Local-scale secondary water inputs modulate seasonal vegetation cover decay rate across Africa

Çağlar Küçük [1,2], Sujan Koirala [1], Nuno Carvalhais [1,3], Diego G. Miralles [2], Markus Reichstein [1], Martin Jung [1]

[1]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, Germany
[2]Hydro-Climate Extremes Lab (H-CEL), Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium
[3]CENSE, Departamento de Ciências e Engenharia do Ambiente, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Caparica, Portugal

Key Points:

- We quantify effects of secondary water inputs on seasonal vegetation cover decay rate on water-limited parts of Africa via machine learning
- Shallow groundwater, topography, and soil properties support vegetation activity over large domains by enhancing surface soil moisture
- 1/3 of seasonal vegetation cover decay rate over the study domain is attributed to secondary water inputs modulated by land properties

Corresponding author: Çağlar Küçük, `ckucuk@bgc-jena.mpg.de`

Abstract

Next to precipitation, secondary water sources emerging from shallow groundwater and lateral redistribution of soil moisture, together with soil properties modulating their accessibility are highly important in water-limited ecosystems. However, effects of these land-associated secondary inputs are not well known over large domains given the mismatch of spatial scales of processes. Here, we quantify the role of land properties on the spatial variations of seasonal decay rate of vegetation cover over water-limited regions of Africa, using machine learning. Over the study domain, 17 % of these variations are directly attributed to land properties, and 16 % are attributed to interaction effects of land properties with climate and vegetation. Locally, total land attributed variations account for more than 60 % in hotspots with different land properties like shallow groundwater, complex topography, and favourable soil properties. Our findings lend empirical evidence for the importance of local-scale secondary water inputs over large domains.

Plain Language Summary

The water needed for vegetation over land is primarily provided by atmosphere as precipitation. However, secondary water inputs enabled by the presence of shallow groundwater or lateral convergence can support the vegetation significantly, especially in dry regions. These secondary inputs, and the soil properties modulating their accessibility to vegetation vary dramatically at local scales. To date, extent of these secondary effects is not well understood over large domains. Here, we quantified the effects of secondary water inputs to the seasonal decay rate of vegetation over water-limited regions of Africa. Using machine learning, we modelled the seasonal decay rate of vegetation as a function of climate, land, and vegetation properties. Over the study domain, we found that secondary water inputs account for 1/3 of the variations in the seasonal decay rate of vegetation. Half of that relates to direct effects of land properties on vegetation, while the other relates to interactions of land with climate and vegetation. Moreover, in local hotspots, secondary inputs control up to 60 % of vegetation cover decay rate. There, shallow groundwater, topography, and soil properties support vegetation against water limitation. Our results indicate importance of representing these local-scale processes accurately to realistically portray large-scale dryland dynamics.

1 Introduction

Drylands cover more than 40 % of land surface globally (D'Odorico et al., 2019). They have a strong impact on the global carbon cycle (Lal, 2019), despite their vulnerability against interannual climatic variations (Brandt et al., 2018). Furthermore, more than 1/3 of the World's population is settled on drylands (Reynolds et al., 2007), 90 % of which on developing countries that strongly rely on ecosystem services (Maestre et al., 2012). Despite their importance, drylands are still not well understood (Maestre et al., 2021). This is particularly the case in Africa, where drylands cover 75 % of the surface and remain severely under-studied (Maestre et al., 2012; Adole et al., 2016; Prăvălie, 2016). Overall, it is crucial to investigate ecosystem dynamics in drylands in order to have a more comprehensive vision of African ecology and biogeography.

Apart from precipitation as the primary component of the terrestrial water cycle, secondary water resources like groundwater (Fan, 2015; Maxwell & Condon, 2016), capillary rise (Koirala et al., 2019), and lateral flow at hillslope scales (Fan et al., 2019), are essential components of the water cycle. Since the functioning of dryland ecosystems is controlled by water availability (Rodriguez-Iturbe & Porporato, 2009), the importance of the secondary water resources can be large. However, land surface models still need to better representation at high spatial scales to be able to capture these secondary, non-trivial, components of the water cycle (Van Dijk et al., 2018; Mu et al., 2021). The stochastic nature of soil moisture, together with an array of factors and processes affecting it, makes modelling soil moisture and capturing its spatial variations in drylands particularly challenging (Rodriguez-Iturbe et al., 2021). This challenge propagates to the representation of land surface heterogeneity and hydrological processes affected in the Earth System Models (M. P. Clark et al., 2015; Fisher & Koven, 2020; Blyth et al., 2021).

