 44 (1), p. W01419

Franz T.E, 2007. *Ecohydrology of the upper Ewaso Ngiro river basin, Kenia*. Doctoral Thesis. Princeton University, Princeton, NJ, USA.

Franz TE, Caylor KK, Nordbotten JM, Rodríguez-Iturbe I., Celia MA., 2010. An ecohydrological approach to predicting regional woody species distribution patterns in dryland ecosystems. *Advances in Water Resources*, 33(2), 215-230

Frassnacht, K. S., Gower, S. T., MacKenzie, M. D., Nordheim, E. V., & Lillesand, T. M., 1997. Estimating the leaf area index of north central Wisconsin forests using the Landsat Thematic Mapper. *Remote Sensing of Environment*, 61, 229 – 245.

Friedl, M. A., Michaelsen, J., Davis, F. W., Walker, H., & Schimel, D. S., 1994. Estimating grassland biomass and leaf area index using ground and satellite data. *International Journal of Remote Sensing*, 15, 1401 – 1420.

Gamon, J.A., Serrano, L., Surfus, J.S., 1997. The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia*, 112, 492-501

García-Arias, A., Francés, F., 2016. The RVDM: modelling impacts, evolution and competition processes to determine riparian vegetation dynamics. *Ecohydrology*, 9, 438–459. DOI: 10.1002/eco.1648

Gigante, V., Milella, p., Iacobellis, V., Manfreda, S., Portoghese, I., 2009. Influences of Leaf Area Index estimations on the soil water balance predictions in Mediterranean regions, Natural Hazard and Earth System Sciences, 9, 979-991, doi: 10,5194/nhess-9-979-2009

GIMHA team (Vélez I., Vélez J., Puricelli M., Montoya JJ., Camilo JC., Bussi G., Medici C., Orozco I., Ruiz-Pérez G.), 2014. *Description of the distributed conceptual hydrological model TETIS v.8*. Universitat Politècnica de València.

Gilleland, E., Ahijevych, D. A., Brown, B. G., Ebert, E.E., 2010. Verifying forecasts spatially, *Bull. Am. Meteorol. Soc.*, 91(10), 1365–1373.

Graf A., Bogena H.R., Drüe C., Hardelauf H., Pütz T., Heinemann G., Vereecken H., 2014. Spatiotemporal relations between water budget components and soil moisture in a forested tributary catchment. *Water Resources Research*, 50(6), 4837-4857. doi: 10.1002/2013WR014516

Gribovszki, Z., Kalicz, P., Szilágyi, J., Kucsara, M., 2008. Riparian zone evapotranspiration estimation from diurnal groundwater level fluctuations, *J. Hydrol.*, 349, 6–17, doi:10.1016/j.jhydrol.2007.10.049.

Gutmann E.D., Small E.E., 2010. A method for the determination of the hydraulic properties of soil from MODIS surface temperature for use in land-surface models. *Water Resources Research*, 46, W06520, doi: 10.1029/2009WR008203

Immerzel, W., Droogers, P., 2008. Calibration of a distributed hydrological model based on satellite evapotranspiration. *Journal of Hydrology*, 349, 411-424

Jasechko S., Sharp Z.D., Gibson J.J., Birks S.J., Yi Y., Fawcett P.J., 2013. Terrestrial water fluxes dominated by transpiration. *Nature*, 496, 347-350, doi: 10.1038/nature11983

Kim G., Barros A.P., 2002. Space-time characterization of soil moisture from passive microwave remotely sensed imagery and ancillary data. *Remote Sensing of Environment*, 81, 393-403

Koch J., Jensen K.H., Stisen S., 2015. Toward a true spatial model evaluation in distributed hydrological modeling: Kappa statistics, Fuzzy theroy, and EOF-analysis benchmarked by the human perception and evaluated against a modeling case study. *Water Resources Research*, 51, 1225-1246, doi: 10.1002/2014WR016607

Koch J., Cornelissen T., Fang Z., Bogena H., Diekkrüger B., Kollet S., Stisen S., 2016a. Inter-comparison of three distributed hydrological models with respect to seasonal variability of soil moisture patterns at a small forested catchment. *Journal of Hydrology*, 533, 234-249.

Koch, J., Siemann, A., Stisen, S., Sheffield, J., 2016b. Spatial validation of large-scale land surface models against monthly land surface temperature patterns using innovative performance metrics. *Journal of Geophysical Research: Atmospheres*, 121(10), 5430-5452. DOI: 10.1002/2015JD024482

Kunnath-Poovakka A., Ryu D., Renzullo L.J., George B., 2016. The efficacy of calibrating hydrologic model using remotely sensed evapotranspiration and soil moisture for streamflow prediction. *Journal of hydrology*, 535, 509-524

Landsberg, J.J; Waring, R.H., 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partiotioning. *Forest Ecolological Management*, 95, 209-228

Law, B. E., Waring, R. H., 1994. Remote sensing of leaf area index and radiation intercepted by understory vegetation. *Ecological Applications*, 4, 272 – 279.

Le Roux X., Bariac t., Mariotti A., 1995. Spatial partitioning of the soil water resource between grass and shrub components in a West African humid savanna. Oecologia, 104, 147-155.

