# Spatial pattern evaluation of a calibrated national hydrological model – a remote sensing based diagnostic approach

Gorka Mendiguren [1], Julian Koch [1], Simon Stisen [1]

5   [1]Department of hydrology, Geological Survey of Denmark and Greenland, Copenhagen, Denmark.

*Correspondence to*: Gorka Mendiguren (gmg@geus.dk)

**Abstract.** Distributed hydrological models are traditionally evaluated against discharge stations, emphasizing the temporal and neglecting the spatial component of a model. The present study widens the traditional paradigm by highlighting spatial patterns of evapotranspiration (ET), a key variable at the land-atmosphere interface, obtained from two different approaches

10   at the national scale of Denmark. The first approach is based on a national water resources model (DK-model), using the MIKE-SHE model code, and the second approach utilizes a two source energy balance model (TSEB) driven mainly by satellite remote sensing data.

The main hypothesis of the study is that while both approaches are essentially estimates, the spatial patterns of the remote sensing based approach are explicitly driven by  observed land surface temperature and therefore represent the most direct

15   spatial pattern information of ET; enabling its use for distributed hydrological model evaluation. Ideally the hydrological model simulation and remote sensing based approach should present similar spatial patterns and driving mechanism of ET. However, the spatial comparison showed that the differences are significant and indicating insufficient spatial pattern performance of the hydrological model.

The differences in spatial patterns can partly be explained by the fact that the hydrological model is configured to run in 6

20   domains that are calibrated independently from each other, as it is often the case for large scale multi-basin calibrations. Furthermore, the model incorporates predefined temporal dynamics of Leaf Area Index (LAI), root depth (RD) and Crop coefficient (Kc) for each land cover type. This zonal approach of model parametrization ignores the spatio-temporal complexity of the natural system. To overcome this limitation, the study features a modified version of the DK-Model in which LAI, RD, and KC are empirically derived using remote sensing data and detailed soil property maps in order to

25   generate a higher degree of spatio-temporal variability and spatial consistency between the 6 domains. The effects of these changes are analyzed by using the empirical orthogonal functions (EOF) analysis to evaluate spatial patterns. The EOF-analysis shows that including remote sensing derived LAI, RD and KC in the distributed hydrological model adds spatial features found in the spatial pattern of remote sensing based ET.

## 1 Introduction

The application of spatially distributed hydrological models has become common practice for a wide range of water resources assessments. Such models are valuable tools to manage terrestrial water resources and to provide insights into the overall water balance as well as the internal distribution of multiple hydrological states and fluxes. The spatial predictability of these models is however severely hampered by the general lack of suitable spatial pattern oriented model evaluation frameworks since evaluation remains focused on spatially aggregated objective functions such as discharge. As stated by Conradt et al. (2013) "In conjunction with distributed hydrological modelling spatial calibration usually means individual multi-site calibration". The neglect of a specific focus on spatial patterns in model evaluation is a paradox in the light of an increasing acknowledgement of the role of patterns in the functioning of hydrological systems (Vereecken et al., 2016). Moreover it is against the rational of developing and applying distributed models (Freeze and Harlan, 1969;Refsgaard, 1997). If the spatial variability of a hydrological system is not of importance to the modeler it seems not worth the effort to apply a distributed model, since numerous studies indicate that equal model fidelity can be achieved with a lumped approach when evaluated solely at the catchment outlet (Stisen et al., 2011;Vansteenkiste et al., 2014)

The concept of spatial pattern comparisons in catchment hydrology was pioneered by Grayson and Blöschl (2000), who developed the theoretical framework and terminology. Since then significant progress has been made in areas such as model code development (Clark et al., 2015;Maxwell and Kollet, 2008;Samaniego et al., 2010), remote sensing (Lettenmaier et al., 2015) and data assimilation (Zhang et al., 2016;Ridler et al., 2014). Nevertheless, explicit spatial pattern evaluation of distributed hydrological models remains a rarity.

In order to perform qualified assessments of simulated spatial patterns reliable observations are a prerequisite. For this purpose satellite remote sensing comes into play as an independent data source with the required spatial resolution and coverage for many catchment scale applications. Satellite imagery has been used for estimation of numerous states and fluxes of interest to hydrological modelling, such as snow cover (Immerzeel et al., 2009), storage change (Chen et al., 2016), soil moisture (SM) (Wanders et al., 2014), vegetation water content (Mendiguren et al., 2015), land surface temperature (LST) (Corbari et al., 2015) and actual evapotranspiration (ET) (Guzinski et al., 2015). The conversions of the remotely sensed signal to hydrological variables is far from trivial, and usually require in situ measurements and observations for model evaluation. However, in spite of their overall uncertainty; satellite based estimates contain valuable pattern information (Mascaro et al., 2015). More precisely, we propose to utilize remote sensing data solely for the purpose of pattern validation, using bias insensitive metrics, and leave the task to validate water balance closure to the more trusted discharge observations.

Several studies have explored the obvious potential in utilizing satellite estimates for hydrological model evaluation and calibration. Some utilize time series of a basin average observation from remote sensing to guide model calibration (Rajib et al., 2016;Rientjes et al., 2013) and therefore rely heavily on the accuracy of the remote sensing estimate while neglecting the spatial pattern information. Others utilize the satellite based estimates to perform a pixel to pixel calibration of the model

(Corbari and Mancini, 2014). The latter approach might explore the full information content of the observations. However, there is an eminent risk of a highly parameterized solution problem and possibly unrealistic spatial parameter distributions in light of the uncertainty of the remote sensing estimates at the pixel level.

Relatively few studies utilize the actual pattern of remote sensing estimates in distributed hydrological model evaluation. Interesting examples are (Li et al., 2009) who used remote sensing derived patterns of actual evapotranspiration, and Hendricks Franssen et al. (2008) who utilized satellite based recharge patterns to constrain the calibration of groundwater models. Immerzeel and Droogers (2008) included actual evapotranspiration estimates in the calibration of the Krishna basin in southern India and obtained a good correlation across 115 sub basins while Githui et al. (2012) successfully applied a multi-objective calibration by combining river discharge and remotely sensed actual evapotranspiration of 59 sub basins in Victoria, Australia. Other pattern oriented model evaluations without calibration were conducted by Bertoldi et al. (2010), Wang et al. (2009) and (Koch et al., 2016); the latter applied different spatial performance metrics in the evaluation of three land surface models over the continental United States based on remotely sensed land surface temperature maps.

The aim of this study is to develop a remote sensing based actual evapotranspiration (AET) dataset suitable for validation and calibration of the spatial patterns of the national hydrological model of Denmark (DK-model). The model evaluation will be based on a diagnostic approach inspired by the study of Schuurmans et al. (2003) who utilized satellite estimates to identify conceptual model errors in a small sub basin of the MetaSWAP model in the Netherlands. The idea is to thoroughly investigate the observed differences in spatial patterns between observations and simulations in order to understand the underlying processes that generate patterns (Conradt et al., 2013). Subsequently, the newly gained insights can guide the development of a new parameterization scheme and calibration framework that can facilitate an improvement in the spatial model performance.

## 2 Methods

In this study two approaches are undertaken to estimate spatial patterns of evapotranspiration (ET) at national scale (Denmark); the first is based on a distributed hydrological model (MIKE-SHE) (described in section 2.1) and a second approach based on a Two Source Energy Balance (TSEB) driven by remote sensing data (Section 2.2). We acknowledge that both approaches are subject to uncertainties; however, the aim of this study is to evaluate the dominating spatial pattern across Denmark and to gain insights into which processes and variables generate these patterns. The evaluation will focus on the spatial pattern itself by neglecting differences in the absolute values of evapotranspiration.

In the last step of the study the current version of the DK-Model is modified by replacing the original root depth (RD), crop coefficient (Kc) and leaf area index (LAI) based on lookup tables and a land cover map by remote sensing based data which features detailed spatio-temporal information. A more detailed description is presented in

(Section 2.2 and 2.3)

## 2.1 Hydrological model

The National Water Resources Model (DK-model) (Højberg et al., 2013;Stisen et al., 2012) was developed at the Geological Survey of Denmark and Greenland in 1996 and updated several times (Henriksen et al., 2003;Højberg et al., 2013). The model is constructed within the hydrological model system MIKE-SHE (Abbott et al., 1986). The model works at 500 m

5 resolution and due to computational efficiency and differences in geology the DK-model was divided into 7 different domains that cover the entire country; however in this study only 6 of the 7 domains were selected covering approximately 98% of the country with and extension of 42.087 km$^2$. Domains used are presented in Fig. 1.

The model is based on a full 3-D finite difference groundwater module that is connected to a simplified two–layer unsaturated zone module (Yan and Smith, 1994). Furthermore, the model was previously calibrated using 191 discharge

10 stations and approximately 17.500 data entries of ground water head (GWH) (Stisen et al., 2012 ). The DK-Model has been extensively used in different applications with different objectives; assessment of climatic change (Karlsson et al., 2016) , water resources management (Henriksen et al., 2008), large scale nitrogen modelling (Windolf et al., 2011;Hansen et al., 2009;van der Keur et al., 2008)  highlighting the importance of the spatial component of the model and its reliability.

## 2.2 TSEB setup and remotely sensed derived inputs

15 The Two Source Energy Balance Model (TSEB) proposed by Norman et al. (1995) is used to retrieve mean monthly maps of ET across Denmark. In our study we have incorporated the code which is provided by the pyTSEB package (https://github.com/hectornieto/pyTSEB last accessed 30/01/2017). The applied model is a two layer model that treats soil and vegetation separately and estimates fluxes on the basis of LST and air temperature ($T_{Air}$) among other input variables. TSEB can successfully be applied with a single LST observation, if the viewing geometry of the observation is available

20 (VZA). The method uses an iterative process of canopy temperature (Tc) and soil temperature (Ts) to find the temperatures that satisfy the energy balance (Norman et al., 1995 ). Several inputs to TSEB are directly obtained from the LST product from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor all at 1 km spatial resolution; day time LST and day time VZA obtained from MOD11A1 and MYD11A1 products flown on TERRA and AQUA respectively. The decision of whether to use LST from TERRA or AQUA is  based on the percentage of high quality pixels available covering

25 Denmark in each scene. The quality flags included in the products is used to select only those pixels with the best observation possible.

LAI is derived using an empirical relationship with the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1973). The study focuses on the growing season from April to September. From a water resources perspective the spatial patterns of ET are regarded as irrelevant for the remaining months of the year due to energy limited conditions and very low

30 potential evapotranspiration. First the Nadir BRDF Adjusted Reflectance (NBAR) from the MCD43B4 product for that time period was used to calculate the NDVI using the following equation Eq. 1:

$$NDVI = \frac{B_2 - B_1}{B_2 + B_1},$$ (1)

where Bx indicates the MODIS band number.