Thanks to recent advancements in remote sensing, high-resolution Earth observation products provide nowadays unprecedented opportunities to improve our understanding in Earth system science. While polar orbiting satellites can be used to monitor the land surface at sub-metre resolutions (e.g., Brandt et al., 2020), geostationary satellites provide insights at high temporal resolution (e.g., Khan et al., 2021; Hashimoto et al., 2021). In recent years, data-driven methods using Machine Learning (ML) and deep learning – which are very powerful in resolving complex interactions in large Earth observation datasets – have been widely used. However, the interpretability of these models is critical and still poses challenges (Rudin et al., 2021), despite the stunning pace of developments in interpretable ML (Molnar, 2019).

In this study, we quantify the effect of land properties (that modulate secondary water resources) on the seasonal decay in vegetation cover (λ) in Africa, estimated using geostationary satellite retrievals (Küçük et al., 2020). Based on an asymptotic exponential decay function, λ quantifies the seasonal decay rate of vegetation cover and creates the possibility to analyse decay dynamics across large domains covering different climate and vegetation types at ca. 5 km spatial resolution. In addition to the strong covariation with climate at large scales, λ also has consistent anisotropic structures at local scales. Initial analysis in Küçük et al. (2020) showed that λ reflects ecosystem scale water use strategies against seasonal water limitation, which may be primarily driven by climate over large scales but also affected by secondary land effects modulating water limitation locally. In order to quantify these secondary effects, we model λ using climate, vegetation, and land properties from an array of products with ML. The main hypothesis of the study is that the spatial variations of λ are primarily driven by water limitation over large parts of Africa, thus any secondary water input supports the ecosystem against water limitation, and leads to a decreased rate of seasonal decay (e.g., the shallower the groundwater, the slower the vegetation decay). We constrain the ML model to follow this hypothesis, regarding secondary water resources, and analyse the model structure to understand the underlying factors. Finally, we quantify land attributed spatial variations of λ and show the sensitivity of the driving factors to climatological aridity.

2 Data and Methods

We used land, climate and vegetation properties over the study domain to model spatial variations of λ that are shown in Table 1 (see Supporting Information for the details of estimations and pre-processing). Regarding the land properties, we used predictors covering (i) groundwater as a secondary water resource, (ii) topographic complexity as a land property that modulates the amount of plant available soil moisture by lateral redistribution of moisture, and (iii) soil hydraulic properties, as the fundamental module in defining the accessibility of soil moisture by plants. In order to incorporate climatological aridity into the model, we used precipitation, temperature and shortwave radiation data across annual and seasonal time scales. Last set of predictors cover vegetation properties shown in Table 1.

After preparing the data to use in modelling, we filtered the study domain for a maximum precipitation value of 1500 mm/year to limit confounding factors affecting λ other than water limitation. Moreover, we excluded λ values with low confidence by filtering for relative standard error less than 1 and at least 3 convergences during the estimation (see Küçük et al., 2020, for product details). Overall, around 730000 grid cells

Table 1: Summary of the dataset used in the study.

Variable	***Data Source***	***Spat. Res.***
Seasonal decay rate of vegetation cover (λ)	Küçük et al. (2020)	5 km
Plant Available Water[1] (PAW)		
Soil hydraulic conductivity at Field Capacity[1] (k_{FC})	Estimated	250 m
Max potential upwards capillary flux[1,2] (I_{cap})		
Water Table Depth (WTD)	Fan et al. (2013)	1 km
Height Above Nearest Drainage (HAND)	Yamazaki et al. (2019)	90 m
Wetlands	Tootchi et al. (2019)	500 m
Topographic Wetness Index (TWI)		
Vectoral Ruggedness Measure (VRM)	Amatulli et al. (2020)	250 m
Magnitude and scale of 3D roughness		
Precipitation[3]	Fick and Hijmans (2017)	
Temperature[3]		5 km
Radiation[3]	Abatzoglou et al. (2018)	
Canopy height	Simard et al. (2011)	1 km
Tree & non-tree vegetation cover	Dimiceli et al. (2015)	
Burned area	Giglio et al. (2015)	250 m
Plant Functional Type	Friedl and Sulla-Menashe (2019)	

[1] Estimated using Hengl et al. (2017), based on Saxton and Rawls (2006)
[2] Based on Richards (1931)
[3] Annual and seasonal scales

with ca. 5 km spatial resolution were kept in the study domain. Map of the target variable after filtration is available in Fig. S1.

