Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

Hydrology and Earth System Sciences
Discussions

Evaluation of soil moisture from CCAM-CABLE simulation, satellite based models estimates and satellite observations: Skukuza and Malopeni flux towers regional case study

Floyd Vukosi Khosa[1,2], Mohau Jacob. Mateyisi[1], Martina Reynita van Der Merwe[1], Gregor Timothy Feig[1,3], Francois Alwyn Engelbrecht[4,5], Michael John Savage[2]

Correspondence to: Floyd Khosa (vukosikhosa@yahoo.com)

[1] CSIR, Natural Resources and the Environment – Global change and ecosystem dynamics, P.O. Box 395, Pretoria 0001, South Africa

[2] Agrometeorology Discipline, School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa

[3] Department of Geography, Geoinformatics and Meteorology, University of Pretoria, South Africa

[4] CSIR, Natural Resources and the Environment – Climate Studies, Modelling and Environmental Health, P.O. Box 395, Pretoria, 0001, South Africa

[5] Unit for Environmental Sciences and Management, North-West University, Potchefstroom, 2520, South Africa

Abstract

Reliable estimates of daily, monthly and seasonal soil moisture are useful in a variety of disciplines. The availability of continuous in situ soil moisture observation records in Southern Africa barely exists. In this regard, process based simulation model outputs turns out to be a valuable source of climate information, which is needed for guiding farming practises and policy interventions at various spatio-temporal scales. Despite their ability to yield historic and future projections of climatic conditions, simulation model outputs often reflect a certain degree of systematic uncertainty hence it is very important to evaluate their representativeness of spatial and temporal patterns against observations. To this effect, this study presents an evaluation of soil moisture outputs from a simulation and satellite data based soil moisture products. The simulation model consists of a global circulation model known as the conformal-cubic atmospheric model (CCAM), coupled to the CSIRO Atmosphere Biosphere Land Exchange model (CABLE). The satellite based soil moisture products include; satellite observations from the European space agency (ESA) and satellite observation based model estimates from the Global Land Evaporation Amsterdam model (GLEAM). The evaluation is done for both the surface (0-10 cm) and root zone (10-100 cm) using in situ soil moisture measurements collected from two savanna sites, located in the Kruger National Park, South Africa. For the two chosen sites with different soil types and vegetation cover, the evaluation considers soil moisture time series aggregated to a monthly time scale from all the data sources. In order to reflect the inter-comparability of CCAM-CABLE simulation output, and GLEAM model estimates, a qualitative analysis of phase agreement, using wavelet analysis is presented. The onset and offset of the wet period, for the two specific sites, is calculated for each of the models and the soil moisture time series covariance between CCAM-CABLE and the GLEAM is discussed. Our results indicate that both the simulation and satellite observation based model outputs are generally consistent with

the in situ soil moisture observations at the two study sites, especially at the surface. CCAM-CABLE and GLEAM inter-comparison also shows that the models are generally in phase, however with a time lag of about 12 and 20 days on average, for the surface and root zone respectively. In general the simulation compare well with the GLEAM model estimates, hence indicating that the key physical processes that drive soil moisture in CCAM-CABLE and GLEAM, at the surface and root zone, lead to an appreciable degree of mutual information. This is reinforced by a predominantly positive measure of covariance between the respective two soil moisture outputs.

Keywords: Soil Moisture, atmospheric model, land surface model, flux tower, cross wavelet

1 Introduction

Accurate estimates of daily, monthly and seasonal soil moisture are important in a number of fields including, but not limited to, agriculture (McNally et al., 2016), water resources planning (Decker, 2015), projection of precipitation (van den Hurk et al., 2012) and quantification of the impacts of extreme weather events such as drought (Sheffield and Wood, 2008), heat waves (Fischer et al., 2007; Lorenz et al., 2010) and floods (Brocca et al., 2011). Soil moisture has been identified as one of the 50 essential climate variables (ECVs) by the Global Climate Observing System (GCOS) and the European Space Agency climate change initiative (ESA-CCI) (McNally et al., 2016). Soil moisture drives fluxes of heat and water at the surface and directly impacts local and regional weather patterns. It is a key parameter to consider in the partitioning of precipitation and net radiation. Precipitation is partitioned into evapotranspiration (ET), infiltration and runoff. Net radiation is partitioned into latent and sensible heat fluxes (Xia et al., 2015; Yuan and Quiring, 2017) at the surface. Root zone soil moisture plays a vital role in the transpiration process of ET especially in arid and semi-arid regions where most of the water loss is accounted for by transpiration during the dry period (Jovanovic et al., 2015; Palmer et al., 2015). The temporal and spatial variation in soil moisture is controlled by vegetation, topography and soil properties (Xia et al., 2015). Regions where soil moisture strongly influences the atmosphere is at a transition between wet and dry climates. This is associated with the strong coupling between ET and soil moisture which is a characteristic of these regions (van den Hurk et al., 2012; Lorenz et al., 2010).

In situ data that is used as a reference in this study consists of surface and root zone soil moisture observations. The in situ data is mostly point based and when it comes to understanding general spatial patterns, this poses significant challenges which are associated with their limited spatio-temporal coverage (Yuan and Quiring, 2017). Direct satellite observations, on the other hand, are presently only available for the surface. To obtain root zone estimate of soil moisture, through model estimates, satellite based surface soil moisture data is used in conjunction with other ground based observations. The modelled soil moisture data are largely dependent on accurate surface forcing data (e.g. air temperature, precipitation and radiation) and the parameterisation of the land surface schemes (Xia et al., 2015). This is done in the frame work of physical based models whose accuracy may vary depending on the response of the models to the forcing data. Due to lack of publically available long term and complete in situ soil moisture measurements in Africa, and the world in general, global climate models (GCMs) are relied on to estimate the land surface states (Dirmeyer et al., 2013). The data produced by land surface-, hydrological- and GCMs have been widely evaluated for many continents and regions (Albergel et al., 2012; An et al., 2016; Dorigo et al., 2015; McNally et al., 2016; Yuan and Quiring, 2017). The evaluations of these soil moisture data products in Africa are sparse, mainly due to the lack of publically available in situ observations (Sinclair and Pegram, 2010). The available studies include those conducted by McNally et al. (2016) and Dorigo et al. (2015a) evaluating ESA-CCI satellite soil moisture products over East and West Africa respectively. This

study is inspired by the notion that the knowledge about soil moisture characteristic patterns, for the study region, can reliably be obtained by making a connection between data from simulation
87 experiments, theoretical or analytical models and in situ observation measurements. Satellite and model based soil moisture products, are capable of providing continuous observations at different temporal and spatial resolutions (Fang et al., 2016). Despite the in situ data being limited in coverage,
90 it is very useful for the calibration and validation of modelled and satellite derived soil moisture estimates (Xia et al., 2015).

The aims of this study are twofold. Firstly, it is to evaluate the ability of the simulated and estimated
93 soil moisture products to capture the observed variability in soil moisture at specific locations. The evaluation is done at two soil depths namely; surface (SSM, i.e., 0-10 cm) and root zone (RZSM, i.e. 10-100 cm), using long term in situ measurements. Much focus is directed to only two study sites that
96 are located in the Kruger National Park. This is due to data constraints. Secondly, the goal is to inter-compare simulated results of soil moisture against satellite based model estimates. This is done primarily at a regional level, where the absence of sufficient in situ observations over space and time
99 presents a major challenge. In particular, we look at spatio-temporal variations in simulated soil moisture data from a coupled land-atmosphere model i.e., conformal cubic atmospheric model (CCAM) coupled to the CSIRO Atmosphere Biosphere Land Exchange (CABLE) model against the
102 three versions of the European Space Agency (ESA) satellite observations (i.e., active, passive and combined), and estimates from three versions of the global land evaporation Amsterdam model (GLEAM). The goal is to obtain a clear picture of how the CCAM-CABLE simulation, satellite
105 derived soil moisture observations and model estimates fare against the in situ data, in capturing the seasonal cycles of soil moisture at a point, and to what extent the simulations and model estimates have mutual information at the regional level within inter-annual time scales. In particular the target is
108 to obtain clear reflection of phase agreement between the respective soil moisture data products and if this is representative of local conditions. In a nutshell, the study seeks to uncover interesting patterns in the observed data, for the study region, and highlight the strengths as well as aspects of the
111 simulation and model estimates which may benefit from continuous testing and improvements.

Clearly the ability of models to capture seasonal cycles of terrestrial processes such as soil moisture is indicative of how well the physical processes that underlie the variability of soil moisture over space
114 and time are represented. A comparison of satellite derived products with in situ observations also yield useful insight on the strengths and weaknesses of various remote sensing techniques that are used. Arguably, a climate models' ability to represent the seasonality of a system could be considered
117 more important than its agreement with observations in absolute values (Fang et al., 2016). The CCAM-CABLE model is specifically parameterised for African climatic and vegetation conditions. Its ability to capture key terrestrial processes such as changes in soil moisture will render it ideal for
120 use by the broader scientific community in understanding terrestrial processes on the continent. For the surrounding area of the study sites, where there are no in situ measurements, it will only suffice to reflect the extent at which two independent approaches for computing soil moisture co-vary, and
123 hence possess mutual information. In particular, we want to uncover how the strength of the covariance differs between the different soil and vegetation types.

2 Materials and methods

126 ### 2.1 Study sites and in situ observations

As mentioned earlier, for this study in situ soil moisture measurements obtained from the CSIR operated network of eddy co-variance flux towers at the Lowveld region of Mpumalanga (i.e.

Skukuza), and Limpopo provinces (i.e. Malopeni), are used. It is worth reiterating at this point that a number of other soil moisture in situ measurements sites exist in the country that are not publically available.

2.1.1 Skukuza

The Skukuza flux tower site is a long term measurement site, located (25.0197° S, 31.4969° E) within the Kruger national park conservation area in South Africa (Fig. 1). The site falls within a semi-arid savanna biome 370 m above sea level, with a mean annual rainfall of 547 mm year^{-1}, and minimum (during the dry season) and maximum (during the wet season) mean annual air temperatures of 14.5 and 29.5˚C respectively. The vegetation is dominated by an overstory of *Combretum apiculatum* (Sond.), and *Sclerocarya birrea* (Hochst.) with a height of approximately 8-10 m, and a tree cover of approximately 30% (Archibald et al., 2009). The understory is a grass layer dominated by *Panicum maximum* (Jacq.), *Digitaria eriantha* (Steud.), *Eragrostis rigidor* (Pilg.) and *Pogonarthria squarrosa* (Roem. and Schult.). The soil is of the Clovelly form with a sandy loam texture (Feig et al., 2008). The Skukuza flux tower site is extensively described in previous studies including those by (Archibald et al., 2009; Scholes et al., 2001). In situ soil moisture data is collected 90 metres north of the tower, and the measurements are taken at two profiles which are 8 m apart. The sensors are located at four different depths on both profiles i.e., 5, 15, 30 and 40 cm (Pinheiro and Tucker, 2001). Time domain reflectometry (TDR) probes (Campbell Scientific CS615L) were used to measure soil moisture at a 30 minute temporal resolution. The half hourly measurements are averaged to a daily time period (using 80% data threshold) to match the resolution of the other soil moisture products. In this study the in situ data from the year 2001 to 2014 is used.