Later, the Savitzky-Golay filter (Savitzky and Golay, 1964) available in the TIMESAT code (Jönsson and Eklundh, 2002;Jönsson and Eklundh, 2004) is selected to smooth the NDVI time series as it preserves the maximum and minimum values of the original dataset and guarantees consistency of time series. In situ measurements of LAI are usually expensive

5 and time consuming, therefore reference LAI was based on the tables used for the Danish National Water Resources model (Stisen et al., 2012) which are based on previous works of Refsgaard et al. (2011). Boegh et al. (2004) successfully derived LAI variability in Denmark using NDVI with an exponential function. In the study a similar approach where coefficients are adjusted to match the data from the National Water Resources model was applied resulting in Eq. 2:

$$LAI = \alpha \cdot e^{\beta \cdot NDVI} \tag{2}$$

10 Where $\alpha$ and $\beta$ are specific parameters for the study case.

Later, coefficients $\alpha$ and $\beta$ were adjusted to maximized the fit between the estimated LAI and reference LAI, leading to Eq. 3:

$$LAI = 0.0633 \cdot e^{5.524 \cdot NDVI} \tag{3}$$

Vegetation Height (VH) was derived assuming a simple linear regression as in Stisen et al. (2011) following Eq. 4 :

15 $$VH = 0.1 \cdot Height_{Max} + 0.9 \cdot Height_{Max} \cdot \left(\frac{LAI}{LAI_{Max}}\right) \tag{4}$$

Where $Height_{Max}$ is a value that changes for each land use class in meters. This relationship was applied on all land use classes except forest, which was set to a year round constant VH of 15m.

$LAI_{Max}$ indicates the maximum LAI value for each class.

The input albedo data was obtained from the MODIS MCD43B3 product and in order to reduce noise mean albedo maps are

20 generated creating 46 mean maps using all the scenes available for each day of the year (DOY) across different years using only the highest quality pixels.

The albedo mean maps for each 8 DOY period are later used as a climatology for the different years processed in TSEB.

Climate forcing data to run the TSEB are obtained from the ERA-Interim reanalysis data set (Berrisford et al., 2011;Dee et al., 2011) provided by the European Centre for Medium-Range Weather Forecast (ECMWF). We have incorporated 10 m

25 horizontal wind speed (10U), 2 m air temperature (2T), surface solar radiation (SSRD) and longwave incoming radiation (SLRD).

Fraction of Green vegetation was derived from LAI by two separate equations for the period defined as 1) beginning of growing season to peak of growing season (the greening phase) and 2) the rest of the year. The equation used outside the greening phase (and for needle leaf forest all year) was defined as Eq.5:

30 $$Fg_i = \frac{LAI_i - LAI_{i,min}}{LAI_{i,min}} \cdot 100 \tag{5}$$

Where $Fg_i$ indicates the Fraction of green for a certain pixel $i$, $LAI_{i,min}$ indicates the minimum LAI value for a pixel i in the time series.

For the greening phase (start to peak of growing season) a different equation was used to describe the very quick increase in fraction of green vegetation that occurs right when the very green plants emerge in the spring. The equation used in that interval is Eq. 6:

$$Fg_i = \frac{LAI_{i,max}}{LAI_{MaxClass}} \cdot (1 - e^{(-2 \cdot LAI_i)})$$ (6)

5    First the dates corresponding to the maximum LAI (Fig. 2) and minimum LAI are identified. Later the date where the LAI increases at least 20% above the minimum is identified as the breakpoint (Fig. 3). The break point is considered to be the start of the growing season and therefore the time when the Fg starts increasing rapidly. Between the breakpoint and maximum LAI equation 5 was used to force an abrupt change in the Fg value and make it rapidly increase to high values before decreasing after reaching a peak in LAI representing the onset of the senescent phase.

10   Finally, instantaneous estimates of ET are converted to daily ET values based on the assumption of a constant value of the evaporative fraction throughout the day (Sugita and Brutsaert, 1991;Brutsaert and Sugita, 1992) following  Eq. 7 :

dET=EF · Net Radiation                                                                                                              (7)

where dET represents the daily ET and EF is the Evaporative Fraction and Net radiation represents the daily Net Radiation. And where EF is calculated as in Eq.8:

15   EF=ET/netRad                                                                                                                              (8)

Where ET is the actual evapotranspiration and netRad is the Net Radiation at the same acquisition time. The assumption of constant EF over the course of a day is often not completely true  and also affects the estimates of daily ET (Gentine et al., 2007) but in the current application is not crucial since only the spatial pattern of the remote sensing estimates is utilized.

Later, the daily maps are aggregated to monthly mean maps by using only those days in which the national coverage

20   exceeded 50%. The monthly maps consist of all available ET estimates for a given month across all years (2001-2014) resulting in just 6 maps (April-Sept.)

This temporal aggregation process is conducted in the same way with simulations from the DK-Model  considering only the same cloud free pixels/grids as from the TSEB estimate and ensuring the comparability of the maps.

Data from three eddy covariance (EC) flux towers is used as a reference to perform a sensitivity analysis and calibration of

25   some of the vegetation parameters of the TSEB. A detailed description of the instrumentation and data processing of each of the flux sites can be found in Ringgaard et al. (2011). Latent heat (LE) measurements is obtained at a frequency of 30 min and the mean value of the observations from 11:00 to 13:00 were used as reference in the calibration of TSEB for the cloud free days where the TSEB estimates are available. The observed eddy covariance data are subject to energy balance closure problems, typically in the order of 20-25%, which is not unusual for this type of measurements (Hendricks Franssen et al.,

30   2008). The observed values used during TSEB evaluation are the Bowen ratio (Bowen, 1926) corrected values and the associated uncertainty estimate is the span between all error in the closure problem being assigned to the latent heat and no error in the latent heat.

### 2.3 Remote sensing derived hydrological input data

Part of the current study is to identify model inadequacies and test possible directions of model parametrization improvements. In an attempt to improve the initial DK-Model, spatially and temporally distributed root depth (RD) maps are generated using remote sensing data using a similar approach to Andersen et al. (2002) and Koch et al. (2017) where RD is calculated by Eq. 9:

$$RD_i = RD_{max} \frac{LAI_i}{LAI_{max}},$$
(9)

where $RD_i$ is RD in pixel $i$ in meters, $LAI_i$ is the LAI in that cell, and $LAI_{max}$ and $RD_{max}$ indicate the maximum values at $i$ in meters. This equation was modified and LAI was substituted by NDVI, leading to Eq. 10

$$RD_i = RD_{max} \frac{NDVI_i}{NDVI_{max}}$$
(10)

and used for forested land cover types, poorly vegetated areas and urban areas. RD can be considered an effective parameter in the DK-model which partly compensates for the lack of a specific vegetation (LAI) component in the evapotranspiration calculations. For instance, eq. 10 equips forest cells with a temporal varying RD which is contrary to our physical understanding, but it compensates for mixed land-use cells and undergrowth. RD of croplands and grasslands, denoted as $RD_{Agri}$ , was estimated by implementing some modifications to equation 10. $RD_{max}$ is known to vary depending on the soil type with shallower roots for sandy soils (Refsgaard et al., 2011) and the opposite for clay soils, therefore an empirical relation between the clay fraction (CF) in the soil and $RD_{max}$ is established for the different soil types. This equation is then included as a substitute of $RD_{max}$. In order to allow RD to reach zero, the second term in equation 11 is normalized by including $NDVI_{min}$ in the Eq. 11,

$$RD_{(agri)i} = [(12 \cdot CF_i) + 0.2] \cdot \frac{NDVI_i - NDVI_{min}}{NDVI_{max} - NDVI_{min}}$$
(11)

Where $CF_i$ indicates the clay fraction at pixel $i$, $NDVI_i$ indicates the NDVI value of pixel $i$, and $NDVI_{Min}$ and $NDVI_{Max}$ represent the minimum and maximum values of NDVI for the same pixels.

The constants of equation 10 are derived by matching the root depths of the original DK-model.

In addition, the crop coefficient (Kc), which is a correction factor for the reference evapotranspiration ($ET_{ref}$) accounting for the difference between a given crop or land surface and the reference crop on which the $ET_{ref}$ is based. Kc is derived from remotely sensed LAI using the following Eq. 12:

$$KC = 0.95 + 0.2 * (1 - e^{(-0.7 * LAI)})$$
(12)

Again the constants of equation 11 are derived to match the Kc-values of the original DK-model.

### 2.4 Optimization of vegetation parameters

A sensitivity analysis of the TSEB model is conducted with the help of PEST (http://www.pesthomepage.org/Home.php), a model-independent parameter estimation and uncertainty analysis tool. PEST evaluates the sensitivity of each parameter by perturbing the value of each of the parameters one at a time and subsequently analyse the response of the performed

perturbation with respect to a change in model performance. At last, the sensitivities are normalized using the most sensitive parameter as reference.

In order to optimize the vegetation parameters, the objective function is set to reduce the differences in monthly ET estimates at three EC measurements sites that represent the three main land cover types in DK (Agriculture, Forest and Meadow).

5   Only 4 parameters are calibrated using PEST. These parameters are: Priestley Taylor parameter for forested areas ($PT_{Forest}$), and the LAI max Class for the three land covers used to estimate the Fg ($LAI_{Max}^{Agriculture}$, $LAI_{Max}^{Forest}$, $LAI_{Max}^{Meadow}$).

## 2.5 Spatial pattern analysis; Empirical Orthogonal Functions (EOF)

The Empirical-Orthogonal-Functions (EOF) analysis is a statistical technique commonly used to evaluate large spatio-temporal datasets of hydrological states and fluxes (Mascaro et al., 2015;Perry and Niemann, 2007;Graf et al., 2014). EOF

10   decomposes the variability of a spatio-temporal dataset in two main components. First, a set of orthogonal spatial patterns (EOFs) which are time invariant and define statistically significant patterns of covariation. Second, a set of loadings that are time variant and specifying the importance of each EOF over time. Graf et al. (2014);Perry and Niemann (2007) described briefly the mathematical background of the EOF analysis. Most commonly, the EOF analysis is applied on either observational or modelled datasets, but recent applications stressed it's usability as a tool for spatial validation of distributed

15   hydrological models at catchment scale (Fang et al., 2015;Koch et al., 2016 ). Koch et al. (2015) suggested performing a joint EOF analysis on an integral data matrix that contains both, observed and simulated data. In this way, the resulting EOF maps honour the spatio-temporal variability of both datasets and the weighted difference between the loadings at specific times can be utilized to derive a quantitative pattern similarity score. The weighting is necessary, because the amount of covariation that lies in each EOF differs, where the first EOF is oriented in the direction of maximum covariation. Therefore,

20   the EOF based similarity score (SEOF) between an observed and a predicted ET map at time x can be formulated as (Eq. 13):

$$S_{EOF}^{x} = \sum_{i=1}^{n} w_i \left| \left( load_i^{simx} - load_i^{obsx} \right) \right| \tag{13}$$

where $w_i$, the covariation contribution of the i'th EOF, is multiplied with the absolute difference between the simulated loading ($load^{sim}$) and the observed loading ($load^{obs}$) of the i'th EOF at time x.

25   The EOF analysis applied in this study evaluated the differences in spatial patterns between the MIKE-SHE outputs in the original configuration and a modified version where three inputs (RD, LAI and Kc) of the model were replaced by those derived from remote sensing data.

## 3 Results and discussion

The results and discussion are presented in two sections; the first focuses on the sensitivity analysis and parameter

30   optimization of the TSEB model and the second features the spatial pattern evaluation of the DK-model using the maps obtained from the TSEB model.

### 3.1 Sensitivity analysis and TSEB calibration

The normalized sensitivity values of the 23 incorporated variables and parameters are illustrated in Fig. 3. The results are presented in three groups depending on the group they belong to: remote sensing data, forcing data and vegetation parameters.