We used XGBoost (Chen & Guestrin, 2016), a recent implementation of gradient boosted regression trees, to model spatial variations of λ with land, climate and vegetation properties. Gradient boosting is a ML method that uses an ensemble of tree-based models generated by subsets of the training data. Tree based regression is a powerful method with high flexibility, designed to minimise output error with a strong gradient search without considering the underlying processes between predictors and target. In order to avoid unlikely attributions to predictors about variation of λ, and ensure the model to consistently reflect the hypothesis between λ and water availability, we constrained the model to have monotonic relationship between λ and land parameters with the principle that any land parameter promoting surface soil moisture via secondary water inputs should correlate positively with λ. In other words, we constrained the model to have positive monotonicity between λ and land parameters, i.e., the larger plant available water the slower vegetation decay, except with WTD and HAND where negative constraints were set, i.e., the deeper the groundwater the weaker its support to surface soil moisture. After setting the constraints, we used 10 % of the grid cells which are randomly selected to build the model and used rest of the grid cells for validation.

Although tree based models are relatively easy to interpret, it is not trivial to estimate importance of predictors of a multi-dimensional and nonlinear ML model in an unbiased way. Lundberg and Lee (2017) suggested using SHapley Additive exPlanation (SHAP) values to address the problem, which is rooted from cooperative game theory (Shapley, 1953) and treats each predictor as a player of a game. Being an additive explanation method, summation of SHAP values of all predictors for an instance, a grid cell in this study, is equal to the deviation of the predicted value of that instance from the mean value of the predictions. Moreover, it is possible to partition the SHAP values for direct and interaction effects. In other words, for a simple modelling scenario of $y_{obs} \approx y_m = f(x_1, x_2)$ where y_{obs} and y_m are the observed and modelled target variable, and x_1 and x_2 are the predictors, $y_m = \overline{y_m} + \phi_{x1-x1} + \phi_{x2-x2} + \phi_{x1-x2}$ where $\overline{y_m}$ is mean of y_m, ϕ_{x1-x1} and ϕ_{x1-x2} are the SHAP values attributed to predictor x_1 alone and to the interaction effects between the two predictors. Lundberg et al. (2020) suggested exploiting model structures of tree based models to approximate SHAP values to avoid computational complexity on large datasets. In order to limit methodological problems related to feature interdependence (see Sec. 3.4) and ease interpretability, we grouped SHAP values of the predictors as land, climate and vegetation properties, to explain the model output as:

$$\lambda \approx \lambda_m = \overline{\lambda_m} + \phi_{land-direct} + \phi_{land-clim} + \phi_{land-veg} + \phi_{clim-direct} + \phi_{clim-veg} + \phi_{veg-direct} \tag{1}$$

Afterwards, in order to quantify the importance of land parameters, we normalised the ϕ values of different sets of features after taking absolute values such as:

$$\Phi_{land-total} = \frac{|\phi_{land-direct}| + |\phi_{land-clim}| + |\phi_{land-veg}|}{|\phi_{land-direct}| + |\phi_{land-clim}| + |\phi_{land-veg}| + |\phi_{clim-direct}| + |\phi_{clim-veg}| + |\phi_{veg-direct}|} \tag{2}$$

Finally, we analysed the sensitivity of $\Phi_{land-total}$ to changes in WTD, topographic complexity, maximum potential capillary flux, and annual precipitation.

3 Results and Discussion

3.1 Model output for λ

The ML model (λ_m, shown in Fig. 1a) captured the continental gradient as well as local variations of λ with 55 % Nash–Sutcliffe modelling efficiency (Nash & Sutcliffe, 1970). However, residuals of the model shows anisotropic structures at local scales (Fig. 1b). This suggests that the model did not capture all the local scale variations, presumably due to incomplete and non-perfect predictors used in the model. After building the model, we analysed λ_m and attributed its spatial variations to predictors by considering the model structure via SHAP values.