Liu Y., 2003. Spatial patterns of soil moisture connected to monthly-seasonal precipitation variability in a monsoon region. *Journal of Geophysical Resources*, 108 (D22), 8856. doi: 10.1029/2002JD003124

Lo, M.H.; Famiglietti, J.S.; Yeh, P.J; Syed, T.H., 2010. Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data. *Water Resources Research* (46), W05517. doi:10.1029/2009WR007855

López-Serrano, F.R., Landete-Castillejos, T., Martínez-Millán, J., Cerro-Barja, A., 2000. LAI estimation of natural pine forest using a non-standard sampling technique. *Agricultural and Forest Meteorology*, 101, 95–111. DOI: 10.1016/S0168-1923(99)00171-9

Manfreda, S., Caylor, K.K., 2013. On The Vulnerability of Water Limited Ecosystems to Climate Change, *Water*, 5(2), 819-833

Manfreda S., Pizzolla T., Caylor K.K., 2013. Modelling vegetation patterns in semiarid environments. *Procedia Environmental Sciences*, 19, 168-177

Manfreda, S., Fiorentino, M., Iacobellis, V., 2005. DREAM: a Distributed model for Runoff, Evapotranspiration and Anteceden Soil Moisture Simulation. *Advances in Geosciences*, 2, 31-39, SRef-ID: 1680-7359/adgeo/2005-2-31

McCabe M.F., E.F. Wood, R. Wojcik, M. Pan, J. Sheffield, H. Gao, H. Su, 2008.Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies, *Remote Sens. Environ.*, 112 (2), 430–444

Medlyn, B.E., 1998. Physiological basis of the light use efficiency model. *Tree Physiology*, 18, 167-176.

Merz, R., Parajka, J., and Bloschl, G., 2009. Scale effects in conceptual hydrological modeling, *Water Resour. Res.,* 45, W09405, doi:10.1029/2009WR007872.

Michaud, J., & Sorooshian, S., 1994. Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed. *Water Resources Research,* 30(3), 593-605.

Montaldo N., Rondena R., Albertson J.D., Mancini M., 2005. Parsimonious modeling of vegetation dynamics for ecohydrologic studies of water limited ecosystems. *Water Resources Research*, 41, W10416, doi: 10.1029/2005WR004094

Nash J.E. and Sutcliffe J.V., 1970. River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, 10 (3), 282-290

Nyholm, T., Rasmussen, K. R., & Christensen, S. 2003. Estimation of stream flow depletion and uncertainty from discharge measurements in a small alluvial stream. *Journal of Hydrology*, 274(1), 129-144.

Ollinger, S.V.; Richardson, A.D.; Martin, M.E.; Hollinger, D.Y.; Frolking, S.; Reich, P.B.; Plourde, L.C.; Katul, G.; Munger, J.W.; Oren, R.; Smith, M.L.; Paw, U.; Bolstad, K.T.; Cook, P.V.; Day, B.; Martin, M.C.; Monson, T.A.; Schmidt, R.K.H.P., 2008. Canopynitrogen, carbon assimilation, and albedo in temperate and boreal forests: functional relations and potential climate feedbacks. *Proc. Natl. Acad. Sci. USA.*, 105, 19335-19340

Pasquato, M.; Medici, M.; Friend A.D.; Francés, F., 2015. Comparing two approaches for parsimonious vegetation modelling in semiarid regions using satellite data. *Ecohydrology*, 8(6), 1024-1036. DOI: 10.1002/eco.1559

Perone, C. S. 2017. http://pyevolve.sourceforge.net/ (on line access on 2017).

Porporato A., Laio F., Ridolfi L., Rodriguez-Iturbe I., 2001. Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress: III. Vegetation water stress. *Advances in Water Resources*, 24(7), 725-744.

Pumo, D., Noto, L. V., Viola, F., 2013. Ecohydrological modelling of flow duration curve in Mediterranean river basins. *Advances in Water Resources*, 52, 314-327.

Preisendorfer, R. W., Mobbley C.D., 1988. *Principal component analysis in meteorology and oceanography* (Vol. 425). (Ed.). Amsterdam: Elsevier

Quevedo, D.I., Francés, F., 2008. A conceptual dynamic vegetation-soil model for arid and semiarid zones. *Hydrol. Earth Syst. Sci.* 12, 1175-1187.

Rientjes, T.H.M., Muthuwatta, L.P., Bos, M. G., Booij, M.J., Bhatti, H.A., 2013. Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration. *Journal of Hydrology*, 505, 276-290

Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, D.-J. Seo, DMIP participants, 2004. Overall distributed model inter-comparison project results. *Journal of Hydrology*, 298(1–4), 27–60

Rodriguez-Iturbe, I.; Porporato, A.; Laio, F; Ridolfi, L. 2001. Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress. I. Scope and general outline. *Advances in Water Resources*, 24, 695–705

Ruiz-Pérez G., González-Sanchis M., Del Campo AD., Francés F., 2016. Can a parsimonious model implemented with satellite data be used for modelling the vegetation dynamics and water cycle in water-controlled environments?. *Ecological modelling*, 324, 45-53

Samaniego L., Kumar R., Jackisch C., 2011. Predictions in a data-sparse region using a regionalized grid-based hydrologic model driven by remotely sensed data. *Hydrology Research*, 42(5), 338-355, doi: 10.2166/nh.2011.156

Sims, D.A.; Luo, H.; Hastings, S.; Oechel, W.C.; Rahman, A.F.; Gamon, J.A., 2006. Parallel adjustments in vegetation greenness and ecosystem CO_2 exchange in response to drought in a Southern California chaparral ecosystem. *Remote Sensing of Environment*, 103, 289-303

Smith, M. B., Koren, V., Reed, S., Zhang, Z., Zhang, Y., Moreda, F.,Cui, Z., Mizukami, N., Anderson, E. A., and Cosgrove, B. A., 2012. The distributed model intercomparison project – Phase 2: Motivation and design of the Oklahoma experiments, *J. Hydrol.*, 418–419, 3–16, doi:10.1016/j.jhydrol.2011.08.055.

Smith, M.B., Koren, V., and a total of 32 coauthors, 2013. The Distributed Model Intercomparison project – Phase 2: Experiment design and summary results of the western basin experiments. *J.Hydrol.,* 207, 300-329. DOI:10.1016/j.hydrol.2013.08.040

Stisen, S., McCabe, M.F., Refsgaard, J.C., Lerer, S. and Butts, M.B., 2011. Model parameter analysis using remotely sensed pattern information in a multi-constraint framework. *Journal of Hydrology*, 409(1), 337-349.