2.1.2 Malopeni

The Malopeni flux tower is located (23.8325° S, 31.2145° E) 130 km north west of the Skukuza flux tower (Fig. 1), at an elevation of 384 m above sea level. The site has a mean annual rainfall of 472 mm year^{-1}, and minimum and maximum mean annual air temperatures ranging between 12.4 and 30.5˚C respectively. The site is dominated by broad leaf *Colophospermum mopane*, which characterise a hot and dry savanna (Ramoelo et al., 2014), *Combretum apiculatum* and *Acacia nigrescens* are also abundant at the site. The grass layer is dominated by *Schmidtia pappophoroides* and *Panicum maximum*. The soil at the site is predominantly of the shallow sandy loam texture. The soil moisture probes are located at four different profiles and depth. The sensors types and depths positioning are consistent between the Malopeni and Skukuza. Soil moisture is collected at four different profiles (i.e. 16 sensors at four depths), and averaged to represent surface and root zone soil moisture at this site. The tower has been collecting data since 2008 to date, however data has not been collected between January of 2010 and January of 2012 due to equipment failure.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

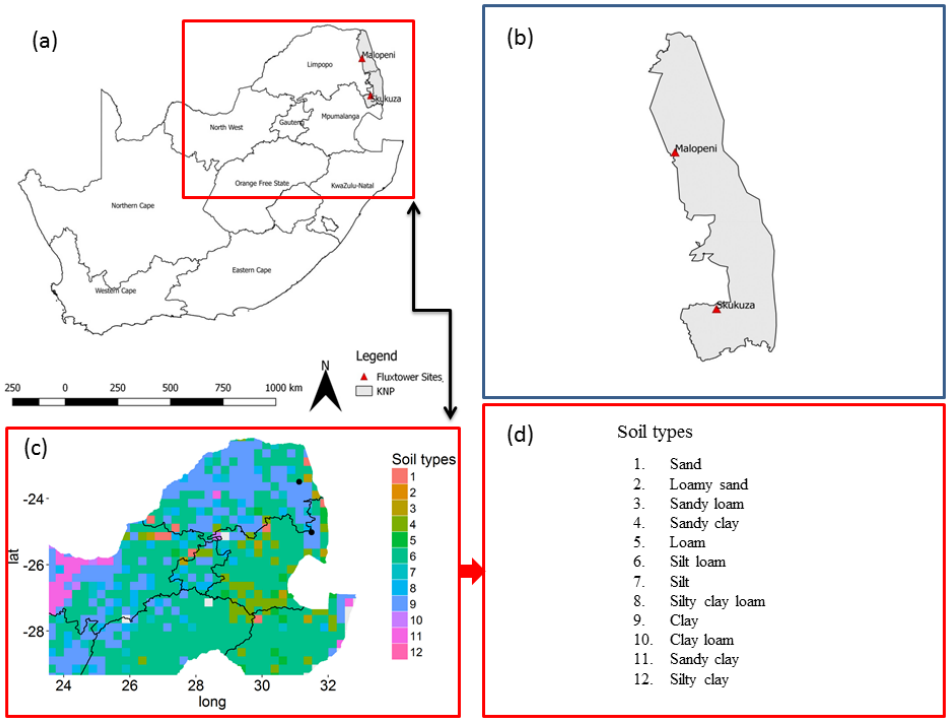


Figure 1. Maps indicating (a) South Africa, (b) Kruger national park (KNP), and flux tower sites; Skukuza and Malopeni, (c) the area considered for grid inter-comparison, with dominant soil types per grid cell, at a resolution of 25 km and (d) a description of soil types. The soil type's data is obtained online from https://soilgrids.org.

2.2 Satellite observations

The European space agency climate change initiative (ESA-CCI), satellite derived soil moisture datasets are used in this study (Liu et al., 2012; Yuan and Quiring, 2017). These global datasets are based on passive and active satellite microwave sensors, and provide surface soil moisture estimates at a resolution of ~25 km (i.e., 0.25˚) (Fang et al., 2016; Yuan and Quiring, 2017). The ESA-CCI merges soil moisture estimates from the active and passive satellite microwave sensors into one dataset (http://www.esa-soilmoisture-cci.org/), using the backward propagating cumulative distribution function method (Dorigo et al., 2015; Fang et al., 2016). The merged active and passive sensors are fully described in Fang et al. (2016), Dorigo et al. (2015) and (Liu et al., 2012). The merging of active and passive sensors is based on their sensitivity to vegetation density, as the accuracy of these product varies as a function of vegetation cover (Liu et al., 2012).

The difference between active and passive sensors is that, passive sensors are dependent on radiation, i.e., sunlight and are only able take measurements during the daylight hours, i.e. these sensors are not able to take measurements at night. Active sensors however are independent of sunlight as they provide their own source of energy to illuminate the objects they observe, and are able to take measurements both during the day and at night. Active sensors as opposed to passive sensors are able to penetrate through; clouds, fog, vegetation and are not affected by bad weather conditions (Fang et al., 2016). A number of studies evaluated these products at a regional and global scale using in situ

186 data, and concluded that passive sensors have better performance over bare to sparsely vegetated regions, whereas the active sensors perform better in moderately vegetated regions (Dorigo et al., 2015; Liu et al., 2012; McNally et al., 2016). Over densely vegetated areas such as tropical forests,
189 neither product produces reasonable estimates. The dense canopy hinders signals reflected from the soil surface (i.e. for passive sensors) or back scattering of active radiation before it reaches the soil surface for active sensors (Liu et al., 2012). The merged data product is used in this study as it has
192 better data coverage compared to the individual products. Missing data in satellite products is not unusual, since retrievals are normally at an interval of 2-3 days (Albergel et al., 2012). However, data from each of the different sensor types are also individually considered for the evaluation of long term
195 seasonal cycles.

2.3 Models for simulating soil moisture

2.3.1 CCAM-CABLE

198 The simulations produced by an ensemble of six downscaled global circulation models (GCMs), with an 8km spatial and six hourly temporal resolution, of the present-day climate, and future climate change over the Limpopo Province (Fig. 1c) are used in this study, focusing on periods from 2001 to
201 2014. A variable resolution GCM (i.e., CCAM), developed by the Commonwealth scientific and industrial research organisation (CSIRO) in Australia (McGregor, 2005; McGregor and Dix, 2001, 2008), is used to dynamically downscale the six GCMs. The six hourly data is averaged to daily. For
204 purposes of this study CCAM was coupled to a dynamic land-surface model (i.e., CABLE) and run in online mode. Simulations of six GCMs from the coupled model inter-comparison project phase five (CMIP5), and assessment report five (AR5), of the Intergovernmental panel on climate change
207 (IPCC), and for emission scenarios described by representative concentration pathways (RCPs) 4.5 and 8.5 (RCP 4.5 and 8.5), were first downscaled to a 50 km resolution globally, and then to 8km resolution for the Limpopo Province. The simulations span the period 1960-2014. The downscaled
210 GCMs includes the Community Climate System Model (CCSM4); the Norwegian Earth System Model (NorESM1-M); the Australian Community Climate and Earth System Simulator (ACCESS1-0); the National Centre for Meteorological Research Coupled Global Climate Model, version 5
213 (CNRM-CM5); the Geophysical Fluid Dynamics Laboratory Coupled Model (GFDL-CM3) and the Max Planck Institute Coupled Earth System Model (MPI-ESM-LR). This experiment follows earlier applications by Engelbrecht et al. (2015) and Muthige et al. (2018), therefore, details about bias
216 correction, the multiple nudging strategy and the computation process can be found in the listed references.

The CABLE soil submodule expresses soil as a heterogeneous system consisting of three component
219 namely water, air and solid (Kowalczyk et al., 2006; Wang et al., 2011). Air and water contest for the pore space with the change in the volume fractions as a result of drainage, precipitation, ET and snow melt. In this model there is no exchange of heat between the soil and the moisture due to the vertical
222 movement of water, as soil moisture is assumed to be at ground temperature. The soil is partitioned into six layers, with the layer thickness of 0.022 m, 0.058 m, 0.154 m, 1.085 m and 2.875 m from the top layer. The top layer contributes to evaporation, while plant roots extraction (i.e., transpiration)
225 occurs water from all the layers depending on the availability of soil water together with the fraction of plant roots in each of the layers (Wang et al., 2011). Soil moisture is solved numerically using the Richard's equation (Kowalczyk et al., 2006).

2.3.2 GLEAM

Global Land Evaporation Amsterdam Model (GLEAM) version three (v3.1) is a set of algorithms used to estimate surface, root-zone soil moisture and terrestrial evaporation using satellite forcing data. Three data sets from the GLEAM namely; v3a, v3b and v3c are used in this study. The data is freely available at www.gleam.eu. Version 3a is based on satellite observed; air temperature, radiation, soil moisture, snow water equivalent, vegetation optical depth and a multi-source precipitation product. Versions 3b and 3c are satellite based with common forcing data excluding soil moisture and vegetation optical depth as these are based on different passive and active microwave sensors, i.e. ESA CCI for v3b and Soil Moisture and Ocean Salinity (SMOS) for v3c (Martens et al., 2017).

Each grid cell in GLEAM contains fractions of four different land cover types namely; open water (e.g. dam, lake), short vegetation (i.e. grass), tall vegetation (i.e. trees) and bare soil. These fractions are based on the global vegetation continuous field product (MOD44B) with the exception of the fraction of open water. The MOD44B product is based on the moderate resolution image spectroradiometer (MODIS) observations (Martens et al., 2017). Soil moisture is estimated separately for each of these fractions and then aggregated to the scale of the pixel based on the fractional cover of each land cover type. Root zone soil moisture is calculated using a multi-layered water balance equation which uses snow melt and net precipitation as inputs, and drainage and evaporation as outputs (Miralles et al., 2011). The depth of soil moisture is a function of land-cover type comprising one layer of bare soil (0-10 cm), two layers for short vegetation (0-10, 10-100 cm) and three layers for tall vegetation (0-10, 10-100, and 100-250 cm) (Martens et al., 2017).

2.4 Analysis approach

2.4.1 Statistical analysis

The first part of the model evaluation focuses on, evaluating the monthly time series data of soil moisture products at the site level. The second part inter-compare modelled simulations and estimates at a regional level. Time series data for the sites were extracted from the soil moisture products, using the flux tower's geographical coordinates. The satellite products present averaged soil moisture data per grid cell at the centre of the grid cells, as opposed to the CCAM-CABLE model that provides four points per grid cell located at the edges. We therefore employed the distance weighted average (DWA) technique on the CCAM-CABLE model simulations to estimate soil moisture values representative of the sites in the grid cells, where the observation sites are located. The DWA method proved to be more representative than the nearest neighbour (NN) method, as the DWA method interpolates at the exact location of the tower in the grid cell, by taking into account all the points of the grid cell. It is noteworthy, that a comparison between the in-situ observations and satellite products, in this study, puts much emphasis on phase agreement as opposed to that of magnitudes. This is with regard to the fact that satellite observations, and GLEAM model estimates are represented as spatial averages for each pixel, in which case an interpolation of such aerial averages to a point (i.e. site), does not add any further information that correspond to the site. This also renders a comparison in magnitude to be rather unfair. However, we anticipate data at the point and grid scales should still comparatively present qualitative features that are characteristic of the climatic system for the region, for example seasonal cycles.