3.2 Importance of land on seasonal decay rate of vegetation cover

Spatial variation of normalised importance of land on λ ($\Phi_{land-total}$, see Eq. 2) is mapped in Fig. 2 together with six zoomed insets and histogram of the values where the mean value over the domain is shown with a dashed line. Over the study domain, 33 % of the variations of λ is attributed to land effects, 17 % of which is direct effects while 16 % is the interaction effects with climate and vegetation. Moreover, we found mesoscale hotspots where this attribution affects more than 60 % of the spatial variation of λ (Fig. 2). Complex but structured distribution of these local-scale hotspots show not only the importance of secondary water resources but also the difficulty to generalise their effects over large domains.

At local scales, regions with shallow groundwater are within these hotspots such as Box-B showing the South of Lake Chad, between the Logone and Chari Rivers and the Sudd Swamp – Fig. 2 (see Fan et al., 2013, for water table depth estimates), which agrees with the literature on the importance of groundwater (Koirala et al., 2017; Roebroek et al., 2020). Additionally, we found strong land effects over the Ethiopian Highlands (Box-E) as well as the Manica Highlands (Box-F) to a lesser extent (see V. Clark et al., 2017, for further information about the Manica Highlands). This is consistent with the literature regarding topographical complexity as an important factor modulating water limitation at hillslope scales by enhancing soil moisture at valleys and riparian zones via lateral convergence of soil moisture (Fan et al., 2019).

Spatial patterns in Fig. 2 bear strong agreement with the secondary evaporation patterns that include permanent or ephemeral waterbodies, groundwater uptake, soil evaporation, and irrigation (Van Dijk et al., 2018). By assimilating remote sensing data with a process-based, hydrological model, Van Dijk et al. (2018) showed secondary water inputs affect plant transpiration globally. The agreement among the findings of our observation-based, ML-leveraged study on the importance of secondary water inputs with a process-based data assimilation study sheds light to the direction of future studies.

In order to understand the driving factors of the normalised importance of land parameters as direct land effects ($\Phi_{land-direct}$) and interaction effects with climate ($\Phi_{land-clim}$) and vegetation ($\Phi_{land-veg}$), we analysed their covariation with topographic complexity, groundwater, and capillary rise. In general, direct land effect ($\Phi_{land-direct}$) is the largest component of normalised importance of land, followed by land and climate interaction effects ($\Phi_{land-clim}$), and finally interaction effects between land and vegetation ($\Phi_{land-veg}$).

Over the entire study domain, we found a robust positive correlation between VRM, a metric summarising topographic complexity, and $\Phi_{land-total}$, largely driven by direct land effects (Fig. 3a) which confirms the previously reported studies at basin scales on positive effects of concentrated soil moisture at hillslope scales due to lateral convergence at the much larger study domain of this study (Hoylman et al., 2018; Tai et al., 2020). Half of the spatial variations of λ is attributed to land in the regions with VRM values greater than 0.85 %. Lower values of $\Phi_{land-total}$ at smaller VRM values suggest other processes become more dominant as the effects of topography reduces.

Secondly, we looked at the same covariation with WTD to relate variations of λ to groundwater (Fig. 3b). Groundwater is an important moisture source for vegetation in water-limited systems and this effect is amplified as it becomes available in shallow depths (Barbeta & Peñuelas, 2017). We observed this effect on the normalised land importance ($\Phi_{land-total}$) with changing WTD where almost half of variations in λ is attributed to land in regions with WTD<1 meter (m). This effect is gradually reduced with deeper groundwater levels up to 16 m. This relation, however, does not hold at WTD levels deeper than 16 m, presumably due to the disconnection between surface and groundwater where other factors become more prominent.

Finally, we observed a similar covariation with the largest gradient with the maximum potential capillary rise (I_{cap}) and $\Phi_{land-total}$, where variations of λ is attributed to land parameters are larger with greater potential of capillary supply (Fig. 3c). Overall, more than half of the variations in λ are attributed to land in regions with $I_{cap} >$ 1 mm/day, due to the physical properties of soil texture. This fits well with the previous studies that soil texture is a key variable mediating the interactions between climate, soil, and vegetation (Fernandez-Illescas et al., 2001).