5   The results show that the most sensitive parameter for the estimation of AET is LST. Interpreting the sensitivity values for each group individually stress that, for the remote sensing input, parameters that are directly related to LST such as emissivity of vegetation (EmissV) and soil (EmissS) are characterized by a high sensitivity as well. The next group, forcing data, exhibited high sensitivity in all parameters, except for wind speed. Overall Air temperature (TempAir) is the most sensitive forcing variable. These results indicate that the algorithm is largely controlled by the LST, LAI and climate forcing

10   data. This finding is considered ideal, since the actual parameters of the algorithm do not dominate the final spatial pattern. In general, the remote sensing and forcing data inputs can be considered observations which are not subject to calibration The least sensitive values are found among the vegetation parameters and some parameters in this group are calibrated to optimize the match to each of the three land cover types. Only minor adjustments of the estimated ET are made by calibrating the $PT_{Forest}$, $LAI_{Max}^{Agriculture}$, $LAI_{Max}^{Forest}$, $LAI_{Max}^{Meadow}$.

15   The results of monthly ET estimates are presented for all three sites in Fig. 4. The bars on the observed values indicate the uncertainty associated with the energy balance closure issues where the upper bound of the uncertainty bar represent the situation in which the residual energy is assigned to the latent heat (LE), whilst the lower bound represent the opposite situation in which all the residual energy is assigned to the sensible heat (H).

Generally the estimated ET values agree well with the EC-measurements especially considering the uncertainties associated

20   with energy balance closure and the spatial scale mismatch between the EC footprint and the remote sensing estimations. In order to minimize the effect of scale issues the EC values of ET at the three sites are compared to the average ET of the surrounding pixels estimated by TSEB. For this comparison, pixels that are considered as purely representative of the specific land cover type and therefore not contaminated by other land cover types are used. The selection of the pixels is carried out manually with the help of a high resolution image of the study area. The comparison is meant as an illustration of

25   the ability of the TSEB to describe the general annual variation and differentiate between land covers. The main aim of the TSEB application is to get robust national maps of growing season ET and the results show agreement on both, the seasonal variation and absolute levels of ET. On the other hand the separation between land covers is somewhat harder to evaluate because all three sites exhibit a similar level of ET.

ET in the forested areas remains mostly constant during the growing season with a tendency to increase at the end.

30   Agricultural areas on the other hand presented much higher variability with a rapid increase at the beginning of the growing season (May-June) and a decrease at the end (August-September). The Wetland shows a similar shape as the forest but with slightly higher ET values, and presents a big increase in the month of August that is not capture in the TSEB.

Mean monthly maps of ET are generated from daily TSEB estimations across all years to ensure consistent spatial patterns for robust spatial model evaluation with the aim to evaluate and improve the model performance. Such an improvement can be facilitated through optimal parametrization, and we therefore focus on the consistent spatial patterns rather than the temporal dynamics of ET variability. This is also reflected in the way the TSEB ET estimates are evaluated.

5 Results indicate that the TSEB ET estimates are within the measurement uncertainty of the EC at the three stations. The only pronounced disagreement is observed in the wetland during the month of August. Ringgaard et al. (2011) showed how the water level of the Skjern River raised during that time of the year and therefore increasing the values of ET in the EC measurements which is located at the bank of the river.

## 3.2 Spatial patterns

10 The mean monthly maps of cloud free ET generated with the TSEB model and DK-Model are presented in Fig. 5 and Fig. 6. The TSEB ET (Fig. 56) exhibits a clear difference between Eastern and Western Denmark with lower ET values in the sandy Western Denmark especially in the peak of the growing season (May-June). The clear E-W pattern identifies by the RSEB model is remarkable considering that it is opposite the general precipitation gradient (Fig 1). This highlights the strong influence of soil properties on the ET pattern across Denmark. Another feature is that forest areas have lower ET for the
15 selected cloud free days where canopy interception is not included.

Regarding the results from MIKE-SHE (Fig. 6) it can be observed that the E-W trend is not noticeable in the maps and the difference between forest and agriculture is less distinguishable. Moreover the effects of the zonal calibration are causing the area of modelling domain 2 (Fig. 1) to have much higher ET in comparison to the other domains, especially in June. In June the pronounced difference between parametrization in domains 4 and 5 is also clearly evident. From figures 5 and 6 it can be
20 extracted that there is almost no resemblance between the spatial patterns of the TSEB ET and the DK-model simulations on the national scale. This seems substantial since the model has been calibrated extensively. However, the applied discharge based calibration is dominated by the winter peak runoff, which conveys little information with respect to the spatial patterns of summer ET. This finding actually highlights the need for spatial pattern evaluation of distributed hydrological models since traditional discharge and groundwater head calibration does not secure reasonable ET patterns.

25 In a first attempt to improve the simulated spatial patterns of the DK-model, new parameterizations of root depth (RD), LAI and crop coefficient (Kc) are prepared based on fully distributed remote sensing and soil data as explained in section 2.3. These contain a higher degree of spatio-temporal detail than the original model input based on predefined tables (Refsgaard et al., 2011) and should reflect distributions that are more realistic and spatially consistent. Fig. 7 shows the modified DK-Model mean maps of ET. The patterns of these maps are more similar to those observed with the TSEB in which the E-W
30 pattern is quite evident.

Fig. 8 shows the spatial differences in growing season average ET between the DK-Model in its original configuration and the modified version based on remote sensing input. It is important to highlight at this point that the modified DK-Model is not recalibrated with the new inputs as this goes beyond the scope of this study. A recalibration may modify the water

balance in comparison to the original setup. However the performed modifications show some relevant features; the most noticeable visual improvement is the much larger gradient in the East-West pattern obtained in the modified DK-model, which emphasizes the more distinct resemblance to the pattern estimated with TSEB. Secondly the abrupt changes between the different model domains of the DK-Model are removed due to a more spatially consistent parameterization. This is

5  especially the case for model domain 2 and the boundary between domain 5 and 6. To analyse the driving mechanism behind the "observed" TSEB patterns and the simulated DK-model patterns the clay fraction used as input to the root depth calculation of the modified DK-Model, the observed average LST input to the TSEB model and the growing season average LAI are illustrated in Fig. 9. These maps reveal interesting findings; first the presence of the East-West pattern in the clay fraction map coincides visually quite well with the TSEB model mean outputs (Fig. 8) in spite of the fact that no soil

10  information has been included in the TSEB ET estimation. This indicates that the general perception, of lower ET for the sandy soils in the West due to soil moisture stress in the summer period, is captured well by the TSEB ET. The East-West pattern is not captured by the DK-model simulation even though the model is based on soil type information on field capacity and wilting point. On the other hand the modified DK-model captures much more of the East-West pattern because the clay fraction information is utilized to stretch the root depth distribution. Moreover, the TSEB pattern is mainly

15  controlled by the LST input combined with a fine scale variability introduced by the LAI patterns (Fig. 9).

Fig. 10 underlines that the changes in the original setup of the DK-Model and the modified version are large when compared using scatter plots using the mean normalized map of TSEB as reference. Even the dispersion in the scatter plots is large, the results reveal an improvement in the Pearson correlation coefficient (from r= 0.07 to r=0.33) and also the points move closer to the 1:1 line.

20  The EOF analysis (Fig. 11) extracts the spatio-temporal similarities and dissimilarities between the two different DK-Model configurations. The analysis is based on monthly mean maps generated using the daily simulations. Only the first three EOFs which, in combination, explain 71% of the total variance are presented. The first EOF captures 45.2% variance and the EOF loadings present very small differences and are equipped with positive sign throughout the entire period. Hence it can be interpreted that EOF 1 addresses the major similarities between the two model configurations. The EOF1 map captured the

25  component of the ET pattern which is driven by the soil properties, as it relates nicely with the mapped clay content in Fig. 10. The loadings of the second EOF in combination with its map add 15.7% to the explained variance. The apparent disagreement in values and sign between the loadings stressed that EOF2 captures the major dissimilarities between the two model configurations. The EOF2 pattern resembles the one found in EOF1, however it was characterized by less contrast and overall, it represents the added spatial detail of RD which is defined as a function of clay content and vegetation. The evident

30  E-W trend is strongest in the first three months of the growing season and afterwards the loadings drop to close to zero for the modified DK-model. The third EOF explains around 10% of variance and further records dissimilarities between the two models. Examining the loadings stresses that the modified DK-model plays a minor role in EOF3, as loadings are close to zero. However the first three months of the original DK-model seems well represented and the map underlines the granularity of the original setup, which is strongly driven by the discrete land-use map. Also, the model boundary between

area 5 and 6 appears in EOF3, which was caused by the zonal calibration of RD in the original DK-model. The overall similarity scores derived by the EOF analysis presented the maximum value for a pattern comparison in June (≈0.11) and the minimum corresponded to a day in April (≈0.02).

The results highlight a soil properties driven spatial pattern which is expected due to larger water holding capacity in clay dominated areas. This relationship is clearly evident in the TSEB data, although soil data do not drive the TSEB algorithm, but this information is embedded in the LST as LST can be used to map soil textures (Wang et al., 2015). In contrast the original DK-model includes soil type information, but clearly the soil parametrization does not have sufficient effect on the simulated patterns of ET. The spatial patterns can probably be improved through calibration, by increasing the contrast in soil parametrizations or by modifying the model formulations on the soil stress function. In the current study, a new root depth parametrization is applied where the spatio-temporal variation in the effective root depth is estimated based on a combination of the clay fraction map and remotely sensed NDVI time series. The simulated ET maps resulting from the new RS-based DK-model (including root depth Kc and LAI) are clearly much more similar to the TSEB estimates, although significant differences still occur in some regions. In order to achieve a better performance the transfer-function of the root depth and Kc parametrizations have to be calibrated against the spatial patterns of the TSEB (in combination with discharge and groundwater head). Unfortunately, re-calibration of the National DK-model goes beyond the scope of the current study, since a single model run of the entire DK-model requires around 40 hours (wall-clock time), but re-calibration will be part of future improvements of the model. This will ensure both spatially consistent parameterizations by utilizing transfer functions inspired by the parameterization scheme of The mesoscale Hydrologic Model (mHM) (Samaniego et al., 2010) and an optimal trade-off between discharge, groundwater head and spatial patterns of ET.

The applied EOF analysis identifies the spatio-temporal similarities and dissimilarities between the two DK-Model configurations. It allows pointing out driving mechanisms behind the simulated spatial patterns, such as the effect of the effective RD in the modified DK-Model in EOF2 or the sharp boundary of simulated ET caused by the zonal calibration of the original DK-Model (EOF3). These findings strengthened the EOF analysis as a suitable tool to meaningfully compare spatial patterns and to diagnose spatial model deficiencies. Recently, the proposed approach was applied by Koch et al. (2017) and Ruiz-Pérez et al. (2016) in a spatial sensitivity analysis and in a spatial pattern oriented model calibration, respectively. In the future the EOF analysis will be considered as a metric to re-calibrate the DK-Model with focus on spatial patterns of ET.

Two different approaches to retrieve ET are compared based on the hypothesis that both models, TSEB and the DK-model should present similar spatial pattern of ET. Results showed that the differences in the outputs were noticeable. These differences can be divided in three groups.