The soil moisture products are first converted to the percentage volumetric soil moisture amounts for comparison purposes. As in Yuan and Quiring (2017) we assume that the soil moisture measurements

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

at the depth of 5 cm represent the 0–10 cm depth. In situ data at the depths of 15, 30 and 40 cm are combined using the depth weighted average method to represent the 10-100 cm depth using Eq. (1):

$$SM_{10-100} = \sum_{i=1}^{n} \frac{LT}{SD} \times SM(i) \quad (1)$$

where SM_{10-100} is the weighted soil moisture, n is the number of layers, LT is the layer thickness calculated as the difference between the soil depths, SD is the total soil depth of the soil profile and $SM(i)$ is the daily in situ soil moisture values at the ith layer. Similarly the data at depths 2.2 and 5.8 cm; and 15.4 and 40.9 cm from CCAM-CABLE are averaged to represent 0-10 and 10-100 cm respectively using Eq. (1). Daily data from all the soil moisture products is averaged to monthly where 80 % of the daily data is available. Months that do not meet the 80 % threshold are excluded in the analysis. Monthly soil moisture data and seasonal cycles are evaluated against measurements.

Table 1. Overview of soil moisture datasets; satellite (grey) in percentage, modelled (blue), simulation (pink) and in situ observations (green) presented as a ratio ($m^3 m^{-3}$) of soil to moisture per unit area.

Soil moisture product	Spatial resolution (Km)	Spatial coverage	Soil depth (cm)	Period
ESA-Combined	25	Global	0-10	1978-2015
ESA-Active	25	Global	0-10	1991-2015
ESA-Passive	25	Global	0-10	1978-2015
CCAM-CABLE	8	Regional	2.2, 5.8, 15.4, 40.9, 108. 5, 287.2 (bedrock)	2000-2014
Skukuza	Point data	Point	5, 15, 30, 40	2000-2017
Malopeni	Point data	Point	5, 15, 30, 40	2008-2017
GLEAM v3a	25	Global	0-10, 10-100	1980-2016
GLEAM v3b	25	Global	0-10, 10-100	2003-2015
GLEAM v3c	25	Global	0-10, 10-100	2011-2015

To evaluate how similar the soil moisture simulation and model estimates are to in situ measurements, we use the stream flow plots, and the coefficient of determination (R^2), as defined in Koirala and Gentry (2012). Additionally, we use the covariance computed on the residuals of the de-trended time series to inter-compare soil moisture products at a regional (or grid) scale. The "stl" package in R is used to de-trend (or decompose) the time series into its components (i.e. seasonal, trend and residual) as discussed in Cleveland et al. (1990).

The "SoilGrids" dataset from the international soil reference information centre (ISRIC) is used in this study, to map soil types (Fig. 1c). The data is available online (https://soilgrids.org), and is described in detail in a study by Hengl et al. (2017). This dataset has a spatial resolution of 250 m and is used in this study to partition the covariance between simulated and modelled soil moisture according to soil type per grid box. Soil is classified into 12 dominant types ranging between sand and silty clay as described in Fig. 1d. The soil types data is available at seven depth (i.e., 0, 5, 15, 30, 60, 100 and 200 cm), here we only consider the data representing the surface (i.e., 0-5 cm). The 250 m dataset is resampled to 25 km, firstly by resampling to 1 km and then to 25 km, using the nearest neighbour method, to match the resolution of the soil moisture products. We acknowledge that resampling from fine to coarse resolution might introduce bias towards certain soil types. However, we believe that, the nearest neighbour method is suitable for resampling categorical data.

2.4.2 Cross-wavelet analysis

The cross-wavelet analysis as described in Rosch and Schmidbauer (2018) is applied in this study, and is computed in R using the "WaveletComp" package. The cross-wavelet method analyses the frequency structure of bivariate time series using the Morlet wavelet (Veleda et al., 2012). The wavelet method is suitable for analysing periodic phenomena of time series data, especially in situations where there is potential of frequency changes over time (Rosch and Schmidbauer, 2018; Torrence and Compo, 1998). This method has been used in other studies, such as those by Koirala and Gentry (2012), aiming to understand the climate change impacts on hydrologic response. Cross-wavelet analysis provides suitable tools to compare the frequency components of two time series, and thereby drawing conclusions about their synchronicity at a given period and time. A continuous wavelet leads to a wavelet transform of a time series which preserves information of both time and frequency resolution parameters. The transform can be partitioned into imaginary (Im) and real (Re) parts, which provide information on both the phase and amplitude over time. This is a prerequisite in the investigation of coherency between two time series (Rosch and Schmidbauer, 2018).

In a geometric sense the cross-wavelet transform is comparable with the covariance. Graphically the cross-wavelet spectrum provides the cone of influence and contour lines indicating significance of joint periodicity and for checks of consistency. Information on the synchronisation of two time series in terms of phase is also presented on the plot. Phase difference of the two time series at each time scale is given by:

$$Angle(T,S) = Arg\left(Wave.xy(T,S)\right) \tag{2}$$

This equals the difference of individual phases, $Phase.x - Phase.y$, when converted to an angle in the interval $[-\pi,\pi]$, this is indicated by arrows (Fig. 2) in the cross-wavelet power plot. The phase is computed using:

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

$$Phase(T,S) = \tan^{-1}\left(\frac{\mathrm{Im}(Wave(T,S))}{\mathrm{Re}(Wave(T,S))}\right) \quad (3)$$

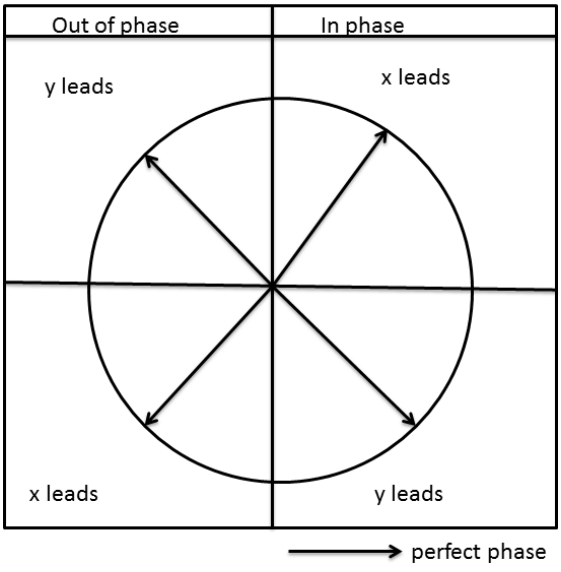


Figure 2. Phase interpretation between two time series x and y. When series x leads, y lags and vice versa. This figure is inspired by a study by (Rosch and Schmidbauer, 2018).

The cross-wavelet analysis in this study is performed between the CCAM-CABLE and GLEAM v3a products for periods between 2001 and 2014, since the technique only works on complete datasets (i.e., without missing values). The cross-wavelet is applied to non-stationary data using the default method (i.e., white noise) with the simulations repeated 10 times. In particular, in this study the cross-wavelet analysis is used to study the periodicity of the soil moisture signal as simulated by the CCAM-CABLE and estimated by GLEAM v3a, and to quantify the phase difference (i.e., lag) between the simulations and estimates.

2.4.3 Onset and offset of the wet period

In addition to analysis of phase agreement as discussed above, we compare the simulation of the onset, and offset of the wet periods by the different soil moisture products. This analysis is performed on daily data for soil moisture product with complete data (i.e., CCAM-CABLE and GLEAM). Instead of using precipitation as discussed in Shongwe et al. (2015) and Liebmann et al. (2007), to identify the onset and offset of the rainy season, we use soil moisture data. This is computed using a cumulative quantity over time as

$$A(day) = \sum_{n=1}^{day}\left[S(n) - \bar{S}\right] \quad (4)$$

where $S(n)$ is the daily soil moisture and $\bar{S}$ denotes the annual daily average. As in Liebmann et al. (2007) we start the calculation on the climatologically driest month, i.e., 1 July (Shongwe et al.,

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

2009), and perform a cumulative sum over a period amounting to a year. The onset of the wet period is defined as the date on which the cumulative sum reaches a minimum, and the offset, as the date on which the cumulative sum reaches the maximum (Shongwe et al., 2015).

3 Results and discussion

3.1 Evaluation of seasonal soil moisture

In this section we discuss how the respective outputs from CCAM-CABLE, ESA and GLEAM reflect the key features of soil moisture for the study sites. As highlighted in the introduction section, the variability of the simulation output, satellite derived data and satellite based model estimates is studied relative to the in situ measurements from the study sites. Much focus is placed on understanding how well the seasonality of the soil moisture is reflected by respective soil moisture data sets. It has also been mentioned in the introductory section that the in situ observations are taken from a semi-arid (i.e., Skukuza), and arid (i.e., Malopeni) savanna sites within the Kruger national park. The patterns of soil moisture at these sites are mainly driven by rainfall which is generally high during the summer season, and low in winter as shown in Fig. 3. The long term surface soil moisture at the respective sites, follow a pattern comparable to that of rainfall. For example, see the soil moisture patterns presented in Fig. 4 (i.e. long term cycles), and the monthly rainfall accumulation (Fig. 3).

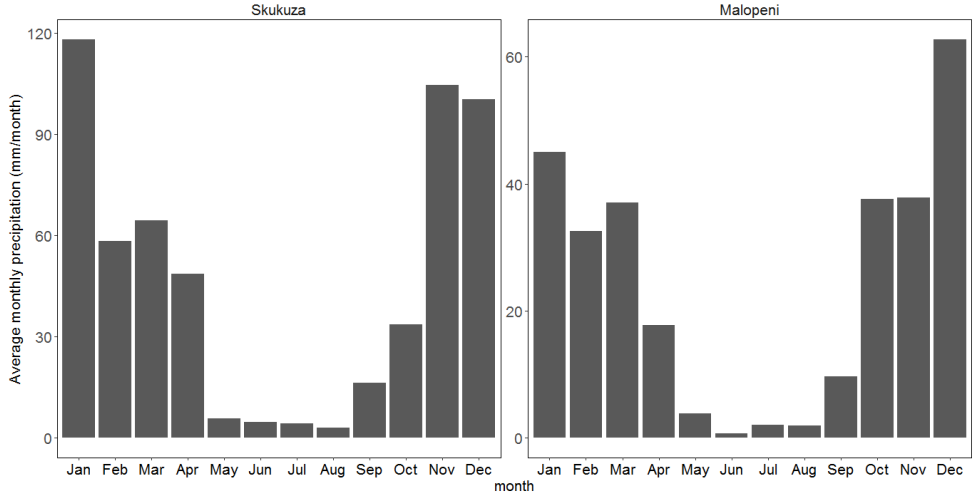


Figure 3. Long term measurements of average monthly precipitation (mm month^{-1}) at Skukuza (2001-2014) and Malopeni (2008-2013) flux towers respectively, the average is computed using daily data with at least 80% available data.