3.3 Effects of aridity to the importance of land parameters

In order to understand the effects of mean annual precipitation, as a simple proxy for climatological aridity, to the importance of land parameters on λ, we analysed the changes on the covariation between normalised importance of land parameters and VRM, WTD, and I_{cap} over a precipitation gradient of 0 to 1500 mm/year.

Sensitivity of the covariation between $\Phi_{land-total}$ and VRM to precipitation suggests that topographic complexity affects λ the most in semi-arid regions (Fig. 4a). Lower $\Phi_{land-total}$ values at higher precipitation values agree with the main hypothesis of the study that λ is derived by water limitation. Moreover, lower $\Phi_{land-total}$ values with very low precipitation values is likely due to the fact that most of the water input is returned to atmosphere locally by soil evaporation under hyper arid conditions (Newman et al., 2006), which reduces the importance of lateral convergence of soil moisture.

Secondly, we analysed sensitivity of the interaction between WTD and land attributed variations of λ to precipitation (Fig. 4b). We found the largest attribution to land parameters in regions with WTD $<$ 1 m, with no clear sensitivity to the precipitation gradient of 0 - 1500 mm/year, suggesting strong effect of groundwater when easily accessible. Except the extreme values of WTD where groundwater is directly available at land surface or disconnected from it, i.e., WTD < 1 m or WTD > 16 m, $\Phi_{land-total}$ values consistently decrease with decreasing climatological aridity. This trend becomes stronger with deeper groundwater levels at larger precipitation values due to lower importance of groundwater with weaker water limitation. These findings agree with previous studies as groundwater subsidises root zone soil moisture and effects of it become more important with stronger aridity (Brooks et al., 2015).

Finally, we analysed the effects of precipitation on the covariation between I_{cap} and importance of land on variations of the λ. We found not only the strongest but also the most consistent gradient between $\Phi_{land-total}$ and precipitation against I_{cap} (Fig. 4c), where the largest land attributed variation of λ occurs in regions with strong climatological aridity and the largest potential of capillary rise. Although land effects become weaker with decreasing climatological aridity, they show the smallest sensitivity against precipitation, showing the importance of capillary rise against water limitation.

3.4 Robustness and limitations

Our machine learning based quantification and analysis of secondary moisture effects on the seasonal vegetation decay over Africa is associated with uncertainties of underlying assumptions and methods. The most fundamental assumption is that the vegetation decay rate (λ) is strongly influenced by plant available moisture. Many studies have found and confirmed that most of African ecosystems are water-limited even though relationships can be complex and diverse (see Küçük et al., 2020, and references therein). We confined the study domain to retain primarily water-limited systems by excluding the wetter tropical regions (see Sec. 2). Key findings of our study, the importance of secondary moisture sources in general, and their decreasing importance with climatological humidity, are consistent with the assumption of dealing with water-limited ecosystems.

The main methodological uncertainties are related to a) the quality and performance of the underlying trained machine learning model, and b) to the correct attribution of modelled lambda variations to land properties. Our machine learning model explained only 55 % of lambda variations based on 10 % of randomly selected pixels for training to avoid overfitting due to spatial auto-correlation (Roberts et al., 2017). This suggests that we are lacking important predictors and/or issues in the quality of data products used as predictors. The model residuals (Fig. 1b) show relatively little large scale patterns but a rather fine grained structure. Thus, we likely underestimate lambda variations due to landscape-scale factors which suggest that our attribution to land properties maybe conservative and even more important in reality. The imperfect representation of surface and subsurface factors governing secondary moisture sources in the predictor set is likely also constrained by the spatial resolution of 3-5 km where likely important sub-grid variations of factors and responses in lambda cannot be resolved adequately.