Tsang, Y. P., Hornberger, G., Kaplan, L. A., Newbold, J. D., and Aufdenkampe, A. K. 2014. A variable source area for groundwater evapotranspiration: Impacts on modeling stream flow. *Hydrological Processes*, 28(4), 2439-2450.

Turner, D. P., Cohen, W. B., Kennedy, R. E., Fassnacht, K. S., & Briggs, J. M., 1999. Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, 70, 52 – 68.

van Dijk A.I.J.M., Renzullo L.J., 2011. Water resource monitoring systems and the role of satellite observations. *Hydrology and Earth system Sciences*, 15, 39-55. doi: 10.5194/hess-15-39-2011

Velpuri N.M., Senay G.B., Asante K.O., 2012. A multi-source satellite data approach for modelling Lake Turkana water level: calibration and validation using satellite altimetry data. *Hydrology and Earth System Sciences*, 16, 1-8, doi: 10.5194/hess-16-1-2012

Wagener, T., Blöschl, G., Goodrich, D. C., Gupta, H., Sivapalan, M., Tachikawa, Y., Weiler, M. 2013. A synthesis framework for runoff prediction in ungauged basins. *chapt, 2*, 11-28.

Wi S., Yang Y.C.E., Steinschneider S., Khalil A., Brown C.M., 2015. Calibration approaches for distributed hydrologic models in poorly gaged basins: implication for streamflow projections under climate change. *Hydrology and Earth System Sciences*, 19, 857-876

Winsemius, H.C., Savenije, H.H.G., Bastiaansen, W.G.M., 2008. Constraining model parameters on remote sensed evaporation: justificacion for distribution in ungauged basins? *Hydrology Earth System Science*, 12, 1403-1413.

Xiao X., Zhang Q., Braswell B., Urbanski S., Boles S., Wofsy S., Moore B., Ojima D., 2004. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. *Remote Sensing of Environment*, 91(2), 256-270

Yang, Y., Shang, S., Jiang, L., 2012. Remote sensing temporal and spatial patterns of evapotranspiration and the responses to water management in a large irrigation district of North China. *Agricultural and forest meteorology*, 164, 112-122.

Yuan W, Liu S, Zhou G, Zhou G, Tieszen LL, Baldocchi D, Bernhofer C, Gholz H, Goldstein AH, Goulden ML, Hollinger DY, Hu Y, Law BE, Stoy PC, Vesala T, Wofsy SC, other AmeriFlux collaborators. 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agricultural and Forest Meteorology*, 143, 189–207. DOI: 10.1016/j.agrformet.2006.12.001

Yuan W., Liu S., Yu G., Bonnefond J-M, Chen J., Davis K., Desai AR., Goldstein A.H., Gianelle D., Rossi F., Suyker AE., Verma SB., 2010. Global estimates of evapotranspiration and gross primary production based on MODIS and global meteorology data. *Remote Sensing of Environment*, 114(7), 1416-1431

Zhang Y.Q., Francis H.S., Chiew H.S., Zhang L., Li H., 2009. Use of remotely sensed actual evapotranspiration to improve rainfall-runoff modeling in southeast Australia. *American Meteorological Society*, 10, 969-980

Zhang Y.Q., Viney N.R., Chiew F.H.S., van Dijk A.I.J.M., Liu Y.Y., 2011. *Improving hydrological and vegetation modelling using regional model calibration schemes together with remote sensing data.* 19th International Congress on Modelling and Simulation, Perth, Australia, 12-16 December

Zhang, Y., Peña-Arancibia, J. L., McVicar, T. R., Chiew, F. H., Vaze, J., Liu, C., Lu, X., Zheng H., Wang, Y., Liu, Y., Miralles, D. G., 2016. Multi-decadal trends in global terrestrial evapotranspiration and its components. *Scientific reports*, 6. doi:10.1038/srep19124

Table I. Summary of the initial values, the search range and the final value of the parameters or correction factors of both sub-models (hydrological and dynamic vegetation sub-models) as well as the units and the reviewed references.

Model	Correction factor or parameter[*]	Units		Initial Value	Search Range	Final Value	References
HYDROLOGICAL SUB-MODEL	FC1-Maximum Static Storage	[-]		1.00	[0.5,2.5]	1.80	[1]
	FC2-Evapotranspiration	[-]		0.70	[0.7,1.2]	1.05	[1]
	FC3-Infiltration	[-]		0.20	[0.01,2]	0.12	[1]
	FC4-Slope velocity	[-]		1.00	[0.1,1.2]	1.00	[1]
	FC5-Percolation	[-]		0.08	[0.001,2]	0.05	[1]
	FC6-Interflow	[-]		140.00	[0.001,100000]	150.12	[1]
	FC7-Deep percolation	[-]		0.06	[0.001,2]	0.04	[1]
	FC8-Connected aquifer	[-]		20.00	[0.001,100000]	16.82	[1]
	FC9-Flow velocity	[-]		1.00	[0.2,1.2]	1.00	[1]
VEGETATION SUB-MODEL	Specific Leaf Storages	mm	Tree	0.50	[0.5,3]	0.43	[2],[3],[4]
			Shrub	2.00	[0.5,3]	2.00	
			Grass	2.00	[0.5,3]	2.00	
	LUE	kg/m^2 MJ	Tree	1.50	[1.2,2.5]	1.14	[5],[6]
			Shrub	1.50	[1.2,2.5]	1.14	
			Grass	1.50	[1.2,2.5]	1.71	
	Coverage factor	[-]	(**)	0.80	[0.1,1.0]	0.90	[3],[4]
	Distribution of roots	[-]	Tree	0.30	[0.0,1.0]	0.10	[3],[4],[7]
			Shrub	0.5	[0.0,1.0]	0.20	
			Grass	0.7	[0.0,1.0]	0.34	
	Maximum LAI	m^2/m^2	Tree	2.50	[0.5,3.5]	3.10	[5],[8],[9],[10]
			Shrub	2.00	[0.5,3.5]	2.00	
			Grass	1.00	[0.5,3.5]	1.50	