3.1.1 Long term seasonal cycles

In general the pattern of the long term average for soil moisture from CCAM-CABLE simulation, ESA satellite observations and GLEAM model estimates are qualitatively comparable to that of in situ observations. Notably, the observed soil moisture signal, for both the surface and root zone, shows an increase in April and October. This is found to be consistent between both at the surface and root zone, whereas the observed rainfall signal shows a similar pattern but in March and November at both sites. This is potentially a signature of soil moisture retention, which relates to the persistence of dry

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

and wet periods at various soil depths (Seneviratne et al., 2006). In light of this, it is also interesting to see how both the simulation and the GLEAM products depict the onset and offset of the wet season, and such a discussion will be dealt with in section 3.2. It is also worth noting that the coupled CCAM-CABLE model does not capture an increase in April strongly for Skukuza while it is completely missed for the Malopeni site. This is probably due to the fact that the CABLE hydrological scheme, within the version of cable used in this study, does not take into account soil resistance (Whitley et al., 2016), this has strong implications on how well the simulation can be comparable with in situ observations in terms of magnitude. Despite this, the long term CCAM-CABLE monthly means of soil moisture are relatively comparable to in situ observation even on magnitude this can be seen in Fig. 4. GLEAM v3c, on the other hand, agrees with in situ measurements on the timing of the April soil moisture rise, but it reflects the November increase in soil moisture, a month earlier (i.e., October). The satellite products (i.e., the active, passive and combined ESA products) and GLEAM models (Fig. 4) displays the same signal as that of the observed soil moisture, indicating that this does not occur by chance. We can safely deduct that the bias in GLEAM v3c is potentially not induced by satellite based forcing data, however this calls for further investigations on the sensitivity of the model to its driving data at a high resolution. We anticipate that at high temporal resolution there is a strong variability on the in situ soil moisture signal which may not entirely be captured, by both CCAM-CABLE and GLEAM due to their relatively low spatial resolution. The low resolution (8 km) in the case of CCAM-CABLE, in particular, potentially has strong implications on how representative the effective drivers of soil moisture such as soil texture and vegetation covers are for the specific sites.

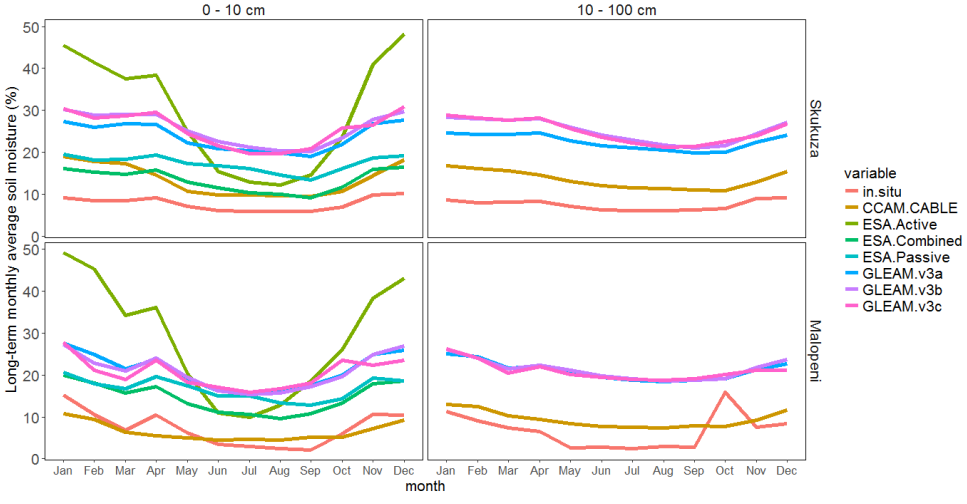


Figure 4. Seasonal variation in long term mean monthly surface (i.e., 0-10 cm) and root zone (i.e., 10-100 cm) soil moisture, based on in situ observations and a variety of soil moisture products. The in situ data is collected from two Skukuza (2001-2014) and Malopeni (2008-2013).

Soil moisture is at its lowest during the dry periods (i.e., May to October) and high during the wet (i.e., November to April) periods. The GLEAM models (Fig. 4) are generally consistent with in situ measurements in estimating soil moisture both in terms of magnitude and phase, both at the surface and root zone. The magnitude of GLEAM v3a root zone estimates is lower than those of the other GLEAM models at the Skukuza site. This can be attributed to the unique multi-source weighted ensemble precipitation (MSWEP) data used to force GLEAM v3a (Martens et al., 2017), which is different to the precipitation forcing data used in GLEAM v3b and GLEAM v3c. We further observe

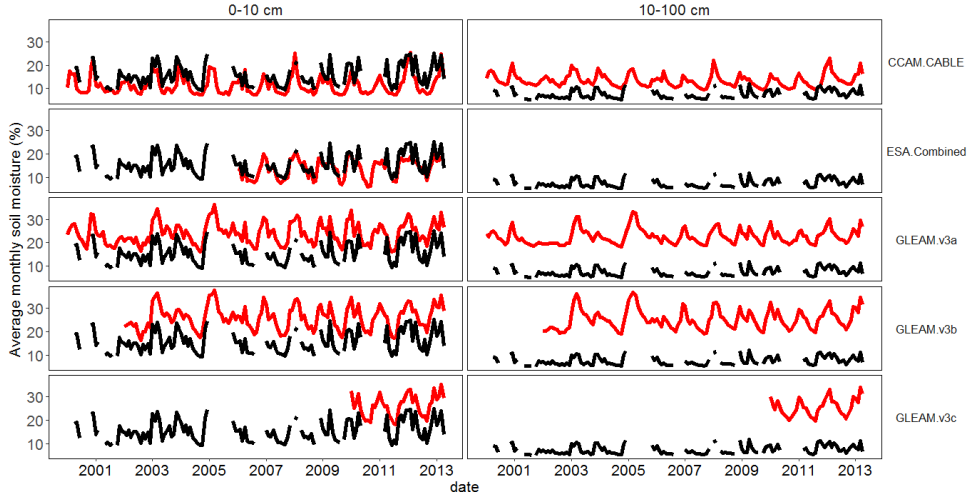
that the GLEAM models, satellite and in situ observations have the same length of the dry period, with the exception of the ESA-Active observation which has a shorter dry period. The ESA-Active satellite product is known to work best for moderate to densely vegetated areas as opposed to savanna sites such as Skukuza and Malopeni where tree cover is sparse (Dorigo et al., 2015) and the vegetation cover changes dynamically due to a combination of factors for example fires and rainfall There is less difference between the ESA-Passive and ESA-Combined satellite products. Generally the ESA-Combined and ESA-Passive have the least difference during the dry period for all sites. The ESA-Combined product shows a strong increase in soil moisture in July which is not observed in the other soil moisture products. Looking at long term averages, it is still interesting to see that both the CCAM-CABLE and GLEAM models are able to capture the intrinsic seasonality of the soil moisture signal for the sites as reflected by both in situ and satellite observations. This is despite their being different both in the forcing data and model structure. Studies by (Wang and Franz, 2017) and (Seneviratne et al., 2010) suggest that local factors (e.g., vegetation, soil and topography) mostly control soil moisture variability at spatial scales lower than 20 km than meteorological forcing. For a fourteen year averaging period, undoubtedly the monthly means are sensitive to anomalously high precipitation, and hence soil moisture, in some months. It is therefore instructive to study how well the simulated and estimated patterns of soil moisture compare to the in situ data on a monthly basis for the respective years.

3.1.2 Short term seasonal cycles

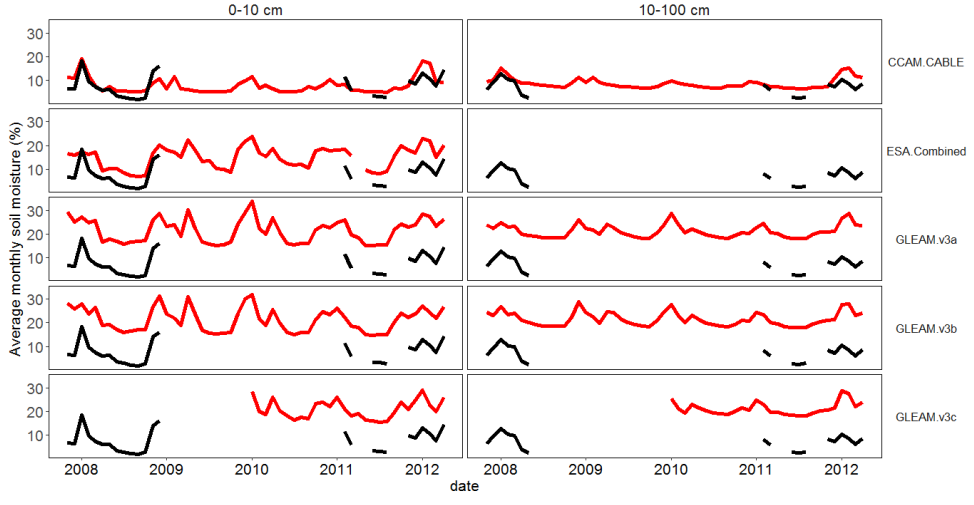
This section presents both a qualitative and quantitative evaluation of the soil moisture signal from CCAM-CABLE simulation and GLEAM estimate at a monthly time scale. To circumvent possible bias that may emanate from missing values, monthly averages are presented for only months where there are observations above 80% data availability threshold. This implies that the number of data points (i.e., sample size), are not equal for all products. For example, GLEAM v3c has the shortest data set spanning between 2011 and 2014, as opposed to other data sets that range between 2001 and 2014 for Skukuza, and 2008 to 2013 for Malopeni. Figures 5 and 6 for Skukuza and Malopeni respectively show that, there is generally a phase agreement which is also noted for long term averages in situ data and the soil moisture products. This is reflected for most of the years. In consistency with the long term pattern, discussed in section 3.1.1, there is a strong agreement in phase between the in situ surface soil moisture signal and that of the simulation, satellite derived products and satellite based model estimates. We further observe that in some instances, there is a lag between the simulations and observations. At the root zone, we observe a decrease in phase similarities between the simulation and model estimates and observations. Clearly this calls for further investigation into water drainage and soil moisture memory which is outside the scope of the discussion in this study. Furthermore, the soil moisture products compare best to in situ data, than they do to CCAM-CABLE, see appendix A Fig. A1 for Skukuza and Fig. A2 for Malopeni.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

Hydrology and Earth System Sciences
Discussions



441 **Figure 5.** Comparison of monthly averaged modelled and satellite data (red line), with in situ observations (black line) of surface (0-10 cm), and root zone (10-100 cm) soil moisture at the Skukuza (2001–2014) site. The gaps in the observations represent months, where the daily data did not meet the 80% averaging threshold.



444

Figure 6. Comparison of monthly modelled and satellite data (red line) with surface (0–10 cm) and root zone (10–100 cm) soil moisture in situ observations (black line) at the Malopeni (2008–2013).

447 On account of missing values, the R^2 values presented in Fig. 7 are based on different sample sizes, therefore, their interpretation is done with this issue in mind. The R^2 values are generally high at Malopeni compared to Skukuza, indicating that the few months where there are observations, there is
450 also a high comparability of the signal. It is however, inconclusive whether the simulations and estimates are more comparable at Malopeni relative to the case in Skukuza. In general, based on Fig. 7, we learn that, all the soil moisture products are able to capture the variability in the observed soil
453 moisture, mainly exceeding the R^2 value of 0.5 (i.e. 50%) both at the surface and root zone. The CCAM-CABLE model mainly presents the lowest, but acceptable R^2 both at the surface and root

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

zone, this is a further reflection that the simulated soil moisture signal is high in magnitudes compared to in situ observations. On the other hand, we find the agreement in the seasonality of CCAM-CABLE interesting in its own right regarding that the model is free running (i.e., it is not initialised with an observed initial state.