While we used Shapley values as state-of-the-art technique for machine learning based attribution to predictors we need to acknowledge that machine learning methods exploit statistical associations without any guarantee of unravelling causal relationships. In our experimental design we aimed at enhanced interpretability of the results by constraining the predictor set to interpretable factors related to our hypothesis, and by constraining the monotonicity of land predictors to lambda according to prior knowledge. These monotonic constraints prescribe only the sign of the response while the shape remains flexible which acts as a causal regularisation in the model training process. However, we cannot claim that our trained machine-learning model is entirely based on causal relationships overall. Some confidence in the qualitative findings of the study originate from the fact that the importance of land properties varies systematically with topographic complexity, water table depth, and maximum capillary rise according to theory and expectations from previous studies (Fig. 3). Please note that this result is not trivial and not enforced by the monotonic predictor constraints since we estimated land importance as mean absolute deviations (Eq. 2).

A key uncertainty of estimating variable importance in machine learning is due to covariations of predictors, including SHAP values (Kumar et al., 2020). We aimed at minimising this issue by analysing the importance of predictor groups, rather than individual predictors based on consistent aggregation of Shapley values (Eq. 2). Therefore, covariation of predictors e.g. within the land group cause no issues and biases of estimated importances. While most co-variation among predictors is within their group, there remains covariation among groups that can potentially lead to some confounding effects.

Given the limitations outlined above, our data-driven findings present hypothesis on the large scale importance of secondary moisture effects on seasonal vegetation decay over Africa. Given that the patterns we found are consistent with theory and literature along with a likely underestimation of the effect of secondary moisture sources due to limited information in the predictors we believe that scrutinising our empirical findings will be critical for improving our understanding of dryland ecohydrology across spatial scales.

4 Conclusions

In this study, we analysed the effects of local scale water resources on seasonal water limitation by analysing the model output of the seasonal vegetation decay rate (λ) of Fractional Vegetation Cover (FVC) over Africa at 5 km spatial resolution. The model output revealed that at local scales, more than 60 % of the variation of λ in space is attributed to land properties in hotspots where land strongly modulates water limitation with different processes, e.g., shallow groundwater or complex topography. Over the study domain, 17 % of variation of λ in space is directly attributed to land while 16 % is attributed to interactions of climate and vegetation properties with land. Moreover, sensitivity of land effects of λ increases with stronger aridity, where contributions of secondary water resources become relatively stronger in water cycle. We found that maximum potential capillary rise of groundwater (I_{cap}) positively correlates with land attributed variations of λ ($\Phi_{land-total}$). 33 % of spatial variations of λ is directly attributed to land effects in regions with $I_{cap} > 1.2$ mm/day. Moreover, this effect becomes larger with stronger aridity. Similarly, land attributed variations of λ correlate negatively with deeper WTD as long as groundwater is connected with surface (WTD < 16 m). Effects of WTD on $\Phi_{land-total}$ reduces with larger annual precipitation values, except shallow groundwater levels (WTD < 1 m). Finally, we found positive correlation between topographic complexity and land attributed variations of λ over the study domain with the largest $\Phi_{land-total}$ in semi-arid regions with complex topography, which shows the importance of lateral moisture convergence due to topography in semi-arid regions. Our findings show the importance of local scale processes affecting water availability in drylands not only at local but also continental to global scales, and the need of bridging processes across spatial scales in ecohydrological studies over large domains.

5 Data Availability Statement

Raster files of raw SHAP values of direct and interaction effects of land, climate, and vegetation and the normalised importance of land effects are available as netCDF format in `https://doi.org/10.6084/m9.figshare.16780405.v1`.

Acknowledgments

Çağlar Küçük acknowledges funding from the International Max Planck Research School for Global Biogeochemical Cycles. Additionally, Çağlar Küçük would like to express his gratitude to the colleagues at the Max Planck Institute for Biogeochemistry helped establishing the scientific base of this data-driven study on ecohydrology.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015. *Scientific Data*, *5*, 1–12. doi: 10.1038/sdata.2017.191

Adole, T., Dash, J., & Atkinson, P. M. (2016). A systematic review of vegetation phenology in Africa. *Ecological Informatics*, *34*, 117–128. doi: 10.1016/j.ecoinf.2016.05.004

Amatulli, G., McInerney, D., Sethi, T., Strobl, P., & Domisch, S. (2020). Geomorpho90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers. *Scientific Data*, *7*(1), 1–18. doi: 10.1038/s41597-020-0479-6