Light extinction coefficient	[-]	All	0.50	[0.4,0.6]	0.52	[11]
SLA	m^2/kg	Tree	4.00	[2.0,5.0]	4.00	
		Shrub	6.00	[4.0,20.0]	10.00	[5],[12]
		Grass	6.00	[6.0,50.0]	30.00	
Optimal temperature	°C	All	16	[10,30]	18	[11]

(*) Regarding to the hydrological sub-model, the table shows the value of the correction factors while regarding to the vegetation sub-model, the table shows the parameter values

(**) The coverage factor depends on the location. The value in the table is the mean value. We used the reported information by [3] and [4].

[1] GIMHA Team, 2014

[2] Van Dijk *et al.,* 2011

[3] Franz *et al.,* 2007

[4] Caylor *et al.,* 2006

[5] TRY Database (www.try-db.org)

[6] Yuan *et al.,* 2007

[7] Le Roux *et al.,* 1995

[8] Pasquato *et al*., 2015

[9] Ceballos and Ruiz de la Torre, 1979

[10] López-Serrano *et al.,* 2000

[11] Ruiz-Pérez *et al.,* 2016

[12] Castro de Costa *et al.,* 2014

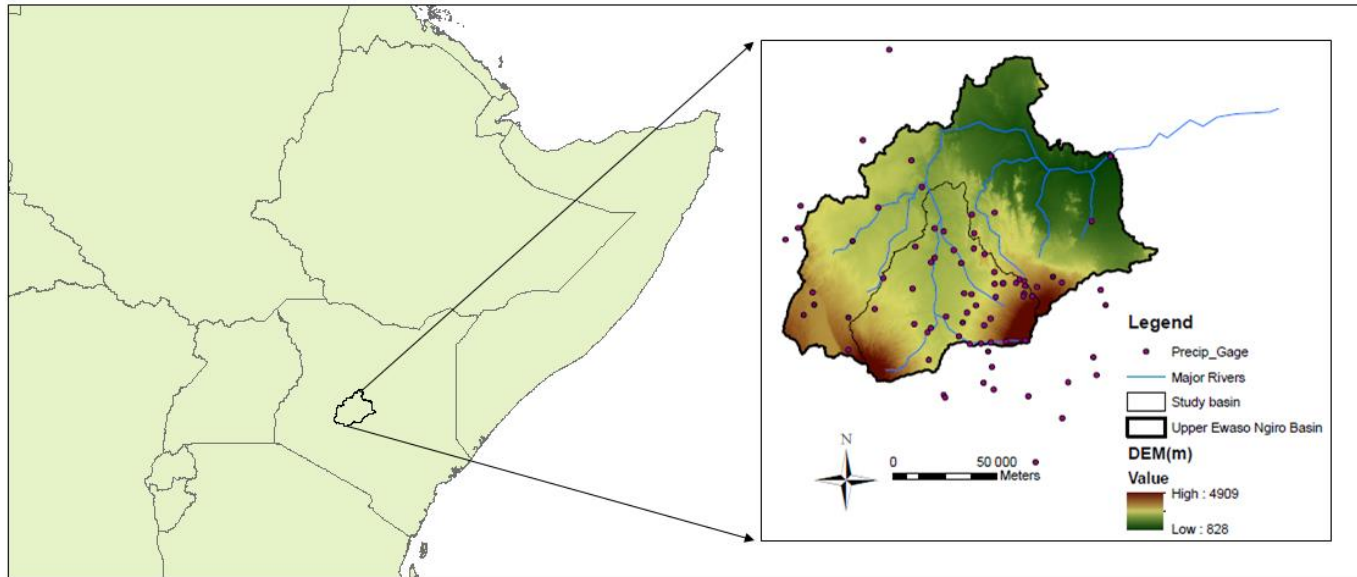



Figure 1. General map location of the Upper Ewaso Ngiro river basin within the boundaries of the East Sub-saharan Africa. The study sub-catchment (study basin) was selected because the density of the rainfall stations (points in dark red).