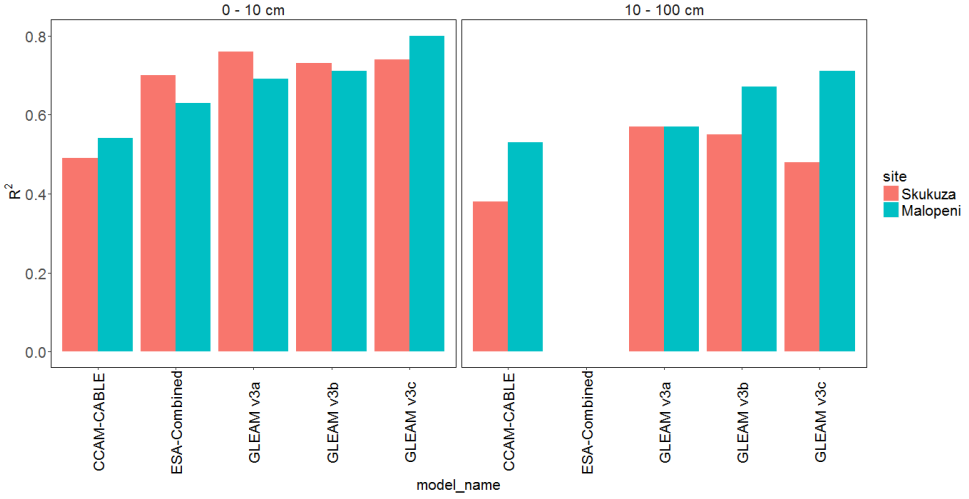


Figure 7. Quantitative comparison between soil moisture products and observations at Skukuza (Red bars) and Malopeni (Green bars), at the surface (0-10 cm) and root zone (10-100 cm), using the coefficient of determination (R^2).

The ESA-combined satellite product is expected to present the best agreement with observations, since it is actual observed data from the sites, this expectation is realised with R^2 values greater than 0.65. Furthermore, the ESA data has been shown to generally capture soil moisture in different regions and climate zones of the world (Loew et al., 2013; McNally et al., 2016; Wang et al., 2016; Zeng et al., 2015). Our study confirmed in Figures 5-7 that the ESA combined product captures local conditions within reason/acceptable amount of certainty. In a study conducted by Yuan and Quiring (2017), assessing the performance of CMIP5 models, both at the surface and root zone concluded that the models performed well at the root zone, compared to at the surface. This is contrary to what we observe in this study, we generally observe better comparison between soil moisture products and in situ measurements at the surface than at the root zone. Based on the extent to which GLEAM products proved to be representative of the qualitative features of the soil moisture signal for different months and seasons, as driven by precipitation at the site, it is compelling to further resolve qualitatively, how the simulated output compare against each other for most of the time periods. To this effect, we next present the results from a cross-wavelet analysis of CCAM-CABLE simulation output and GLEAM estimates for the two study sites.

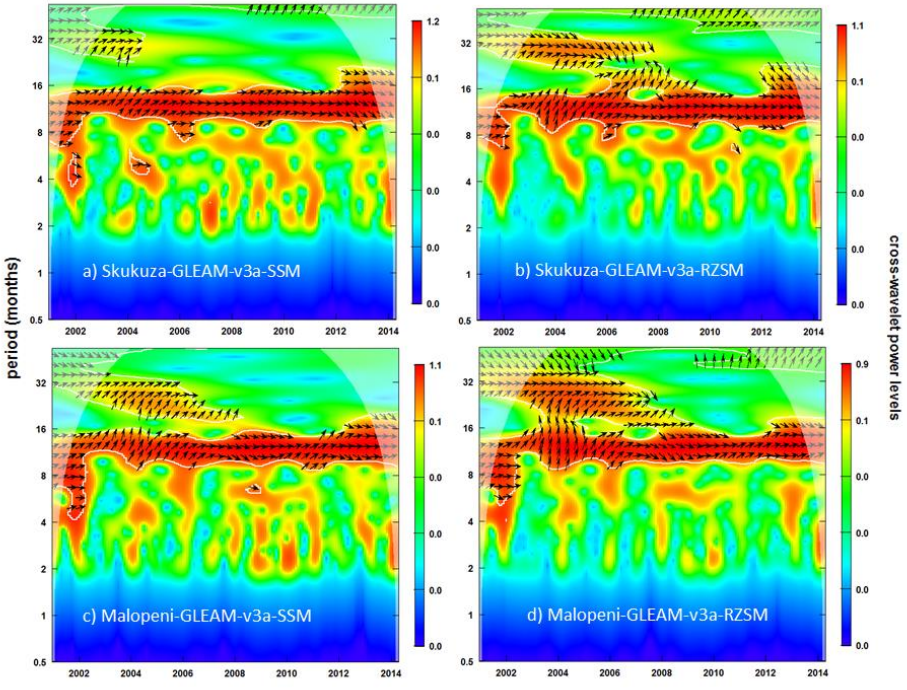
3.2 Cross-wavelet analysis

The cross-wavelet power spectrum (Fig. 8) reveals that, generally the time series of CCAM-CABLE simulations and GLEAM v3a soil moisture estimates are in phase, both at the respective sites and the soil depths investigated. This is indicated by the arrows generally pointing to the right, as illustrated in Fig. 2. The arrows are plotted between the white contour lines indicating areas of significance, and joint periodicity at 10 % (i.e., 90 % confidence level). This area of significance is generally between

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

Hydrology and Earth System Sciences
Discussions

the periods of 8 and 15 months (y-axis). Although the time series are in phase most of the time, in some instances there is a lag. This is identified by the direction of the arrows, in which case the
486 arrows are inclined upwards or downwards at different margins. For example, between the year 2002 and 2005 periods of about 12 months, we observe that GLEAM v3a leads and CCAM-CABLE lags by 6 days on average at the surface at Skukuza (Fig. 8a) and by 10 days at Malopeni (Fig. 8c). The
489 cross correlation also shows that, there are other cyclical responses of the soil moisture signal with a periodicity of approximately two years. This becomes apparent for some years when The CCAM-CABLE and GLEAM v3a signal have statistically significant periodic features which repeat after 28
492 months. In this case CCAM-CABLE leads GLEAM v3a by 5 days on average. A plot of the phase differences, between CCAM-CABLE (red) and GLEAM-v3a (blue), for the years 2002 to 2014 associated with soil moisture patterns with a characteristic period of 12 months, is illustrated in Fig.
495 B1 in appendix B. Between 2007 and 2011 we note that there is no lag between the two series, particularly for the annual cycles i.e., repeating features of the respective signals which repeat between 11 and 13 months for the surface.



498

Figure 8. Cross wavelet power spectrum of surface (SSM, 0-10 cm) and root zone (RZSM, 10-100 cm) soil moisture between CCAM-CABLE, and GLEAM v3a at Skukuza (a, b) and Malopeni (c, d) respectively. The
501 white contour lines indicate periods of significance at 10 %. The arrows pointing to the right indicates that the models are in phase, anti-phase point left, CCAM-CABLE leading GLEAM v3a is indicated by arrows pointing straight down. The dome shape (shaded areas) represents the cone of influence between 2001 and 2014.

504 At the root zone we see a similar pattern as that of the surface soil moisture. Most statistically significant shared periodic feature between CCAM-CABLE and GLEAM v3a, on the soil moisture signal, have periods mainly between 10 and 16 months. This is true for the entire time series (i.e.,
507 2001-2014). The cross-wavelet analysis in this case picks the characteristic annual pattern of soil

moisture which is effectively repeated for different years. The time series are in phase for the whole analysis period generally without any lag between 2007 and 2012 for periods ranging between 9 and 15 months. For the feature of the signal with a 12 month period, there is on average a time lag of 10 and 19 days at Skukuza (Fig. 8b) and Malopeni (8d) respectively. We further note that the significant periodic features of the signal generally increase from the surface to root zone. This is potentially associated with differences in the drivers of soil moisture between the respective layers. Root zone soil moisture, for instance, is likely to respond to plants driven moisture demands in a slightly different manner in comparison to the surface layer. An accurate attribution of soil moisture patterns per layer to the respective drivers, in this context, is a rather complex problem and demands a separate investigation. The simulation and the models estimates show coherency in capturing periodic patterns, at least those that are recurrent on an annual time scale. The cross-wavelet analysis successfully reveal that there is a similarity in the patterns of surface and root zone soil moisture over time at both sites albeit negligible differences across-sites for events that are recurrent on periods below or exceeding 12 months. The existence of a time lag or differences in phase in the soil moisture signal between the simulation and GLEAM model outputs, is likely a result of CCAM-CABLE being a free run model as highlighted in the section 3.1.2. A simulation model forced with a non-observed climate states is still anticipated to capture most of the characteristic feature of a climatic systems such as the seasons. Its output may not match satellite derived observations on certain aspects including inter-annual variability. Next we will explore how these differences reflect on the onset and offset of the wet period calculated from the simulation and GLEAM models.

The results in Fig. 4 indicated that the modelled and satellite derived soil moisture products generally capture the length of both the dry and the wet period. However, Fig. 8 shows that there is a lag between the time series of CCAM-CABLE and GLEAM estimates which indicate uncertainty in phase agreement. Figure 9 show that the GLEAM models have a relatively low uncertainly for the onset and offset of the wet period. This is expected as these models use similar forcing data. For example, GLEAM v3a and v3b agrees on the onset of the wet period during the following years; 2003, 2004, 2005, 2009, 2013 and 2014. Furthermore, GLEAM v3a and v3b agrees on the offset of the wet period during 2010 and 2011. The CCAM-CABLE and GLEAM products predominantly differ by a factor not exceeding 30 days on the timing of the onset the wet period. There is a very noticeable uncertainty in the timing of the cessation of the wet period among all approaches. These analysis yield results that are consistent with those observed in Fig. 4. The study sites mainly experience summer rainfall, commonly occurring between November and April. The CCAM-CABLE model generally shows a consistent the length of the wet period at both study sites for most of the years. The GLEAM models generally present an early onset in October and offset in May. This is consistent with the difference in phase between CCAM-CABLE and other products presented in Fig. 8.

Looking at the agreement between in situ observation, CCAM-CABLE simulation output and the GLEAM model estimates, when it comes to the main periodicity of the soil moisture signal portrayed in Fig. 8 as well as the results of the onset and cessation of the wet period Fig. 9, we find it indisputable that the two modelling approaches are representative of the key features of the soil moisture signal. It is also interesting to note that the level of uncertainty between the two modelling approaches, as reflected by the onset of the wet period in Fig. 9, is within an acceptable level i.e., it predominantly lies within days not exceeding a month. The uncertainty gets more pronounced when it comes to the cessation of the wet period. This is indicative of differences in inter-annual variation of the soil moisture signal which is expected, to a certain extent, due to the different input data used and the mathematical structure of the models. Clearly there is need for an understanding of how the noted

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

Hydrology and Earth System Sciences
Discussions

uncertainty could be attributed to various factors from forcing data or soil moisture drivers. It would be very important to understand, in particular, as to how much uncertainty is inherent in the individual coupled model key components. An in-depth investigation of various sources of model uncertainty is indeed a topical issue (Fang et al., 2016) which deserves a lot of attention but such a discussion will not be dealt with in this study. It is interesting to establish whether the insight gained in understanding the level of inter-comparability of the soil moisture signal, at the two respective sites, will hold at the regional level i.e., we want to know if the mutual agreement between simulation and model estimates persists for the study regions indicated in Fig. 1c. A natural starting point is to look at the covariance between the simulated and estimated soil moisture signal for the region.

Figure 9. Onsets (triangles) and offsets (circles) of the wet period at Skukuza and Malopeni as simulated by; CCAM-CABLE (2001-2014, red), GLEAM-v3a (2001-2014, green), GLEAM-v3b (2003-2014, blue) and GLEAM-v3c (2011-2014, purple).