- Differences due to model setup: The DK-model is an aggregate of 6 domains. Each of these domains is calibrated individually, which leads to inconsistent spatial distributions of hydrological properties across domains. This increases the accuracy and performance of the model when evaluated only using discharge stations and ground water heads, but

ignores the spatial component as it is an aggregated evaluation. On the contrary, when the ET maps are obtained with TSEB these problems are not present as all the study area is treated the same way and with the same parameterization.

- Spatial differences due to the land cover parameterizations are also important and are clearly evident in the case of the TSEB maps when comparing forest and agriculture areas. In this study three different EC datasets are used to calibrate the vegetation parameters and these sites are assumed to be representative of each land cover at a national scale. In some cases this assumption might not be adequate as soil type/ forest type and other variables might affect the plant response to ET. The TSEB was adjusted to show this pattern which might in some cases be overestimated and therefore enhancing the contrast of the TSEB between two land cover types.

- Differences due to the models: Estimations of ET in the DK-Model are mainly driven by precipitation, root depth and soil properties and represent a residual in the water balance equation. On the other hand the TSEB relies mostly on forcing data and LST to estimate ET as a residual in the energy balance equation and does not take into account any soil information or rainfall. The similar features found between the mean annual maps of TSEB and clay fraction may indicate that the LST, and thereby TSEB, is sensitive to some soil properties. On the other hand these similarities between LST and soil property patterns can also be explained by the fact that areas with sandy soils and low clay fraction are coincident with areas with lower agricultural production and higher risk of summer drought and vegetation under soil water stress.

## 4 Conclusions

In this study the potential of remote sensing outputs to evaluate spatial patterns of hydrological models has been shown. The information derived from remote sensing data can be used as a diagnostic tool for revealing model structural insufficiencies and inconsistencies. Additionally, remote sensing derived variables can be used in hydrological models and hence adding spatial information that is finally translated to the outputs of the models. The use of spatial metrics is beneficial to identify spatial model deficiencies. Furthermore such metrics are required for a spatial pattern oriented model calibration in order to meaningfully compare the changes in the spatial patterns.

Hydrological model evaluations have traditionally focussed on the temporal dynamics of the outputs and not so much on the spatial component. This study shows that more attention must be given to the spatial pattern evaluation as traditional calibration does not ensure a realistic spatially representation. If the spatial component of the model is neglected, the use of distributed hydrological models is not always meaningful and therefore the use of more simple models could be more appropriate.

Model re-calibration should focus on a combination of improved parameter regionalization including clay fraction and other derived variables using remote sensing data, spatially countrywide consistency in parametrization and inclusion of dedicated spatial pattern oriented objective functions in combination with discharge and groundwater head observations.

This study was conducted over an energy limited region and over a specific time of the year where ET pays a more important role in the water cycle. Extending the study to other areas with different ecosystems that combine energy and water limited ecosystems will provide a wider overview on the factors controlling the ET spatial patterns.