Manual Calibration Results

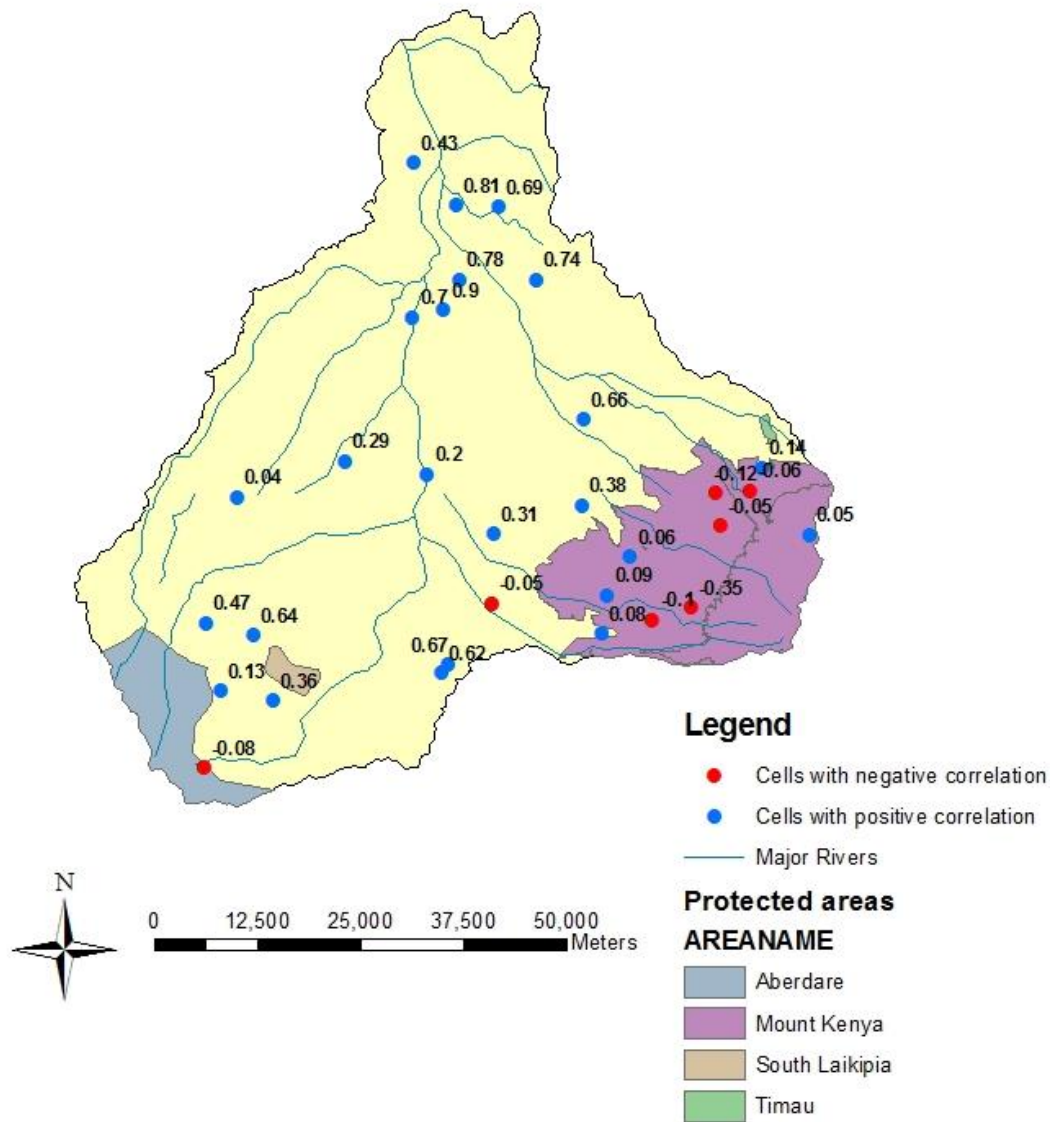



Figure 2. Location of the points where the manual calibration was carried out. The value of the Pearson correlation coefficient between the satellite NDVI and the simulated LAI appears together to the point used to the manual calibration of the model