3.3 Regional inter-comparison

In Fig. 10, a plot of covariance between soil moisture from CCAM-CABLE simulations outputs and GLEAM model estimates is portrayed. The covariance is computed from the residuals of the de-trended series of both CCAM-CABLE and GLEAM models. For this analysis only data from 2011 to 2014 is used, as it is common between all the soil moisture products. We generally observe in Fig. 10 that at the surface (SMsurf, 0-10 cm), the covariance between the soil moisture CCAM-CABLE and GLEAM is high, compared to the root zone (RZSM, 10-100 cm), implying that shared information between GLEAM and CCAM-CABLE is predominantly more pronounced at the surface compared to the root zone. This signals differences in the representation of soil moisture drainage at the root zone between the simulation and satellite data based model estimates. In order to see if there are any major differences in the simulation and GLEAM models estimates that can be associated with differences in soil types, we further partition the covariance between the models at various soil types (Fig. 1c) of the grid. The results are presented in Fig. 11.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

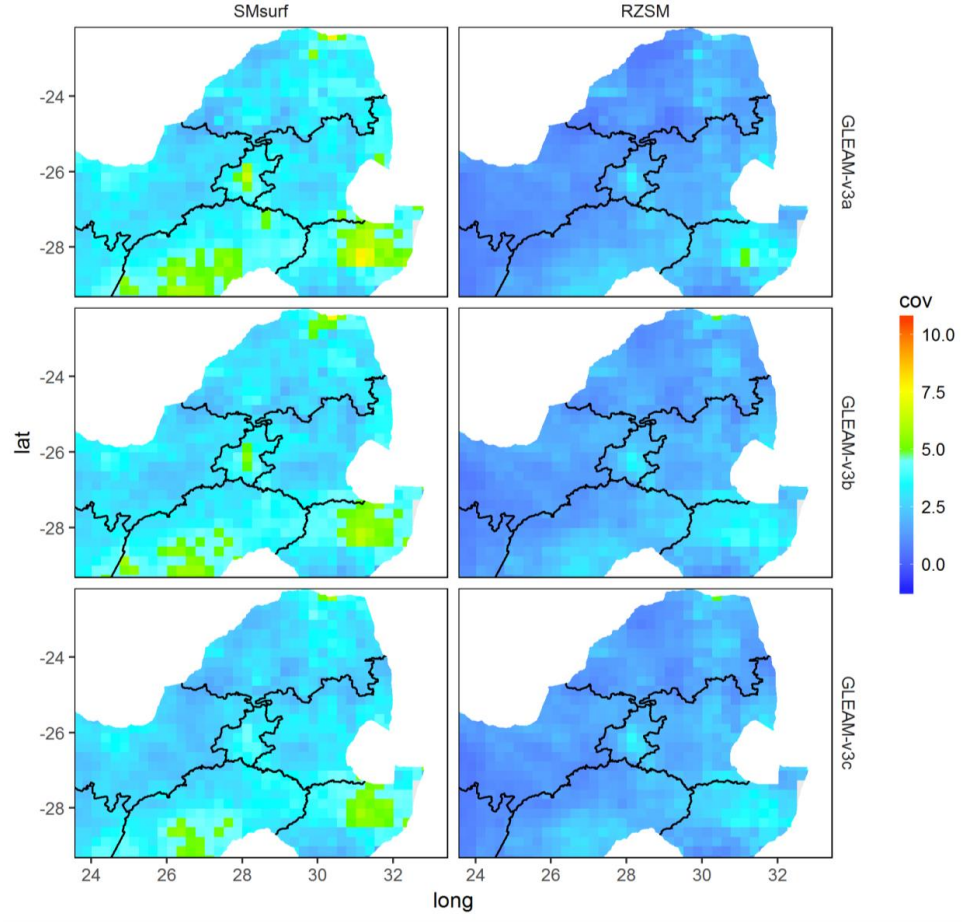


Figure 10. Covariance (cov) computed on the residuals of monthly time series (2011-2014) of surface (SMsurf, 0-10 cm) and root zone (RZSM, 10-100 cm) soil moisture, between CCAM-CABLE simulations and GLEAM models estimates.

The dominant soil types in the region include loam, silt and clay (Fig. 1c). The dominant soil types 5 (loam) and 6 (silt loam); and soil types 8 (silty clay loam) and 9 (clay) as presented in Fig. 1d are associated with the grassland and savanna biomes respectively, as shown in a study by Stevens et al. (2015), the grass land and savanna biomes are dominant in this study area.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

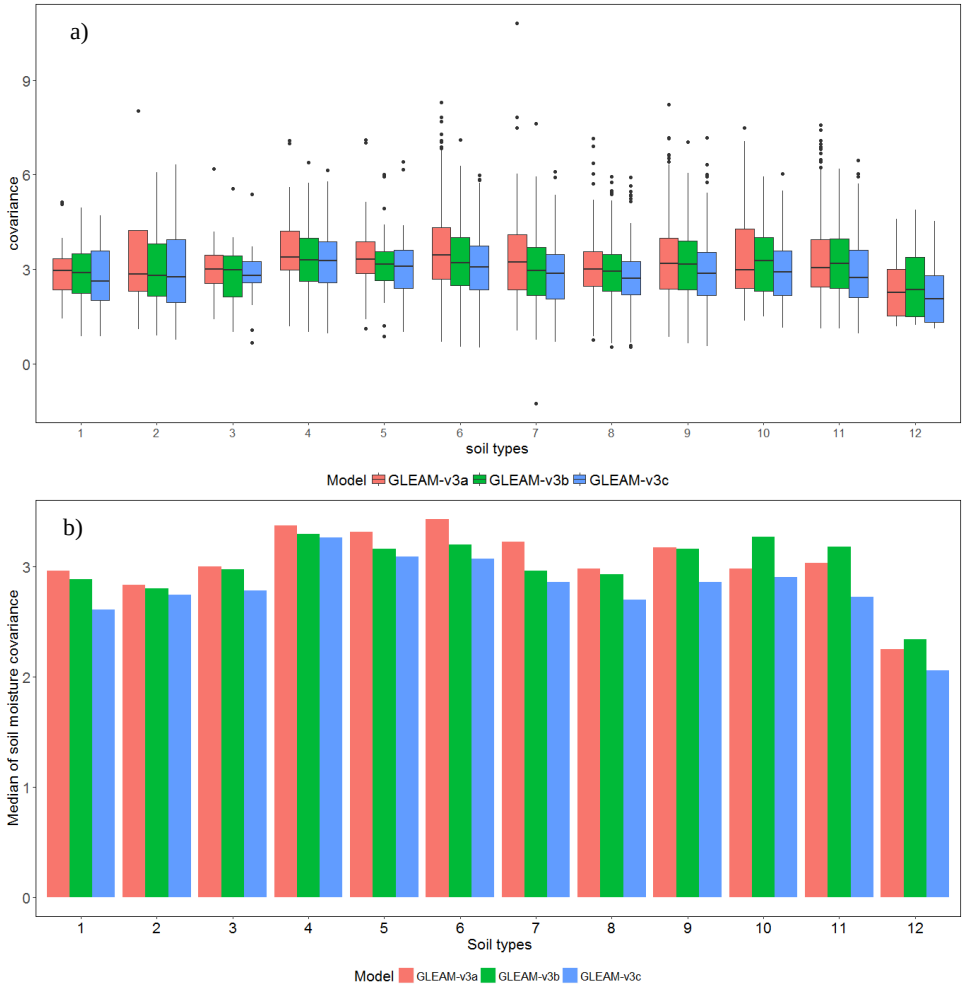


Figure 11. The covariance of soil moisture, per soil type computed between the residuals of CCAM-CABLE and GLEAM-v3a (red), GLEAM-v3b (green) and GLEAM-v3c (blue) soil moisture products, a) boxplots showing the spread of the covariance per model and soil type, the horizontal lines in the box plots represent the median, and b) bar plots of the median of the covariance.

The spread of the covariance, as grouped by soil types, is presented in Fig. 11a. In general, the covariance between the residuals of CCAM-CABLE and GLEAM models is positive. In particular, the spread of the soil moisture covariance generally ranges between 2 and 4. This is indicative that there exists mutual information or joint variation even on time scales shorter than seasons between the respective signals. Based on the inter-quartile range (i.e., height of the bars of the boxplots) we observe that there is a pronounced variability in soil moisture covariance between CCAM-CABLE and all the GLEAM products across different grid cells. However, across the soil types the spread is mostly comparable. The comparability is in the sense that all box-and-whisker plots have appreciable overlaps within the inter-quartile range. This indicates comparable spatial uncertainty on the covariance among soil types and models. Clearly, the averaging done for the soil classes also introduce a certain level of soil moisture covariance. It is therefore instructive to look at the patterns

central tendency as reflected by the median covariance of Fig. 11b. The covariance of the median values is a preferable measure as it is sensitive to outliers (i.e., values beyond the whiskers ends), as opposed to the mean per soil category.

It is worth noting that median covariance between CCAM-CABLE and all the GLEAM models outputs occur mostly at the dominant soil types (i.e. 4-sandy clay; 5-loam; 6-silt loam and 7-silt). This is indicative that there is a fair amount of data points that lie further apart from their associated grid cell mean. This could mean that the respective distributions, for a specific grid cell whose covariance is calculated, have comparable synchronous points that lie apart from the mean or one of the soil type distributions is having such outlying points. The latter is likely to be predominant in the case where there are some time lags as demonstrated in Fig. 8. This alludes to differences in the representation of inter-annual variation by the simulation and GLEAM model estimates as highlighted in section 3.1. This indicates that despite the fact that there is joint variation between the simulation and GLEAM models there exists non-negligible variability within these respective soils types which can potentially be uncovered by studying responses to various soil moisture drivers as modelled through the respective approaches. As mentioned earlier, these dominating soil types are generally important for agricultural purposes which make it very relevant for further investigations. The least pronounced soil class covariance mean between the simulation and GLEAM models occurs in soil type 12 (silty clay). This is one of the least dominating types. In this case there are very few grid-points for a meaningful comparison against other categories. In summary, from the regional covariance calculation we learn that there is existence of joint variation at short time scales between the simulation and GLEAM estimates. Apart from this, there is a fairly modest level of model uncertainty between CCAM-CABLE and GLEAM products, which is comparably reflected across all soil types. We strongly feel that attribution of its inherent sources merit a further investigation.

4 Conclusions

In this study, the ability of a process based simulation model (CCAM-CABLE), satellite data driven models estimates (GLEAM) and satellite observations (ESA-Active, -Passive and -Combined) are evaluated against site specific in situ observations from two flux tower sites namely, Skukuza and Malopeni. The sites are situated within the Kruger national Park in South Africa. The evaluation is done for two soil depths namely; surface (SSM, i.e., 0-10 cm) and root zone soil moisture (i.e. 10-100 cm) with the objective of understanding how the respective data products capture characteristic patterns of soil moisture within a 25 km grid box that enclose the study sites. The evaluation includes an assessment of qualitative features of long term (i.e. multi-year) and short term (i.e., monthly) averages of the soil moisture signal relative to in situ measurements from each of the two flux tower sites. We learn that all the soil moisture products at all depths present higher magnitudes of soil moisture compared to observations, though the simulation output at Malopeni flux site, which is closer to observations in magnitude. The difference in magnitude may be attributed to difference in length scale between in situ measurements and the rest of the products. The study therefore placed much focus on features of the soil moisture signal which may be attributed to as responses to generic influences of the climatic systems for the region. The coefficient of determination (R^2), however reveals that most of the soil moisture products for the sites have an appreciable level of similarity (mostly $R^2 > 0.5$) at all depths. A qualitative analysis of the time averaged soil moisture signal, for all the products, indicates that satellite observation and satellite based model estimates capture most of the inter-annual structure of the soil moisture signal. We also learn from this study that all GLEAM models compare well with the in situ observation in reflecting the seasonality of soil moisture. In particular, the products portray that there is a more pronounced soil moisture, for the Southern

hemisphere during the period October-March relative to May-September. It is therefore recommendable that satellite derived model estimates be used as surrogate observations for forcing process based models. In particular, this can be done for instances where inter-annual variability as well as seasonal patterns of the soil moisture are of particular interest.