**5 Acknowledgments**

5    The present work has been carried out under the SPACE (SPAtial Calibration and Evaluation in distributed hydrological modelling using satellite remote sensing data) project; The SPACE project is funded by the Villum Foundation (http://villumfonden.dk/ ) through their Young Investigator Programme.
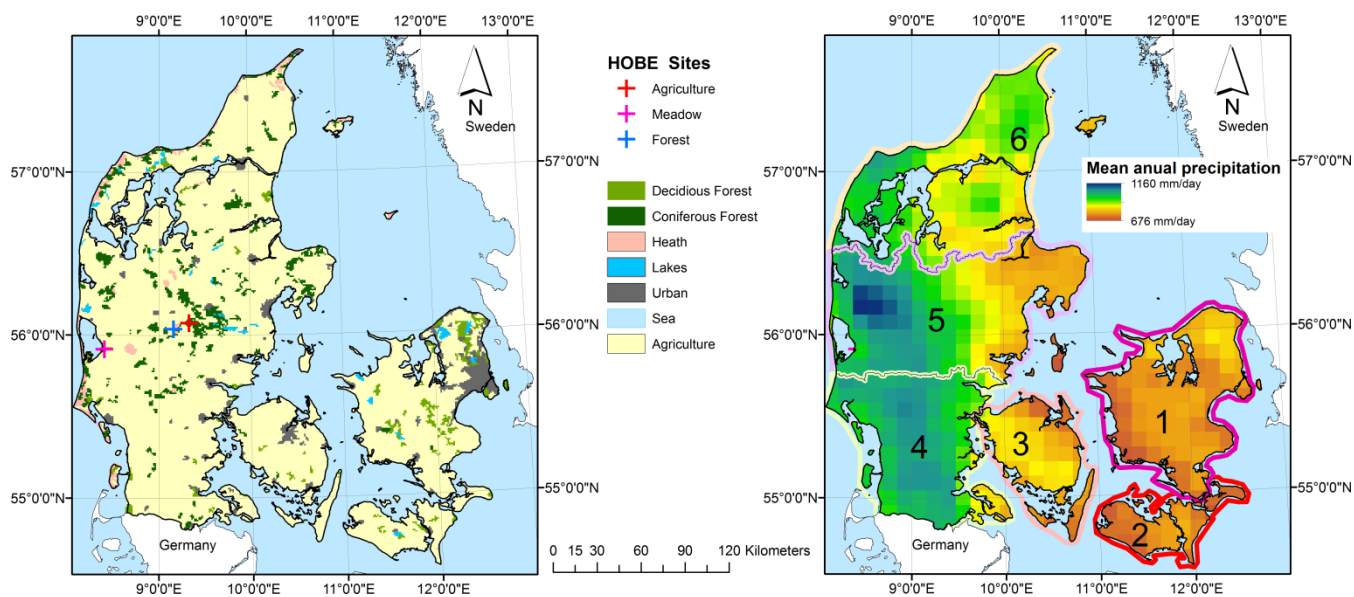
Figure **1. Left: Land cover map and model domains of the National Water Research Model of Denmark (Bornholm Island excluded in the figure. Right: National Water Research Model of Denmark domains and mean annual precipitation.**
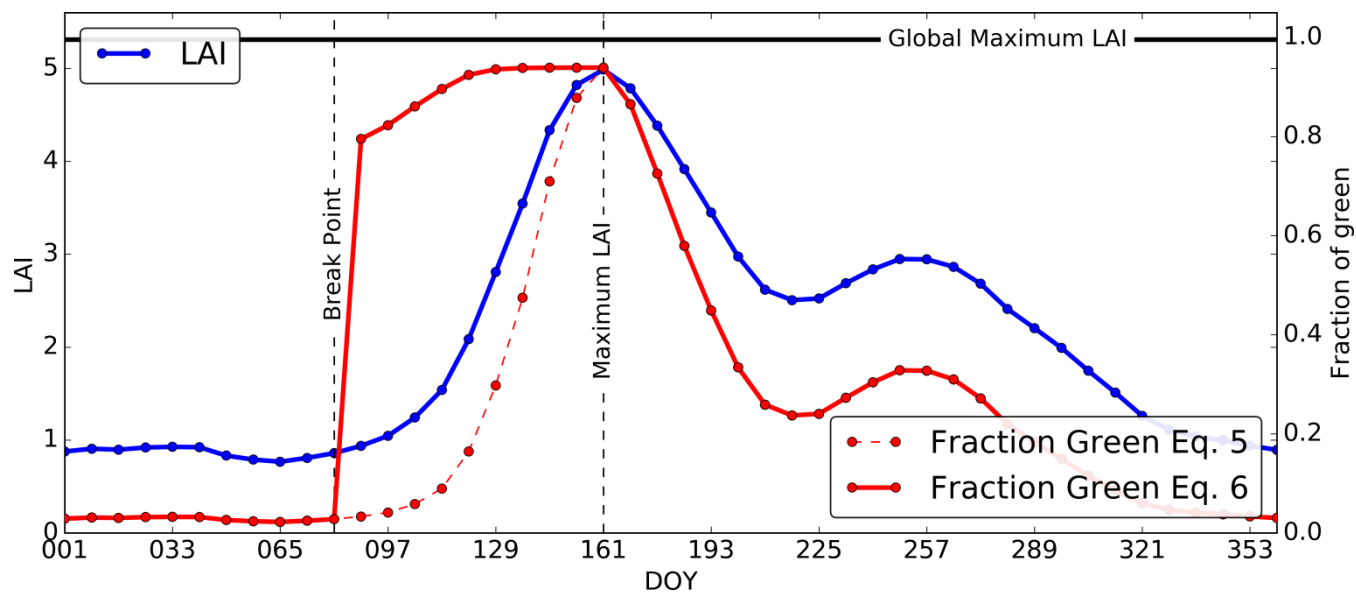
**Figure 2. Diagram of Fraction of Green (Fg) calculation based on the leaf area index (LAI). Data presented corresponds to an Agricultural pixel (row 100, coloum 84) from the dataset.**
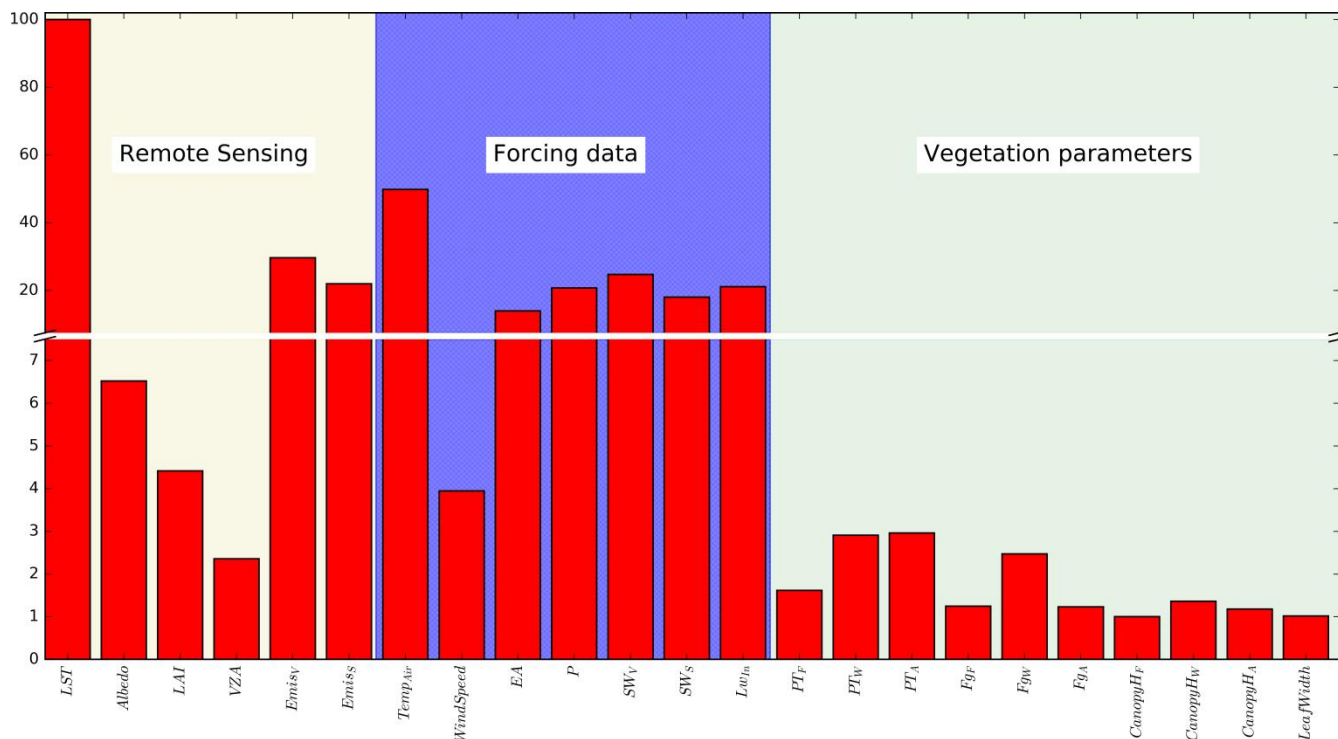
**Figure 3. Sensitivity of 28 TSEB model inputs obtained with PEST. Results are normalized using the most sensitive as reference. Acronims used: LST (Land Surface Temperature), LAI (Leaf Area Index), VZA (View Zenithal Angle), Emiss$_V$ (Emissivity of Vegetation), Emiss$_S$ (Emissivity of Soil), Temp$_{Air}$ (Air Temperature), EA (Water Vapor pressure above canopy), P (Atmospheric pressure), SW$_V$ (Short wave incoming radiation for vegetation), SW$_S$(Short wave incoming radiation for soil), LW$_{In}$ (Long wave incoming radiation), PT$_F$ (Pristley Taylor parameter for Forest), PT$_W$ (Pristley Taylor parameter for meadow), PT$_A$ (Pristley Taylor parameter for Agriculture), Fg$_F$ (Fraction of green vegetation for forest), Fg$_W$ (Fraction of green vegetation for meador), Fg$_A$ (Fraction of green vegetation for Agriculture), CanopyH$_F$ (Canopy height for forest), CanopyH$_W$ (Canopy height for meadow), CanopyH$_A$ (Canopy height for Agriculture).**
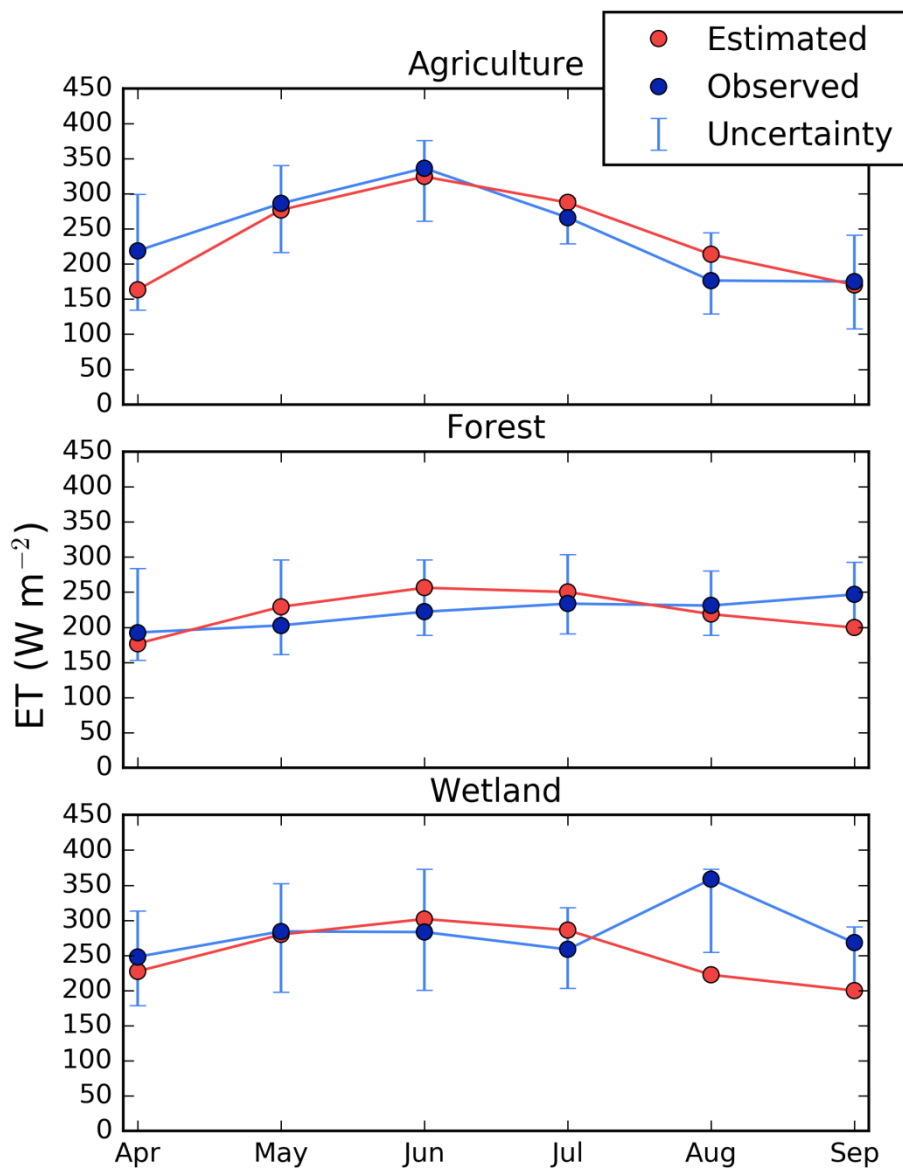