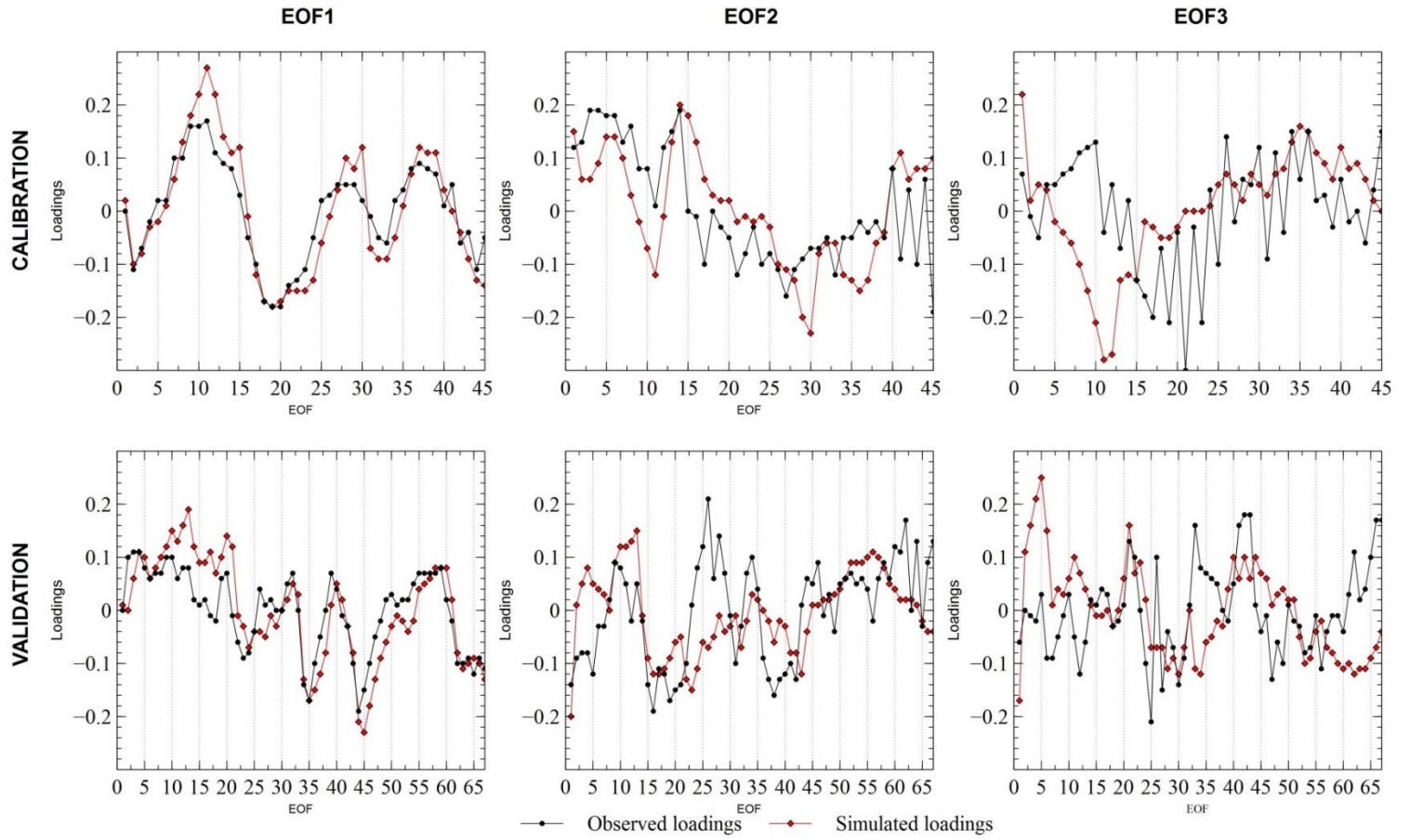



Figure 3. The three first EOFs during the calibration (upper part) and during the validation (lower part) are represented. The y-axes reflect the unitless loadings of each EOF. The x-axes reflect the time step.

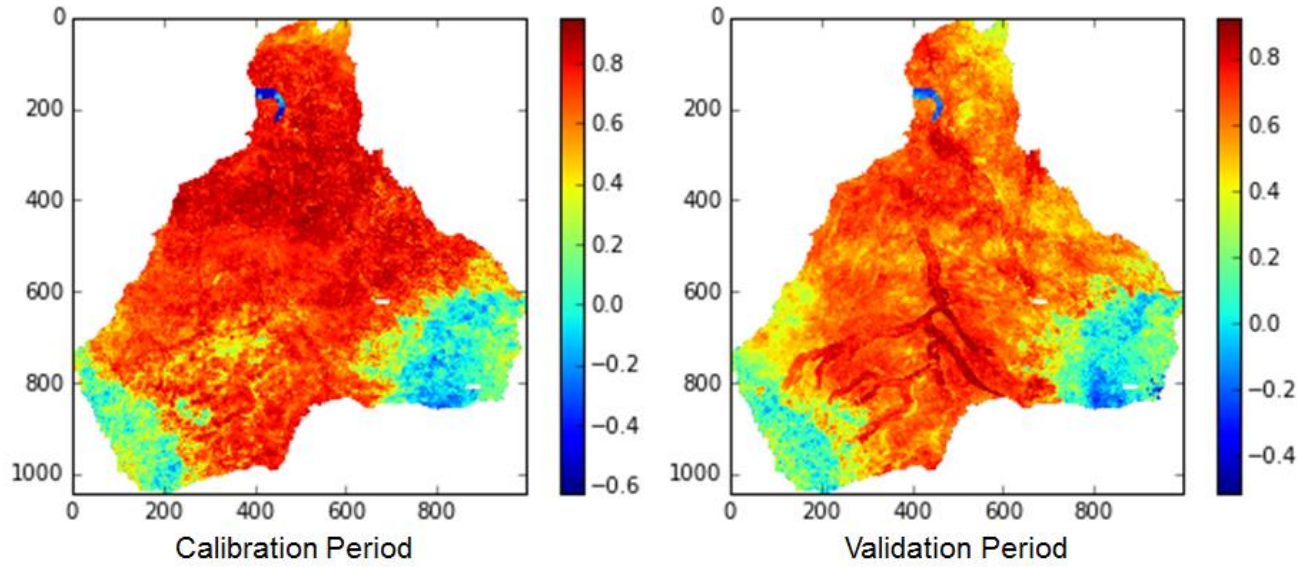



Figure 4. Temporal Pearson correlation coefficient between the NDVI provided by MODIS and the LAI simulated by the model during the calibration and validation period. The two areas with negative values correspond to the Mount Kenya and Aberdare Mountains.