The CCAM-CABLE simulation model effectively represents the seasonality of the soil moisture signal for the sites, however the simulation model fails to reflect short time scale effects such as the rise in soil moisture observed in April and November. The simulation strength in reflecting the changes in soil moisture across seasons demonstrates that it could be used to test the implications of long term land cover changes on soil moisture patterns. There is however, a need for further investigation into the sensitivity of the simulation model's hydrological scheme to changes in the soil moisture drivers within inter-annual time scales.

The study also investigated the level of uncertainty between GLEAM models and the CCAM-CABLE simulation. In particular the wavelet analysis is used to reveal, at a qualitative level, how periodic features of the soil moisture signal compared between the simulation and the estimates produced by GLEAM models. In this case, the emphasis is on evaluating the extent to which both approaches have a joint variation or shared mutual information. The analysis has successfully revealed that both the simulation and model estimates equally reflect the periodic seasonal pattern of soil moisture, however there is a predominant time lag between GLEAM products and CCAM-CABLE. The time lag is mostly of a time scale not exceeding a month at all soil depths (i.e., it mostly lies between 5 and 20 days) during the studied years (2001-2014). We conclude that the major difference in the long term and short term feature of the soil moisture signal, between CCAM-CABLE and GLEAM models estimates can be attributed to, among other factors, their difference in capturing inter-annual patterns of the soil moisture signal. This is also supported by the existence of a non-negligible level of uncertainty on the onset and offset of the wet period which is calculated for the CCAM-CABLE and all GLEAM models outputs.

Despite the existence of uncertainty, we affirm that there is appreciable mutual information on the soil moisture signal from the simulation and GLEAM models. This is also reflected by the regional patterns of the covariance between the CCAM-CABLE and GLEAM-v3c signal. The covariance of soil moisture between CCAM-CABLE and GLEAM is obtained from the residuals of the de-trended soil moisture time series. It is found to be mostly positive for all soil and vegetation types. Looking at the spread of the covariance values within the study region, as well as their associated median values as grouped by soil types, we conclude that the extent of the shared features is not limited to the seasonal time scale. The covariance does not vary too strongly among the dominating soil types. In general the covariance matric ranges between 2 and 4.

The difference in the soil moisture signal structure, at inter-annual time scale, between the simulation and GLEAM models, opens-up an interesting question relating to the extent to which the influence of different drivers of soil moisture is represented by the studied simulation and estimation approaches. To understand this, future research will benefit from investigating the influence of changes in soil moisture drivers, particularly change in vegetation cover and soil type, on soil moisture memory. It will also be interesting to unearth the effects of extreme weather and climate change induced pattern on the long term soil moisture pattern persistence. In this regard, it would be interesting to uncover the tipping or breaking points of trends in soil moisture. To this effect, we find the GLEAM products soil moisture patterns worth further investigation using various statistical approaches including machine and deep learning algorithms to gain a deeper understanding of soil moisture response to climatic and land management related effects.

Team list and Author contribution:

- Floyd – Developed research questions, analysed the data and compiled the manuscript.
- Marna and Mohau –suggested datasets to be explored, reviewed the manuscript, and made inputs on data analysis approaches and research questions formulations.
- Gregor – Provision of in situ data, inputs of formulation of research questions and critical review of the manuscript.
- Francois – Led the CCAM-CABLE model simulations runs, and introduced the lead author to the model structure and the dynamical downscaling methods.
- Michael – Supervisory role and manuscript review.

Competing interests: The authors declare that they have no conflict of interest.

Disclaimer:

Acknowledgements: This work was funded by the EEGC030 project of the CSIR. The authors wish to acknowledge Ms Humbelani Thenga and Dr Marc Pienaar for their contributions.

Data availability:

In situ data

Data from the CSIR owned flux towers (i.e. Skukuza and Malopeni) can be requested from Ms Humbelani Thenga (HThenga@csir.co.za).

ESA CCI

The data is available online from www.esa-soilmoisture-cci.org/

GLEAM

The data is available online from www.gleam.eu

References

Albergel, C., de Rosnay, P., Gruhier, C., Munoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y. and Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sens. Environ., 118, 215–226, doi:10.1016/j.rse.2011.11.017, 2012.

An, R., Zhang, L., Wang, Z., Quaye-Ballard, J. A., You, J., Shen, X., Gao, W., Huang, L. J., Zhao, Y. and Ke, Z.: Validation of the ESA CCI soil moisture product in China, Int. J. Appl. Earth Obs. Geoinf., 48, 28–36, doi:10.1016/j.jag.2015.09.009, 2016.

Archibald, S. A., Kirton, A., van der Merwe, M. R., Scholes, R. J., Williams, C. A. and Hanan, N.: Drivers of inter-annual variability in net ecosystem exchange in a semi-arid savanna ecosystem, South Africa, Biogeosciences, 6(2), 251–266, doi:10.5194/bg-6-251-2009, 2009.

Brocca, L., Melone, F. and Moramarco, T.: Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting, Hydrol. Process., 25(18), 2801–2813, doi:10.1002/hyp.8042, 2011.

Cleveland, R. B., Cleveland, W. S., McRae, J. E. and Terpenning, I.: STL: A seasonal-trend decomposition procedure based on loess, J. Off. Stat., 6(1), 3–73, doi:citeulike-article-id:1435502, 1990.

Decker, M.: Development and evaluation of a new soilmoisture and runoff parameterization for the CABLE LSM including subgrid-scale processes, J. Adv. Model. Earth Syst., 7, 513–526, doi:10.1002/2015MS000507, 2015.

Dirmeyer, P. A., Jin, Y., Singh, B. and Yan, X.: Trends in land–atmosphere interactions from CMIP5 simulations, J. Hydrometeorol., 14(3), 829–849, doi:10.1175/JHM-D-12-0107.1, 2013.

Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M. and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, Remote Sens. Environ., 162, 380–395, doi:10.1016/j.rse.2014.07.023, 2015.

Engelbrecht, F., Adegoke, J., Bopape, M. J., Naidoo, M., Garland, R., Thatcher, M., McGregor, J., Katzfey, J., Werner, M., Ichoku, C. and Gatebe, C.: Projections of rapidly rising surface temperatures over Africa under low mitigation, Environ. Res. Lett., 10(8), 1–16, doi:10.1088/1748-9326/10/8/085004, 2015.

Fang, L., Hain, C. R., Zhan, X. and Anderson, M. C.: An inter-comparison of soil moisture data products from satellite remote sensing and a land surface model, Int. J. Appl. Earth Obs. Geoinf., 48, 37–50, doi:10.1016/j.jag.2015.10.006, 2016.

Feig, G. T., Mamtimin, B. and Meixner, F. X.: Soil biogenic emissions of nitric oxide from a semi-arid savanna in South Africa, Biogeosciences, 5(6), 1723–1738, doi:10.5194/bg-5-1723-2008, 2008.

Fischer, E. M., Seneviratne, S. I., Vidale, P. L., Lüthi, D. and Schär, C.: Soil moisture-atmosphere interactions during the 2003 European summer heat wave, J. Clim., 20(20), 5081–5099, doi:10.1175/JCLI4288.1, 2007.

Hengl, T., De Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S. and Kempen, B.: SoilGrids250m: Global gridded soil information based on machine learning, PLoS One, 12(2), 1–40, doi:10.1371/journal.pone.0169748, 2017.

van den Hurk, B., Doblas-Reyes, F., Balsamo, G., Koster, R. D., Seneviratne, S. I. and Camargo, H.: Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe, Clim. Dyn., 38(1–2), 349–362, doi:10.1007/s00382-010-0956-2, 2012.

Jovanovic, N., Mu, Q., Bugan, R. D. H. and Zhao, M.: Dynamics of MODIS evapotranspiration in South Africa, Water SA, 41(1), 79–91, doi:10.4314/wsa.v41i1.11, 2015.

Kowalczyk, E. A., Wang, Y. P. and Law, R. M.: The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model, CSIRO Mar. Atmos. Res. Pap., 13(November 2015), 1–42, doi:1921232390, 2006.

Liebmann, B., Camargo, S. J., Seth, A., Marengo, J. A., Carvalho, L. M. V., Allured, D., Fu, R. and Vera, C. S.: Onset and end of the rainy season in South America in observations and the ECHAM 4.5 atmospheric general circulation model, J. Clim., 20(10), 2037–2050, doi:10.1175/JCLI4122.1, 2007.

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P. and van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sens. Environ., 123, 280–297, doi:10.1016/j.rse.2012.03.014, 2012.

Loew, A., Stacke, T., Dorigo, W., De Jeu, R. and Hagemann, S.: Potential and limitations of multidecadal satellite soil moisture observations for selected climate model evaluation studies, Hydrol. Earth Syst. Sci., 17(9), 3523–3542, doi:10.5194/hess-17-3523-2013, 2013.

Lorenz, R., Jaeger, E. B. and Seneviratne, S. I.: Persistence of heat waves and its link to soil moisture

memory, Geophys. Res. Lett., 37(9), 1–5, doi:10.1029/2010GL042764, 2010.

Majozi, N. P., Mannaerts, C. M., Ramoelo, A., Mathieu, R., Nickless, A. and Verhoef, W.: Analysing surface energy balance closure and partitioning over a semi-arid savanna FLUXNET site in Skukuza, Kruger National Park, South Africa, Hydrol. Earth Syst. Sci., 21(7), 3401–3415, doi:10.5194/hess-21-3401-2017, 2017.

Martens, B., Miralles, D. G., Lievens, H., Van Der Schalie, R., De Jeu, R. A. M. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., Verhoest, N. E. C. C., Férnandez-Prieto, D., Beck, H. E., Dorigo, W. A. and Verhoest, N. E. C. C.: GLEAM v3: Satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev. Discuss., 10(5), 1903–1925, doi:10.5194/gmd-10-1903-2017, 2017.

McNally, A., Shukla, S., Arsenault, K. R., Wang, S., Peters-Lidard, C. D. and Verdin, J. P.: Evaluating ESA CCI soil moisture in East Africa, Int. J. Appl. Earth Obs. Geoinf., 48, 96–109, doi:10.1016/j.jag.2016.01.001, 2016.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A. and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci., 15(2), 453–469, doi:10.5194/hess-15-453-2011, 2011.

Muthige, M. S., Malherbe, J., Englebrecht, F. A., Grab, S., Beraki, A., Maisha, T. R. and Van der Merwe, J.: Projected changes in tropical cyclones over the South West Indian Ocean under different extents of global warming, Environ. Res. Lett., 13(6), 065019, doi:10.1088/1748-9326/aabc60, 2018.