**Figure 4. Comparison of TSEB ET estimates in different land cover types. Uncertainty bars limits represent two situations, the upper in which all the residual is assigned to the latent heat and the lower one in which the residual is assigned to the sensible heat flux.**
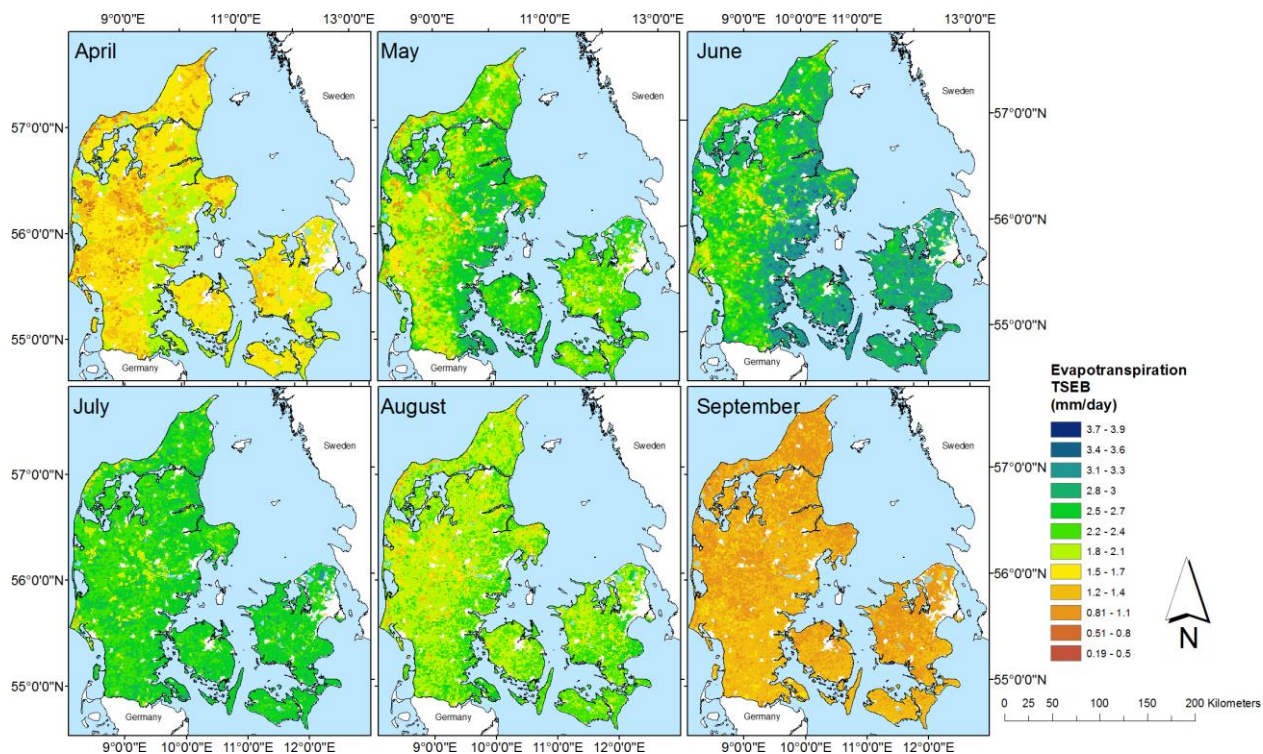
**Figure 5. Mean monthly TSEB ET maps in which urban areas have been masked out.**
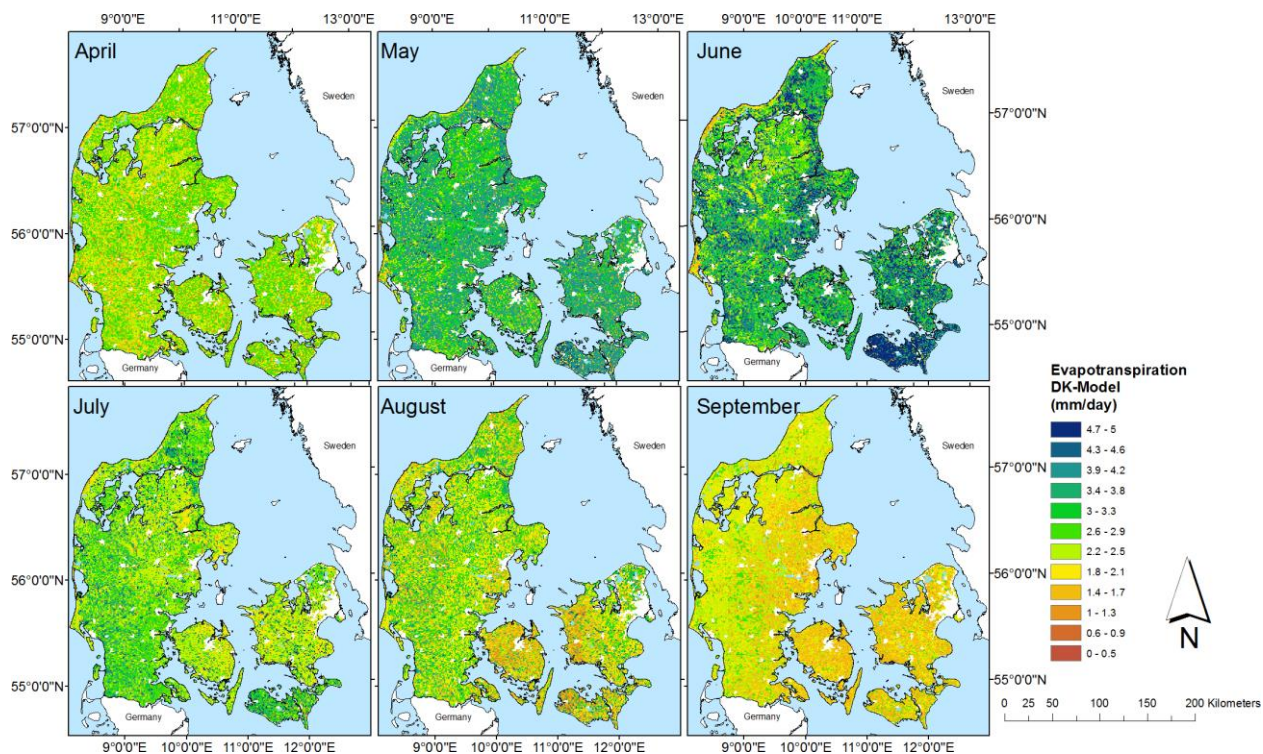
**Figure 6. Mean monthly DK-Model ET maps in which urban areas have been masked out.**
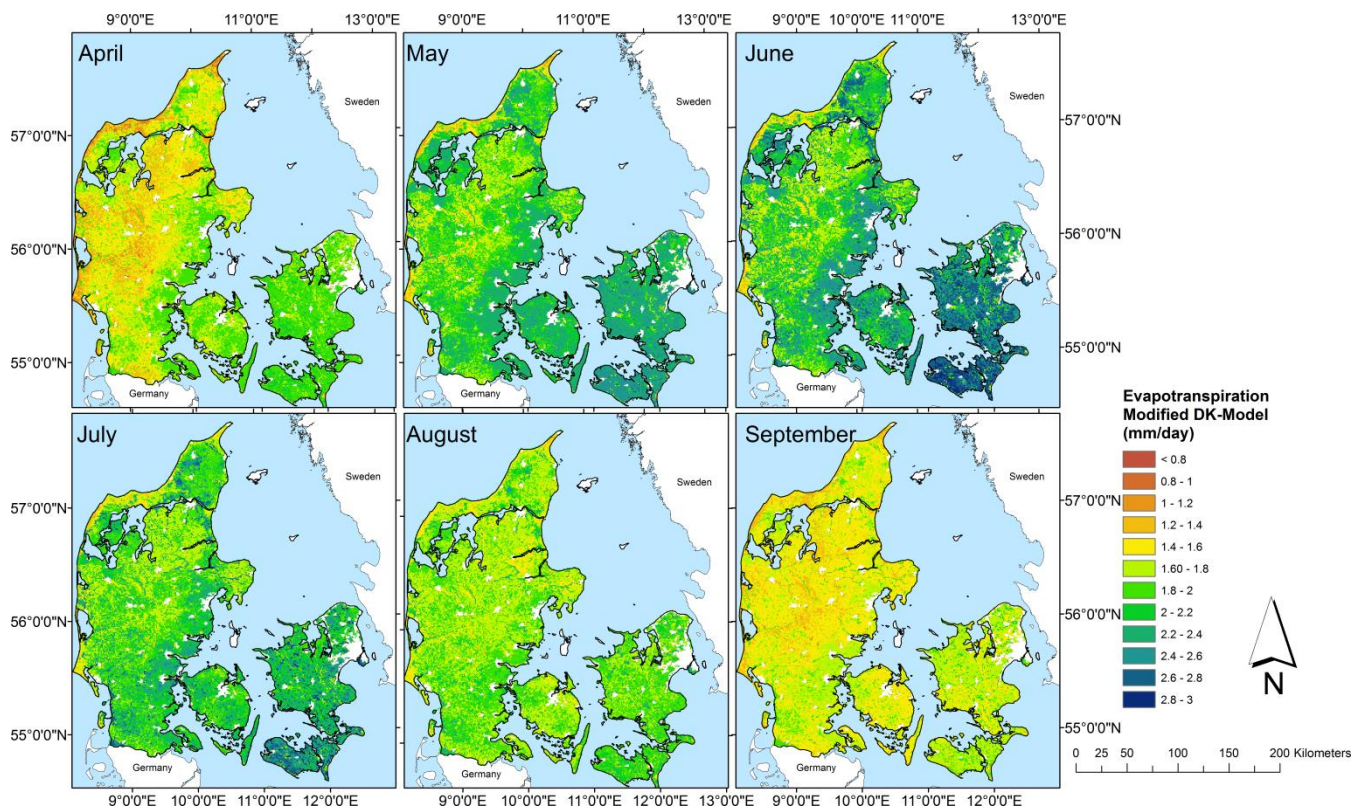
**Figure 7. Modified DK-Model ET mean maps in which urban areas have been masked out.**
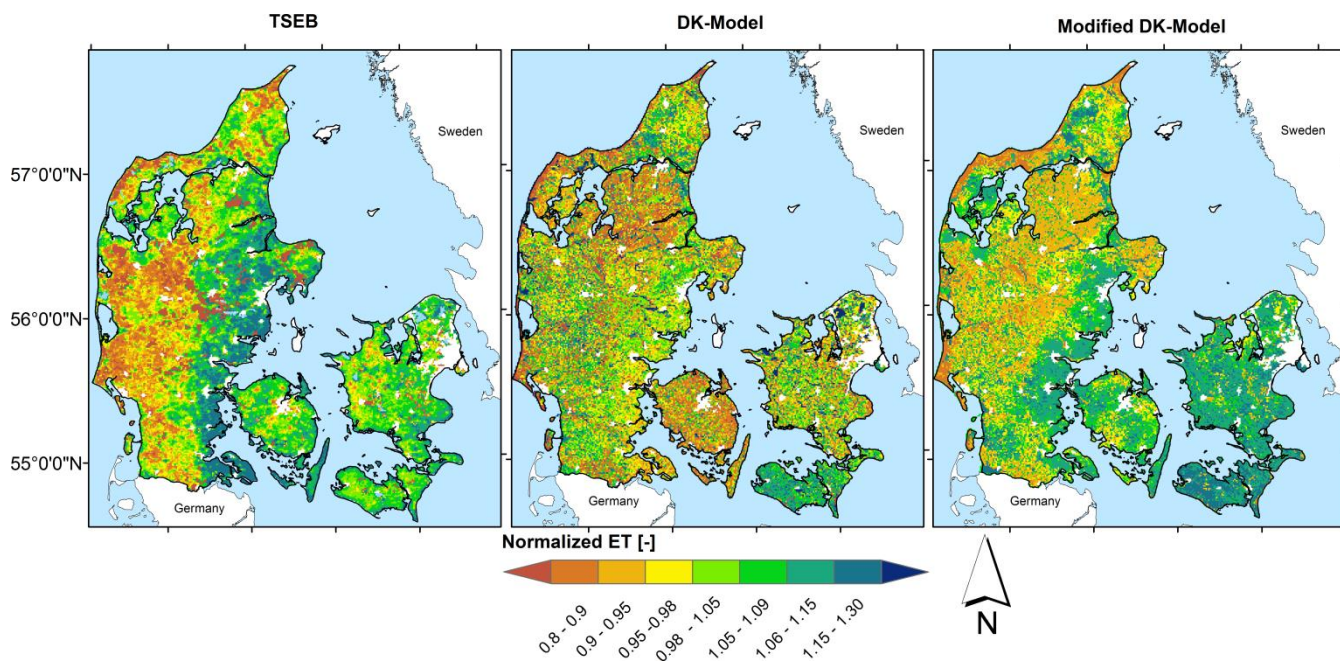
**Figure 8. Normalized growing season maps for the TSEB , DK-Model and the modified version of the DK-Model.**
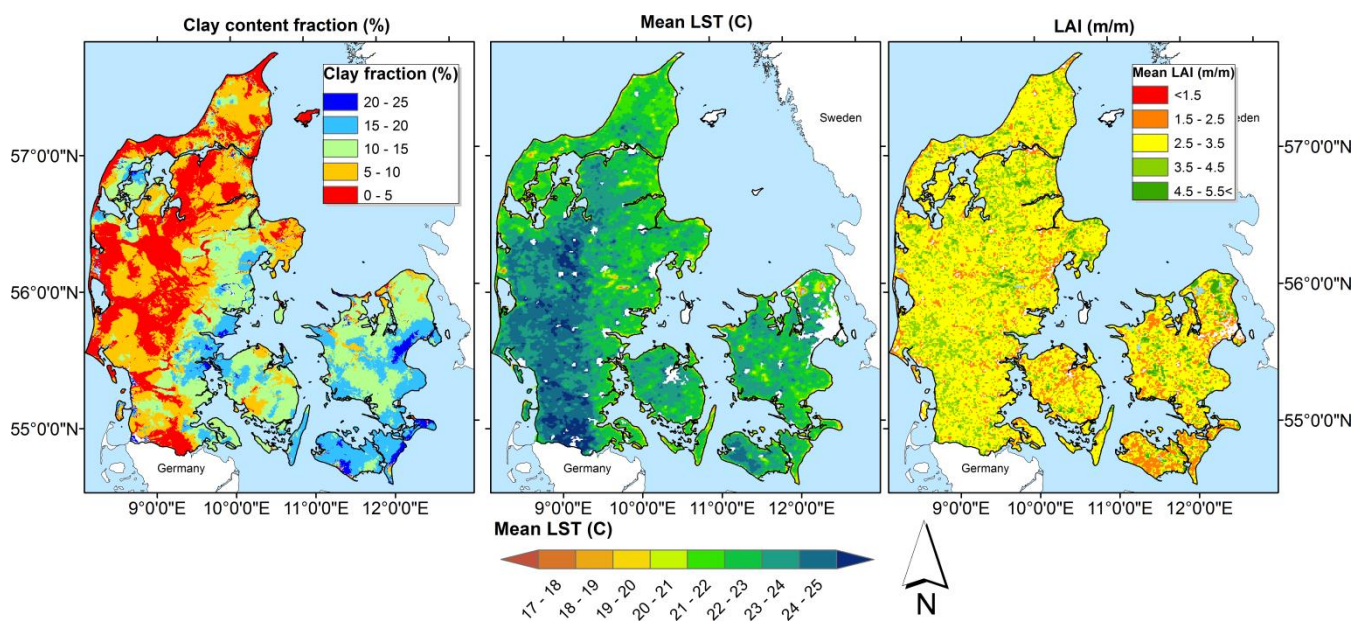
**Figure 9. Maps of three different parameters used in this study. Left map shows the clay fraction distribution. Center map**
5  **displays the mean values of LST during the growing season and right map displays the mean values of LAI during thre growing**
**season and used in the TSEB.**
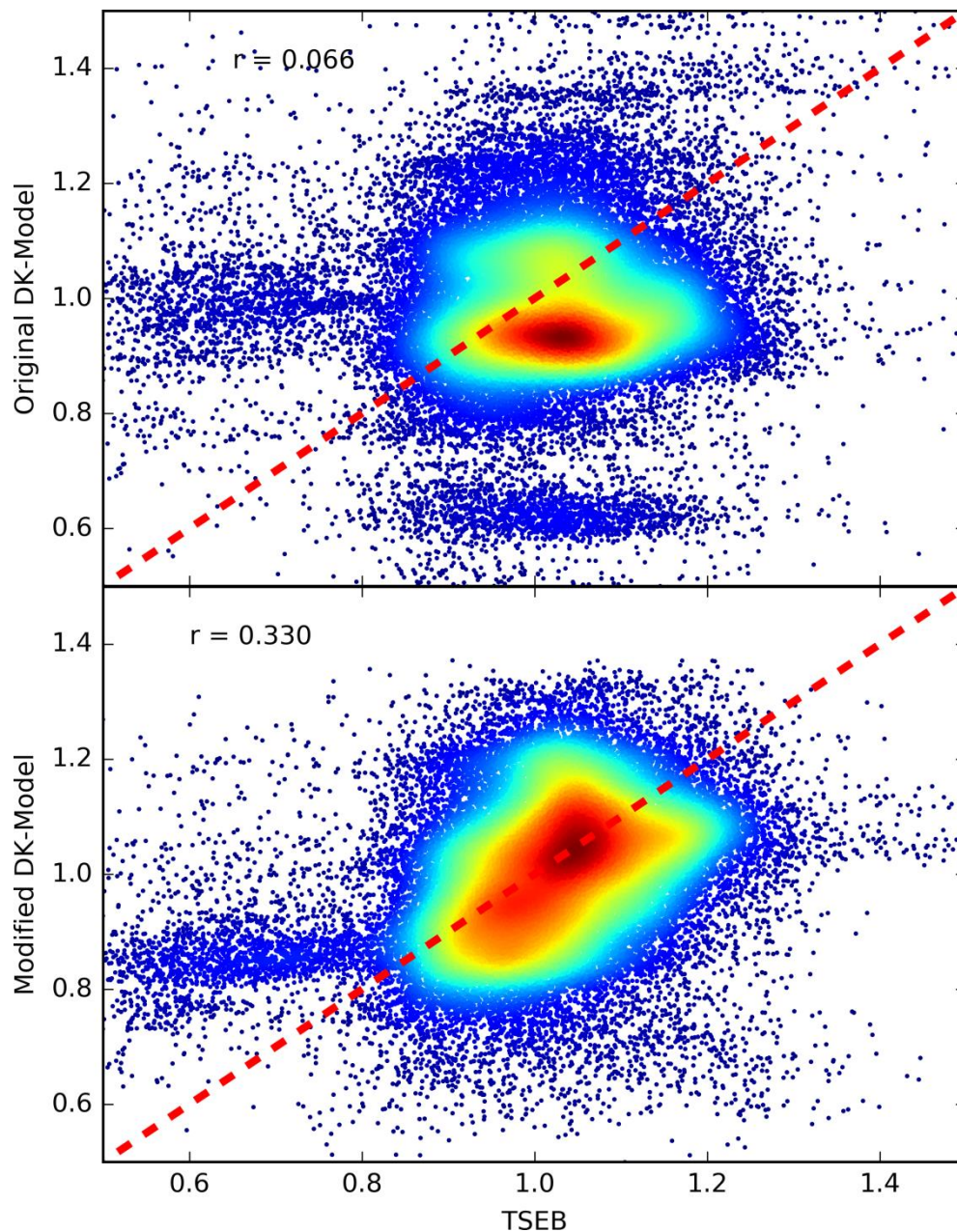
**Figure 10. Scatter plots of normalized ET showing the comparison of the TSEB against the two DK-Molde configurations. Upper scatter plot compares the TSEB against the original DK-Model, and lower compares against the modified DK-Model version.**
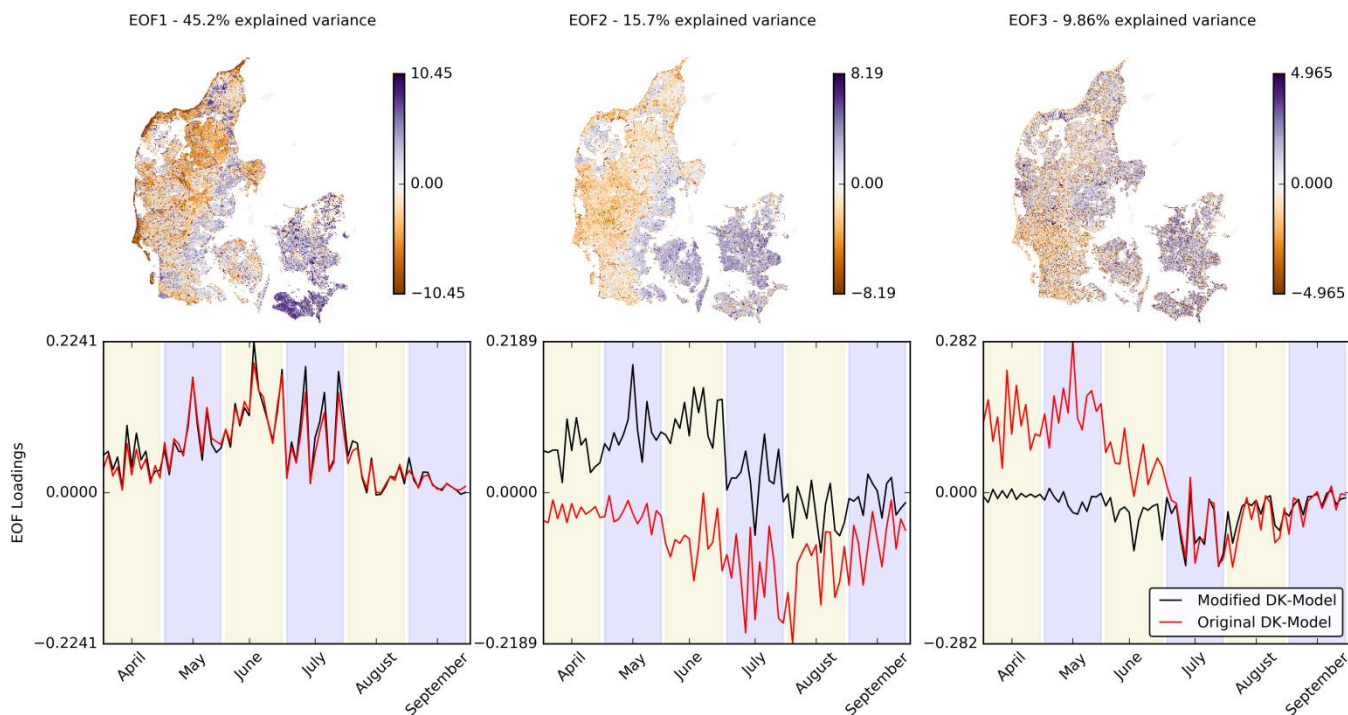
**Figure 11. Maps of the first three EOFs comparing the original setup of the MIKE-SHE model and the modified version of the model.**