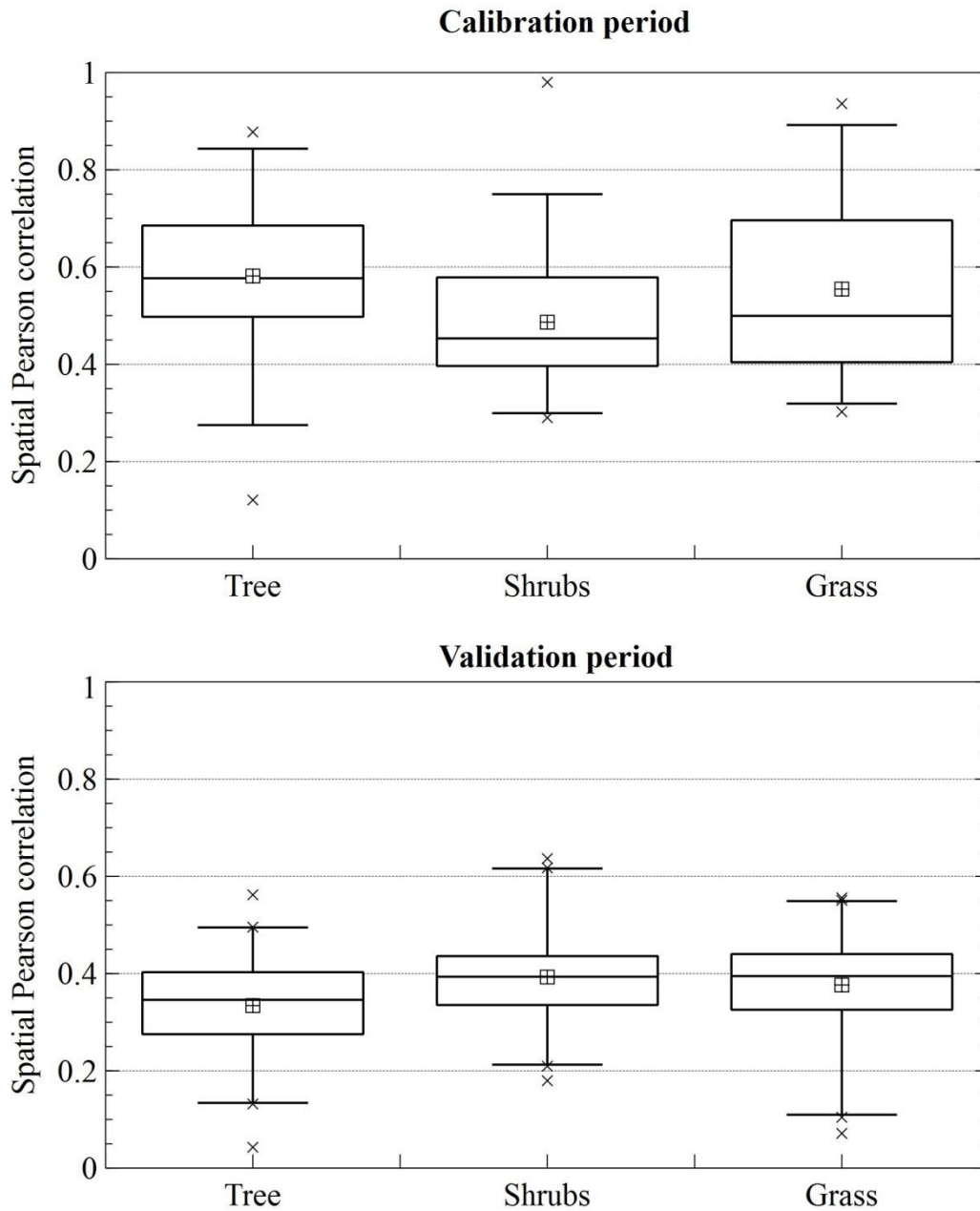



Figure 5. Spatial Pearson correlation coefficient during the calibration (upper panel) and during the validation (lower panel) distinguishing between the main land covers: tree, shrubs and grass. The whiskers were calculated according to the 98% percentile and the outliers were plotted as x. The median is the line inside boxplot and the mean is the quadrangle.

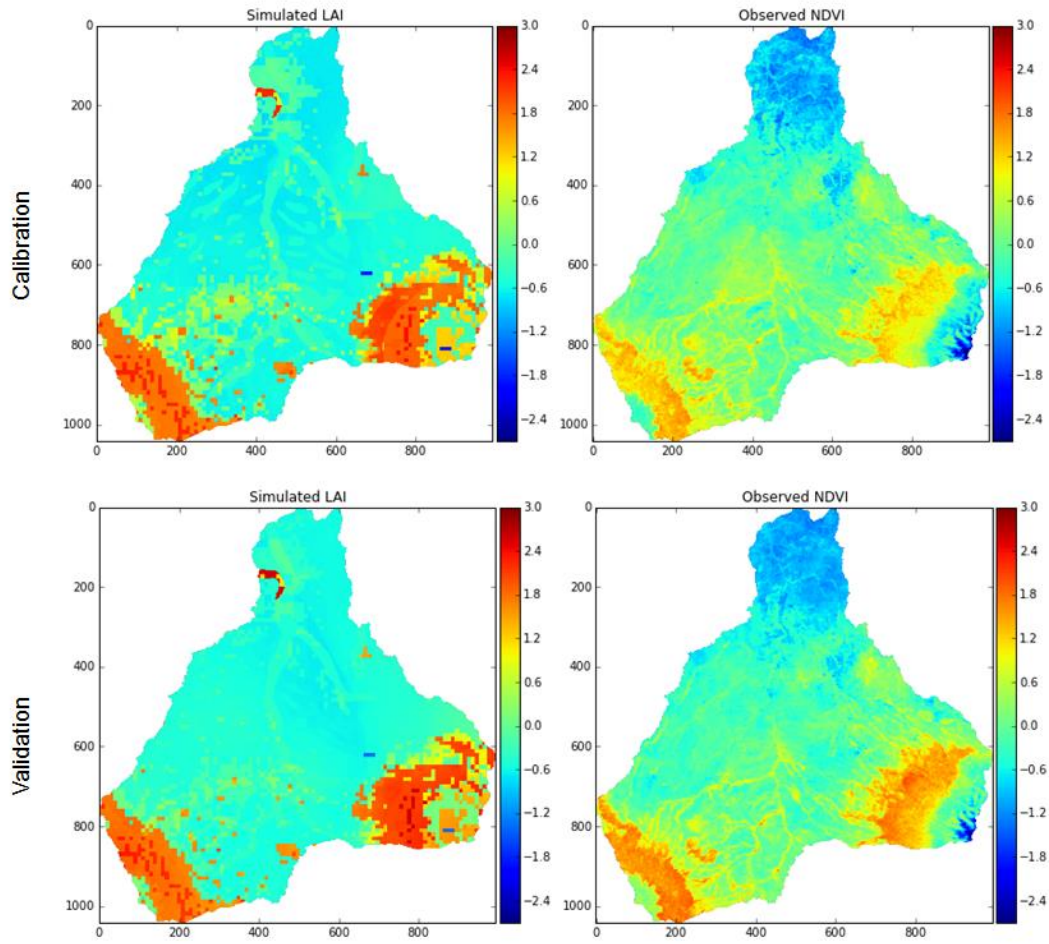



Figure 6. Comparison between the maps where each pixel contains the difference between the temporal mean calculated in this particular pixel and the general mean calculated using the all dataset of the simulated LAI (left) and observed NDVI (right) in both periods: calibration (upper panels) and validation (lower panels). This difference is a measure of spatial gradient of both variables (LAI and NDVI).