Palmer, A. R., Weideman, C., Finca, A., Everson, C. S., Hanan, N. and Ellery, W.: Modelling annual evapotranspiration in a semi-arid, African savanna: functional convergence theory, MODIS LAI and the Penman–Monteith equation, African J. Range Forage Sci., 32(1), 33–39, doi:10.2989/10220119.2014.931305, 2015.

Pinheiro, A. C. and Tucker, C. J.: Assessing the relationship between surface temperature and soil moisture in southern Africa, Remote Sens. Hydrol., 2000(267), 296–301, 2001.

Raj Koirala, S. and W. Gentry, R.: SWAT and wavelet analysis for understanding the climate change impact on hydrologic response, Open J. Mod. Hydrol., 02(02), 41–48, doi:10.4236/ojmh.2012.22006, 2012.

Ramoelo, A., Majozi, N., Mathieu, R., Jovanovic, N., Nickless, A. and Dzikiti, S.: Validation of global evapotranspiration product (MOD16) using flux tower data in the African savanna, South Africa, Remote Sens., 6(8), 7406–7423, doi:10.3390/rs6087406, 2014.

Rosch, A. and Schmidbauer, H.: WaveletComp : A guided tour through the R-package, , 1–38 [online] Available from: http://www.hs-stat.com/projects/WaveletComp/WaveletComp_guided_tour.pdf, 2018.

Scholes, R. J., Gureja, N., Giannecchinni, M., Dovie, D., Wilson, B., Davidson, N., Piggott, K., McLoughlin, C., Van der Velde, K., Freeman, A., Bradley, S., Smart, R. and Ndala, S.: The environment and vegetation of the flux measurement site near Skukuza, Kruger National Park, Koedoe, 44(1), 73–84, doi:10.4102/koedoe.v44i1.187, 2001.

Seneviratne, S. I., Koster, R. D., Guo, Z., Dirmeyer, P. A., Kowalczyk, E., Lawrence, D., Liu, P., Mocko, D., Lu, C.-H., Oleson, K. W. and Verseghy, D.: Soil moisture memory in AGCM simulations: Analysis of global land–atmosphere coupling experiment (GLACE) data, J. Hydrometeorol., 7(5), 1090–1112, doi:10.1175/JHM533.1, 2006.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B. and Teuling, A. J.: Investigating soil moisture-climate interactions in a changing climate: A review, Earth-

819 Science Rev., 99(3–4), 125–161, doi:10.1016/j.earscirev.2010.02.004, 2010.

Sheffield, J. and Wood, E. F.: Global trends and variability in soil moisture and drought characteristics, 1950-2000, from observation-driven simulations of the terrestrial hydrologic cycle, J.
822 Clim., 21(3), 432–458, doi:10.1175/2007JCLI1822.1, 2008.

Shongwe, M. E., van Oldenborgh, G. J., van den Hurk, B. J. J. M., de Boer, B., Coelho, C. A. S. and van Aalst, M. K.: Projected changes in mean and extreme precipitation in Africa under global
825 warming. Part I: Southern Africa, J. Clim., 22(13), 3819–3837, doi:10.1175/2009JCLI2317.1, 2009.

Shongwe, M. E., Lennard, C., Liebmann, B., Kalognomou, E.-A., Ntsangwane, L. and Pinto, I.: An evaluation of CORDEX regional climate models in simulating precipitation over Southern Africa,
828 Atmos. Sci. Lett., 16(3), 199–207, doi:10.1002/asl2.538, 2015.

Sinclair, S. and Pegram, G. G. S.: A comparison of ASCAT and modeled soil moisture over South Africa, using TOPKAPI in land surface mode, Hydrol. Earth Syst. Sci. Discuss., 6, 7439–7482, 2010.

831 Stevens, N., Bond, W., Hoffman, T. and Midgley, G.: Change is in the air : Ecological trends and their drivers in South Africa, South African Environmental Observation Network. [online] Available from: http://www.saeon.ac.za/Change is in the air_WEB VERSION.pdf, 2015.

834 Torrence, C. and Compo, G. P.: A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc., 79(1), 61–78, doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2, 1998.

Veleda, D., Montagne, R. and Araujo, M.: Cross-wavelet bias corrected by normalizing scales, J.
837 Atmos. Ocean. Technol., 29(9), 1401–1408, doi:10.1175/JTECH-D-11-00140.1, 2012.

Wang, S., Mo, X., Liu, S., Lin, Z. and Hu, S.: Validation and trend analysis of ECV soil moisture data on cropland in North China Plain during 1981–2010, Int. J. Appl. Earth Obs. Geoinf., 48, 110–121,
840 doi:10.1016/j.jag.2015.10.010, 2016.

Wang, T. and Franz, T. E.: Evaluating climate and soil effects on regional soilmoisture spatial variability using EOFs, Water Resour. Res., (1), 5375–5377, doi:10.1002/2013WR014979.Reply,
843 2017.

Wang, Y. P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M. R., Pak, B., Van Gorsel, E. and Luhar, A.: Diagnosing errors in a land surface model (CABLE) in the time and frequency
846 domains, J. Geophys. Res. Biogeosciences, 116(1), 1–18, doi:10.1029/2010JG001385, 2011.

Whitley, R., Beringer, J., Hutley, L. B., Abramowitz, G., De Kauwe, M. G., Duursma, R., Evans, B., Haverd, V., Li, L., Ryu, Y., Smith, B., Wang, Y. P., Williams, M. and Yu, Q.: A model inter-
849 comparison study to examine limiting factors in modelling Australian tropical savannas, Biogeosciences, 13(11), 3245–3265, doi:10.5194/bg-13-3245-2016, 2016.

Xia, Y., Ek, M. B., Wu, Y., Ford, T. and Quiring, S. M.: Comparison of NLDAS-2 simulated and
852 NASMD observed daily soil moisture. Part I: comparison and analysis, J. Hydrometeorol., 16(5), 1962–1980, doi:10.1175/JHM-D-14-0096.1, 2015.

Yuan, S. and Quiring, S. M.: Evaluation of soil moisture in CMIP5 simulations over the contiguous
855 United States using in situ and satellite observations, Hydrol. Earth Syst. Sci., 21(4), 2203–2218, doi:10.5194/hess-21-2203-2017, 2017.

Zeng, J., Li, Z., Chen, Q., Bi, H., Qiu, J. and Zou, P.: Evaluation of remotely sensed and reanalysis
858 soil moisture products over the Tibetan Plateau using in-situ observations, Remote Sens. Environ., 163, 91–110, doi:10.1016/j.rse.2015.03.008, 2015.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

5 Appendix A

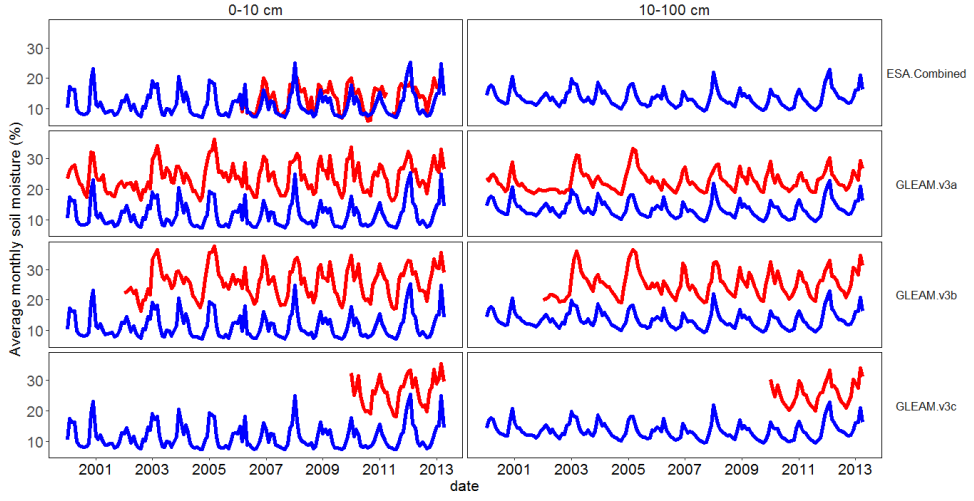


Figure A1. Comparison of monthly modelled and satellite products (red line) with CCAM-CABLE (blue line) surface (0-10 cm), and root zone (10-100 cm) soil moisture at the Skukuza (2001–2014) site.

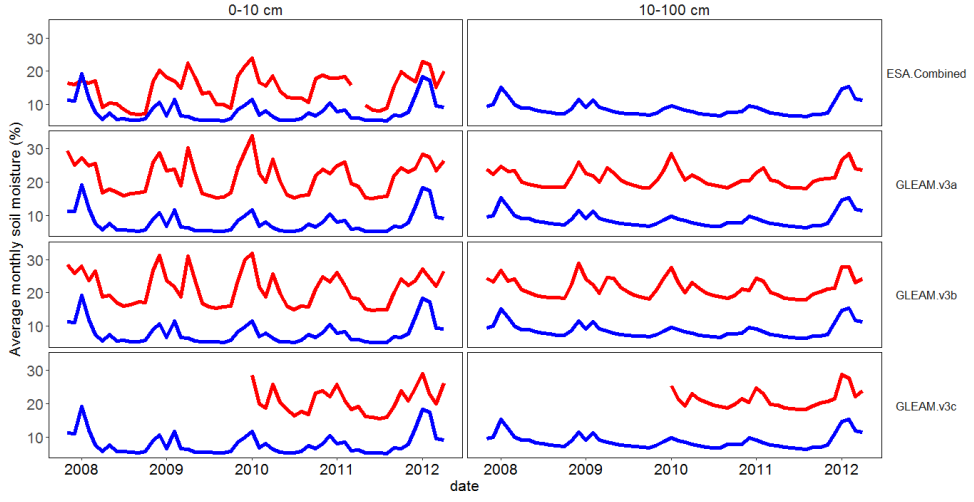
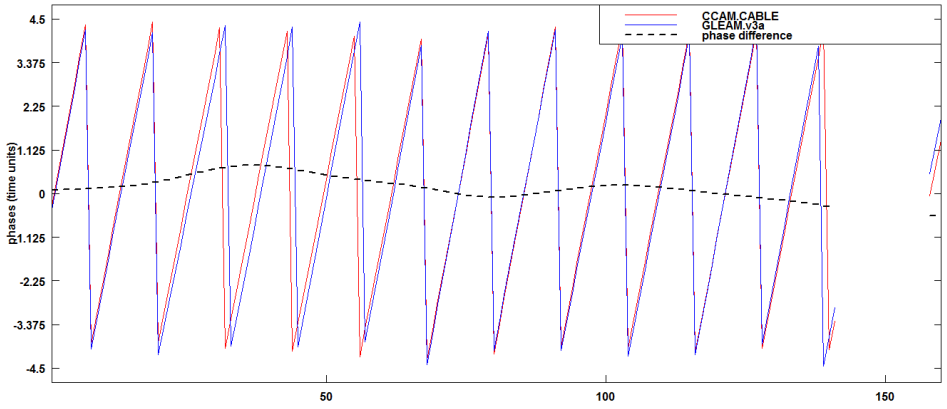


Figure A2. Comparison of monthly modelled and satellite products (red line) with CCAM-CABLE (blue line) surface (0-10 cm), and root zone (10-100 cm) soil moisture at the Malopeni (2008–2013) site.

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-546
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 2 November 2018

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

6 Appendix B



870 **Figure B1.** Phase difference between surface soil moisture simulated using CCAM-CABLE, and GLEAM v3a at Skukuza between 2001, and 2014 at period 12 at the surface.