# References

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., and Rasmussen, J.: An introduction to the European Hydrological System — Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system, Journal of Hydrology, 87, 45-59, http://dx.doi.org/10.1016/0022-1694(86)90114-9, 1986.

Andersen, J., Dybkjaer, G., Jensen, K. H., Refsgaard, J. C., and Rasmussen, K.: Use of remotely sensed precipitation and leaf area index in a distributed hydrological model, Journal of Hydrology, 264, 34-50, http://dx.doi.org/10.1016/S0022-1694(02)00046-X, 2002.

Berrisford, P., Dee, D. P., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Kållberg, P. W., Kobayashi, S., Uppala, S., and Simmons, A.: The ERA-Interim archive Version 2.0, in: ERA Report Series, ECMWF, Shinfield Park, Reading, 23, 2011.

Bertoldi, G., Notarnicola, C., Leitinger, G., Endrizzi, S., Zebisch, M., Della Chiesa, S., and Tappeiner, U.: Topographical and ecohydrological controls on land surface temperature in an alpine catchment, Ecohydrology, 3, 189-204, 10.1002/eco.129, 2010.

Boegh, E., Thorsen, M., Butts, M. B., Hansen, S., Christiansen, J. S., Abrahamsen, P., Hasager, C. B., Jensen, N. O., van der Keur, P., Refsgaard, J. C., Schelde, K., Soegaard, H., and Thomsen, A.: Incorporating remote sensing data in physically based distributed agro-hydrological modelling, Journal of Hydrology, 287, 279-299, http://dx.doi.org/10.1016/j.jhydrol.2003.10.018, 2004.

Bowen, I. S.: The ratio of heat losses by conduction and by evaporation from any water surface, Physical Review, 27, 779-787, 10.1103/PhysRev.27.779, 1926.

Brutsaert, W., and Sugita, M.: Application of self-preservation in the diurnal evolution of the surface energy budget to determine daily evaporation, Journal of Geophysical Research: Atmospheres, 97, 18377-18382, 10.1029/92JD00255, 1992.

Chen, J., Famiglietti, J. S., Scanlon, B. R., and Rodell, M.: Groundwater Storage Changes: Present Status from GRACE Observations, Surveys in Geophysics, 37, 397-417, 10.1007/s10712-015-9332-4, 2016.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J., and Rasmussen, R. M.: A unified approach for process-based hydrologic modeling: 1. Modeling concept, Water Resources Research, 51, 2498-2514, 10.1002/2015WR017198, 2015.

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

Corbari, C., and Mancini, M.: Calibration and validation of a distributed energy-water balance model using satellite data of land surface temperature and ground discharge measurements, Journal of Hydrometeorology, 15, 376-392, 10.1175/JHM-D-12-0173.1, 2014.

Corbari, C., Mancini, M., Li, J., and Su, Z.: Can satellite land surface temperature data be used similarly to river discharge measurements for distributed hydrological model calibration?, Hydrological Sciences Journal, 60, 202-217, 10.1080/02626667.2013.866709, 2015.

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., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553-597, 10.1002/qj.828, 2011.

Fang, Z., Bogena, H., Kollet, S., Koch, J., and Vereecken, H.: 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, Part 3, 1754-1767, http://dx.doi.org/10.1016/j.jhydrol.2015.08.011, 2015.

Freeze, R. A., and Harlan, R. L.: Blueprint for a physically-based, digitally-simulated hydrologic response model, Journal of Hydrology, 9, 237-258, 1969.

Gentine, P., Entekhabi, D., Chehbouni, A., Boulet, G., and Duchemin, B.: Analysis of evaporative fraction diurnal behaviour, Agricultural and Forest Meteorology, 143, 13-29, http://dx.doi.org/10.1016/j.agrformet.2006.11.002, 2007.

Githui, F., Selle, B., and Thayalakumaran, T.: Recharge estimation using remotely sensed evapotranspiration in an irrigated catchment in southeast Australia, Hydrological Processes, 26, 1379-1389, 10.1002/hyp.8274, 2012.

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

Grayson, R. B., and Blöschl, G.: Spatial modelling of catchment dynamics, in: Spatial Patterns in Catchment Hydrology: Observations and Modelling, edited by: Grayson, R. B., Blöschl, G. (Eds.), Cambridge University Press, 51–81, 2000.

Guzinski, R., Nieto, H., Stisen, S., and Fensholt, R.: Inter-comparison of energy balance and hydrological models for land surface energy flux estimation over a whole river catchment, Hydrol. Earth Syst. Sci., 19, 2017-2036, 10.5194/hess-19-2017-2015, 2015.

Hansen, J. R., Refsgaard, J. C., Ernstsen, V., Hansen, S., Styczen, M., and Poulsen, R. N.: An integrated and physically based nitrogen cycle catchment model, Hydrology Research, 40, 347-363, 10.2166/nh.2009.035, 2009.

Hendricks Franssen, H. J., Brunner, P., Makobo, P., and Kinzelbach, W.: Equally likely inverse solutions to a groundwater flow problem including pattern information from remote sensing images, Water Resources Research, 44, 10.1029/2007WR006097, 2008.