Barbeta, A., & Peñuelas, J. (2017). Relative contribution of groundwater to plant transpiration estimated with stable isotopes. *Scientific Reports*, *7*(1), 1–10. doi: 10.1038/s41598-017-09643-x

Blyth, E. M., Arora, V. K., Clark, D. B., Dadson, S. J., De Kauwe, M. G., Lawrence, D. M., … Yuan, H. (2021). Advances in land surface modelling. *Current Climate Change Reports*, *7*(2), 45–71. doi: 10.1007/s40641-021-00171-5

Brandt, M., Tucker, C. J., Kariryaa, A., Rasmussen, K., Abel, C., Small, J., … Fensholt, R. (2020). An unexpectedly large count of trees in the West African Sahara and Sahel. *Nature*, *587*(7832), 78–82. doi: 10.1038/s41586-020-2824-5

Brandt, M., Wigneron, J. P., Chave, J., Tagesson, T., Penuelas, J., Ciais, P., … Fensholt, R. (2018). Satellite passive microwaves reveal recent climate-induced carbon losses in African drylands. *Nature Ecology and Evolution*, *2*(5), 827–835. doi: 10.1038/s41559-018-0530-6

Brooks, P. D., Chorover, J., Fan, Y., Godsey, S. E., Maxwell, Reed M., McNamara, J. P., & Tague, C. (2015). Hydrological partitioning in the critical zone: Recent advances and opportunities for developing transferable understanding of water cycle dynamics. *Water Resources Research*, *51*, 6973–6987. doi: 10.1002/2015WR017039

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. doi: 10.1145/2939672.2939785

Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., … Zeng, X. (2015). Improving the representation of hydrologic processes in Earth System Models. *Water Resources Research*, *51*(8), 5929–5956. doi: 10.1002/2015WR017096

Clark, V., Timberlake, J., Hyde, M., Mapaura, A., Palgrave, M. C., Wursten, B., … Barker, N. (2017). A first comprehensive account of floristic diversity and endemism on the Nyanga Massif, Manica Highlands (Zimbabwe–Mozambique). *Kirkia*, *19*(1), 1–53. Retrieved from `https://www.jstor.org/stable/48595380`

Cronberg, G., Gieske, A., Martins, E., Prince Nengu, J., & Stenström, I.-M. (1995). Hydrobiological studies of the Okavango Delta and Kwando/Linyati/Chobe River, Botswana I surface water quality analysis. *Botswana Notes and Records*, *27*. Retrieved from `http://www.jstor.org/stable/40980045`

Dimiceli, C., Carroll, M., Sohlberg, R., Kim, D. H., Kelly, M., & Townshend, J. R. G. (2015). *MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V006.* NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MOD44B.006

D'Odorico, P., Porporato, A., & Runyan, C. (2019). Ecohydrology of arid and semiarid ecosystems: An introduction. In P. D'Odorico, A. Porporato, & C. Wilkinson Runyan (Eds.), *Dryland ecohydrology* (pp. 1–27). Springer Inter-