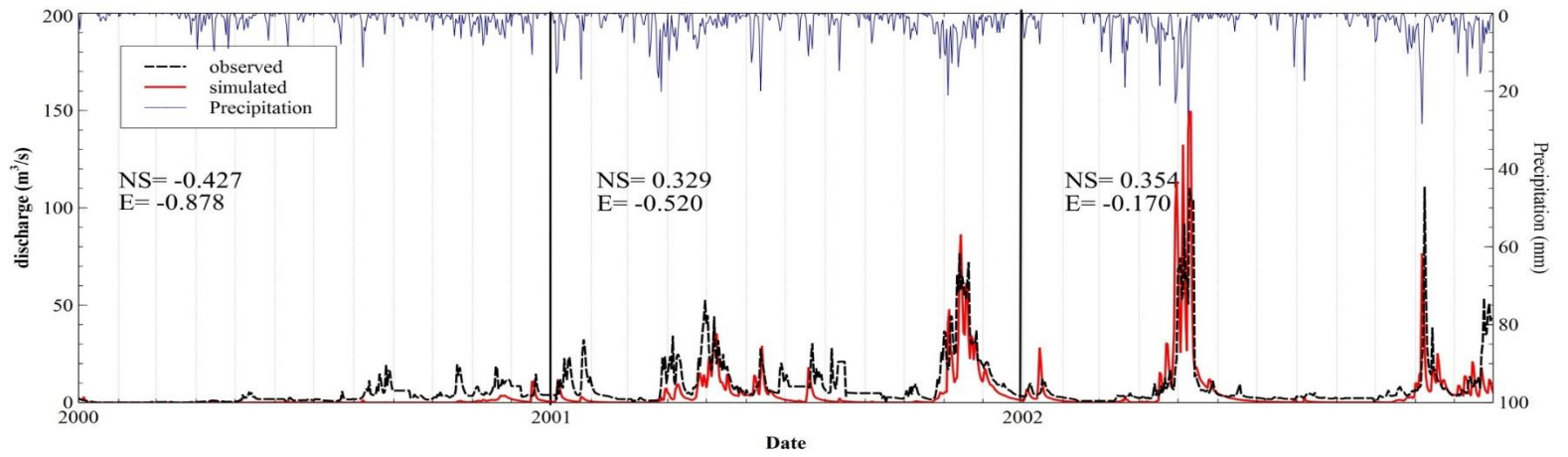
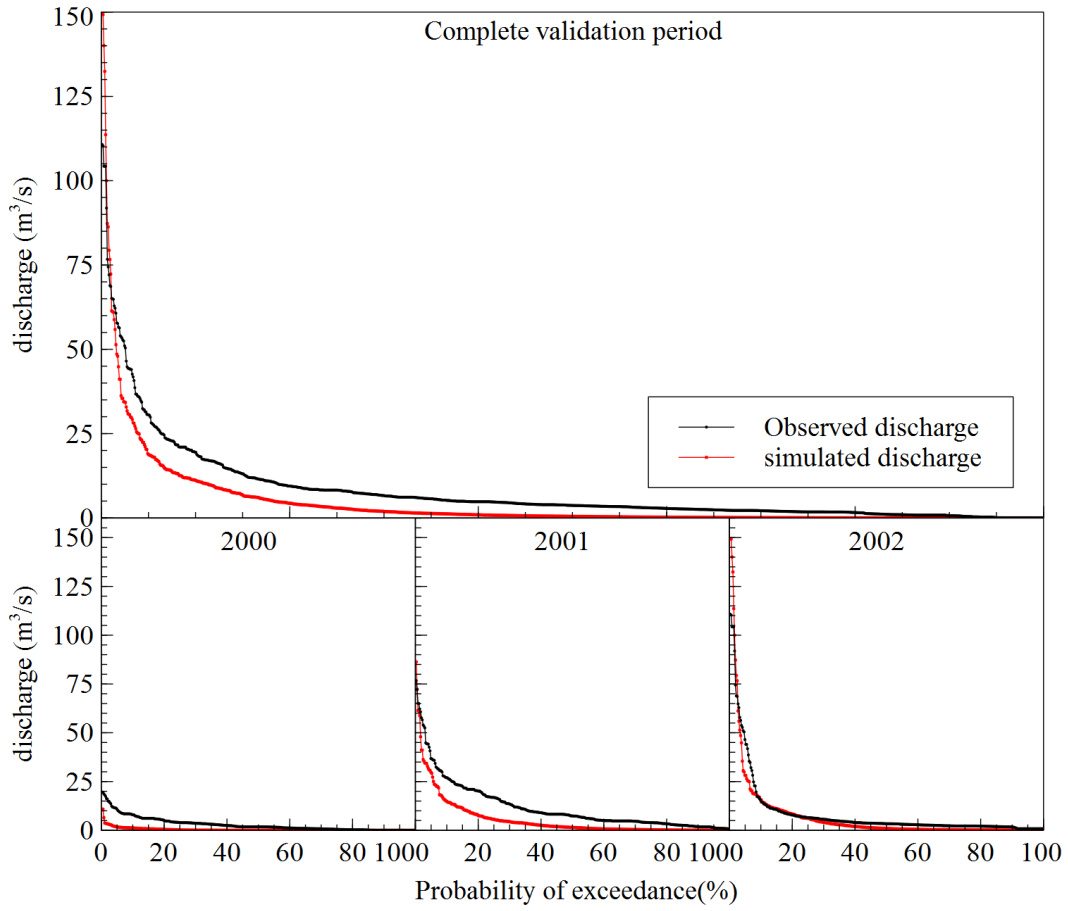



Figure 7. Time series of rainfall and observed and simulated daily discharge (m³/s) during the validation period (2000,2001 and 2002)

26   **Figure 8. Observed (in black) and simulated (in red) flow duration curves for the whole validation period (upper**
27   **panel) and for the corresponding three years in isolation (lower three panels): 2000, 2001 and 2002 respectively**