5  Henriksen, H. J., Troldborg, L., Nyegaard, P., Sonnenborg, T. O., Refsgaard, J. C., and Madsen, B.: Methodology for construction, calibration and validation of a national hydrological model for Denmark, Journal of Hydrology, 280, 52-71, http://dx.doi.org/10.1016/S0022-1694(03)00186-0, 2003.

Henriksen, H. J., Troldborg, L., Højberg, A. L., and Refsgaard, J. C.: Assessment of exploitable groundwater resources of Denmark by use of ensemble resource indicators and a numerical groundwater–surface water model, Journal of Hydrology, 348, 224-240,
10  http://dx.doi.org/10.1016/j.jhydrol.2007.09.056, 2008.

Højberg, A. L., Troldborg, L., Stisen, S., Christensen, B. B. S., and Henriksen, H. J.: Stakeholder driven update and improvement of a national water resources model, Environmental Modelling & Software, 40, 202-213, http://dx.doi.org/10.1016/j.envsoft.2012.09.010, 2013.

Immerzeel, W. W., and Droogers, P.: Calibration of a distributed hydrological model based on satellite evapotranspiration, Journal of
15  Hydrology, 349, 411-424, 10.1016/j.jhydrol.2007.11.017, 2008.

Immerzeel, W. W., Droogers, P., de Jong, S. M., and Bierkens, M. F. P.: Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing, Remote Sensing of Environment, 113, 40-49, http://dx.doi.org/10.1016/j.rse.2008.08.010, 2009.

Jönsson, P., and Eklundh, L.: Seasonality extraction by function fitting to time-series of satellite sensor data, IEEE Transactions on
20  Geoscience and Remote Sensing, 40, 1824-1832, 10.1109/TGRS.2002.802519, 2002.

Jönsson, P., and Eklundh, L.: TIMESAT—a program for analyzing time-series of satellite sensor data, Computers & Geosciences, 30, 833-845, http://dx.doi.org/10.1016/j.cageo.2004.05.006, 2004.

Karlsson, I. B., Sonnenborg, T. O., Refsgaard, J. C., Trolle, D., Børgesen, C. D., Olesen, J. E., Jeppesen, E., and Jensen, K. H.: Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change, Journal of
25  Hydrology, 535, 301-317, http://dx.doi.org/10.1016/j.jhydrol.2016.01.069, 2016.

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

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

Koch, J., Mendiguren, G., Mariethoz, G., and Stisen, S.: Spatial Sensitivity Analysis of Simulated Land Surface Patterns in a Catchment Model Using a Set of Innovative Spatial Performance Metrics, Journal of Hydrometeorology, 18, 1121-1142, 10.1175/jhm-d-16-0148.1, 2017.

35  Lettenmaier, D. P., Alsdorf, D., Dozier, J., Huffman, G. J., Pan, M., and Wood, E. F.: Inroads of remote sensing into hydrologic science during the WRR era, Water Resources Research, 51, 7309-7342, 10.1002/2015WR017616, 2015.

Li, H. T., Brunner, P., Kinzelbach, W., Li, W. P., and Dong, X. G.: Calibration of a groundwater model using pattern information from remote sensing data, Journal of Hydrology, 377, 120-130, http://dx.doi.org/10.1016/j.jhydrol.2009.08.012, 2009.

Mascaro, G., Vivoni, E. R., and Méndez-Barroso, L. A.: Hyperresolution hydrologic modeling in a regional watershed and its
40  interpretation using empirical orthogonal functions, Advances in Water Resources, 83, 190-206, http://dx.doi.org/10.1016/j.advwatres.2015.05.023, 2015.

Maxwell, R. M., and Kollet, S. J.: Interdependence of groundwater dynamics and land-energy feedbacks under climate change, Nature Geosci, 1, 665-669, 2008.

Mendiguren, G., Pilar Martín, M., Nieto, H., Pacheco-Labrador, J., and Jurdao, S.: Seasonal variation in grass water content estimated
45  from proximal sensing and MODIS time series in a Mediterranean Fluxnet site, Biogeosciences, 12, 5523-5535, 10.5194/bg-12-5523-2015, 2015.

Norman, J. M., Kustas, W. P., and Humes, K. S.: Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature, Agricultural and Forest Meteorology, 77, 263-293, http://dx.doi.org/10.1016/0168-1923(95)02265-Y, 1995.

50  Perry, M. A., and Niemann, J. D.: Analysis and estimation of soil moisture at the catchment scale using EOFs, Journal of Hydrology, 334, 388-404, http://dx.doi.org/10.1016/j.jhydrol.2006.10.014, 2007.

Rajib, M. A., Merwade, V., and Yu, Z.: Multi-objective calibration of a hydrologic model using spatially distributed remotely sensed/in-situ soil moisture, Journal of Hydrology, 536, 192-207, http://dx.doi.org/10.1016/j.jhydrol.2016.02.037, 2016.

Refsgaard, J. C.: Parameterisation, calibration and validation of distributed hydrological models, Journal of Hydrology, 198, 69-97,
55  http://dx.doi.org/10.1016/S0022-1694(96)03329-X, 1997.

Refsgaard, J. C., Stisen, S., Højberg, A. L., Olsen, M., Henriksen, H. J., Børgesen, C. D., Vejen, F., Kern-Hansen, C., and Blicher-Mathiesen, G.: DANMARKS OG GRØNLANDS GEOLOGISKE UNDERSØGELSE RAPPORT 2011/77, Geological Survey of Danmark and Greenland (GEUS), 2011.

Ridler, M.-E., Madsen, H., Stisen, S., Bircher, S., and Fensholt, R.: Assimilation of SMOS-derived soil moisture in a fully integrated
5   hydrological and soil-vegetation-atmosphere transfer model in Western Denmark, Water Resources Research, 50, 8962-8981, 10.1002/2014WR015392, 2014.

Rientjes, T. H. M., Muthuwatta, L. P., Bos, M. G., Booij, M. J., and Bhatti, H. A.: Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration, Journal of Hydrology, 505, 276-290, http://dx.doi.org/10.1016/j.jhydrol.2013.10.006, 2013.

10  Ringgaard, R., Herbst, M., Friborg, T., Schelde, K., Thomsen, A. G., and Soegaard, H.: Energy Fluxes above Three Disparate Surfaces in a Temperate Mesoscale Coastal Catchment, Vadose Zone Journal, 10, 54-66, 10.2136/vzj2009.0181, 2011.

Rouse, J. W., Haas, R. H., Deering, D. W., and Schell, J. A.: Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation, Goddard Space Flight Center, Greenbelt, MD, 87, 1973.

Ruiz-Pérez, G., Koch, J., Manfreda, S., Caylor, K., and Francés, F.: Calibration of a parsimonious distributed ecohydrological daily model
15  in a data scarce basin using exclusively the spatio-temporal variation of NDVI, Hydrol. Earth Syst. Sci. Discuss., 2016, 1-33, 10.5194/hess-2016-573, 2016.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, 10.1029/2008WR007327, 2010.

Savitzky, A., and Golay, M. J. E.: Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Analytical Chemistry,
20  36, 1627-1639, 10.1021/ac60214a047, 1964.

Schuurmans, J. M., Troch, P. A., Veldhuizen, A. A., Bastiaanssen, W. G. M., and Bierkens, M. F. P.: Assimilation of remotely sensed latent heat flux in a distributed hydrological model, Advances in Water Resources, 26, 151-159, http://dx.doi.org/10.1016/S0309-1708(02)00089-1, 2003.

Stisen, S., McCabe, M. F., Refsgaard, J. C., Lerer, S., and Butts, M. B.: Model parameter analysis using remotely sensed pattern
25  information in a multi-constraint framework, Journal of Hydrology, 409, 337-349, http://dx.doi.org/10.1016/j.jhydrol.2011.08.030, 2011.

Stisen, S., Hojberg, A. L., Troldborg, L., Refsgaard, J. C., Christensen, B. S. B., Olsen, M., and Henriksen, H. J.: On the importance of appropriate precipitation gauge catch correction for hydrological modelling at mid to high latitudes, Hydrology and Earth System Sciences, 16, 4157-4176, 10.5194/hess-16-4157-2012, 2012.

Sugita, M., and Brutsaert, W.: Daily evaporation over a region from lower boundary layer profiles measured with radiosondes, Water
30  Resources Research, 27, 747-752, 10.1029/90WR02706, 1991.

van der Keur, P., Hansen, J. R., Hansen, S., and Refsgaard, J. C.: Uncertainty in Simulation of Nitrate Leaching at Field and Catchment Scale within the Odense River Basin, Vadose Zone Journal, 7, 10-21, 10.2136/vzj2006.0186, 2008.

Vansteenkiste, T., Tavakoli, M., Van Steenbergen, N., De Smedt, F., Batelaan, O., Pereira, F., and Willems, P.: Intercomparison of five lumped and distributed models for catchment runoff and extreme flow simulation, Journal of Hydrology, 511, 335-349,
35  http://dx.doi.org/10.1016/j.jhydrol.2014.01.050, 2014.

Vereecken, H., Pachepsky, Y., Simmer, C., Rihani, J., Kunoth, A., Korres, W., Graf, A., Franssen, H. J. H., Thiele-Eich, I., and Shao, Y.: On the role of patterns in understanding the functioning of soil-vegetation-atmosphere systems, Journal of Hydrology, 542, 63-86, http://dx.doi.org/10.1016/j.jhydrol.2016.08.053, 2016.

Wanders, N., Bierkens, M. F. P., de Jong, S. M., de Roo, A., and Karssenberg, D.: The benefits of using remotely sensed soil moisture in
40  parameter identification of large-scale hydrological models, Water Resources Research, 50, 6874-6891, 10.1002/2013WR014639, 2014.

Wang, D.-C., Zhang, G.-L., Zhao, M.-S., Pan, X.-Z., Zhao, Y.-G., Li, D.-C., and Macmillan, B.: Retrieval and Mapping of Soil Texture Based on Land Surface Diurnal Temperature Range Data from MODIS, PLOS ONE, 10, e0129977, 10.1371/journal.pone.0129977, 2015.

Wang, L., Koike, T., Yang, K., and Yeh, P. J.-F.: Assessment of a distributed biosphere hydrological model against streamflow and MODIS land surface temperature in the upper Tone River Basin, Journal of Hydrology, 377, 21-34,
45  http://dx.doi.org/10.1016/j.jhydrol.2009.08.005, 2009.

Windolf, J., Thodsen, H., Troldborg, L., Larsen, S. E., Bøgestrand, J., Ovesen, N. B., and Kronvang, B.: A distributed modelling system for simulation of monthly runoff and nitrogen sources, loads and sinks for ungauged catchments in Denmark, Journal of Environmental Monitoring, 13, 2645-2658, 10.1039/c1em10139k, 2011.

Yan, J., and Smith, K. R.: SIMULATION OF INTEGRATED SURFACE WATER AND GROUND WATER SYSTEMS - MODEL
50  FORMULATION1, JAWRA Journal of the American Water Resources Association, 30, 879-890, 10.1111/j.1752-1688.1994.tb03336.x, 1994.

Zhang, D., Madsen, H., Ridler, M. E., Kidmose, J., Jensen, K. H., and Refsgaard, J. C.: Multivariate hydrological data assimilation of soil moisture and groundwater head, Hydrol. Earth Syst. Sci., 20, 4341-4357, 10.5194/hess-20-4341-2016, 2016.