national Publishing. doi: 10.1007/978-3-030-23269-6_1
Fan, Y. (2015). Groundwater in the Earth's critical zones: Relevance to large-scale patterns and processes. *Water Resources Research*, 3052–3069. doi: 10.1002/2015WR017037
Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., ... Yamazaki, D. (2019). Hillslope hydrology in global change research and earth system modeling. *Water Resources Research*, 1737–1772. doi: 10.1029/2018WR023903
Fan, Y., Li, H., & Miguez-Macho, G. (2013). Global patterns of groundwater table depth. *Science*, *339*(6122), 940–943. doi: 10.1126/science.1229881
Fernandez-Illescas, C. P., Porporato, A., Laio, F., & Rodriguez-Iturbe, I. (2001). The ecohydrological role of soil texture in a water-limited ecosystem. *Water Resources Research*, *37*(12), 2863–2872. doi: 10.1029/2000WR000121
Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. doi: 10.1002/joc.5086
Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of Land Surface Models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*. doi: 10.1029/2018ms001453
Friedl, M., & Sulla-Menashe, D. (2019). *MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006.* NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MCD12Q1.006
Giglio, L., Justice, C., Boschetti, L., & Roy, D. (2015). *MCD64A1 MODIS/Terra+Aqua Burned Area Monthly L3 Global 500m SIN Grid V006.* NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MCD64A1.006
Hashimoto, H., Wang, W., Takenaka, H., Higuchi, A., Dungan, J. L., Li, S., ... Nemani, R. R. (2021). New generation geostationary satellite observations support seasonality in greenness of the Amazon evergreen forests. *Nature Communications*. doi: 10.1038/s41467-021-20994-y
Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Gonzalez, M. R., Kilibarda, M., Blagotić, A., ... Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE*, *12*(2), 1–40. doi: 10.1371/journal.pone.0169748
Hoylman, Z. H., Jencso, K. G., Hu, J., Martin, J. T., Holden, Z. A., Seielstad, C. A., & Rowell, E. M. (2018). Hillslope Topography Mediates Spatial Patterns of Ecosystem Sensitivity to Climate. *Journal of Geophysical Research: Biogeosciences*, *123*(2), 353–371. doi: 10.1002/2017JG004108
Khan, A., Stoy, P., Douglas, J., Anderson, M., Diak, G., Otkin, J., ... McCorkel, J. (2021). Reviews and syntheses: Ongoing and emerging opportunities to improve environmental science using observations from the Advanced Baseline Imager on the Geostationary Operational Environmental Satellites. *Biogeosciences*. doi: 10.5194/bg-2020-454
Koirala, S., Jung, M., Reichstein, M., de Graaf, I. E., Camps-Valls, G., Ichii, K., ... Carvalhais, N. (2017). Global distribution of groundwater-vegetation spatial covariation. *Geophysical Research Letters*, *44*(9), 4134–4142. doi: 10.1002/2017GL072885
Koirala, S., Kim, H., Hirabayashi, Y., Kanae, S., & Oki, T. (2019). Sensitivity of global hydrological simulations to groundwater capillary flux parameterizations. *Water Resources Research*, *55*(1), 402–425. doi: 10.1029/2018WR023434
Küçük, Ç., Koirala, S., Carvalhais, N., Miralles, D. G., Reichstein, M., & Jung, M. (2020). Characterising the response of vegetation cover to water limitation in Africa using geostationary satellites. *Earth and Space Science Open Archive*, 38. doi: 10.1002/essoar.10504964.2
Kumar, I. E., Venkatasubramanian, S., Scheidegger, C., & Friedler, S. (2020, 13–18

Jul). Problems with shapley-value-based explanations as feature importance measures. In H. D. III & A. Singh (Eds.), *Proceedings of the 37th international conference on machine learning* (Vol. 119, pp. 5491–5500). PMLR. Retrieved from `https://proceedings.mlr.press/v119/kumar20e.html`

Lal, R. (2019). Carbon cycling in global drylands. *Current Climate Change Reports*, *5*(3), 221–232. doi: 10.1007/s40641-019-00132-z

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... Lee, S.-I. (2020). From local explanations to global understanding with explainable ai for trees. *Nature Machine Intelligence*, *2*(1), 2522-5839. doi: 10.1038/s42256-019-0138-9

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems* (pp. 4768–4777).

Maestre, F. T., Benito, B. M., Berdugo, M., Concostrina-Zubiri, L., Delgado-Baquerizo, M., Eldridge, D. J., ... Soliveres, S. (2021). Biogeography of global drylands. *New Phytologist*, *231*(2), 540–558. doi: 10.1111/nph.17395

Maestre, F. T., Salguero-Gómez, R., & Quero, J. L. (2012). It is getting hotter in here: Determining and projecting the impacts of global environmental change on drylands. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1606), 3062–3075. doi: 10.1098/rstb.2011.0323

Maxwell, R. M., & Condon, L. E. (2016). Connections between groundwater flow and transpiration partitioning. *Science*, *353*(6297), 377–380. doi: 10.1126/science.aaf7891

Molnar, C. (2019). *Interpretable machine learning.* (`https://christophm.github.io/interpretable-ml-book/`)

Mu, M., De Kauwe, M. G., Ukkola, A. M., Pitman, A. J., Gimeno, T. E., Medlyn, B. E., ... Ellsworth, D. S. (2021). Evaluating a land surface model at a water-limited site: implications for land surface contributions to droughts and heatwaves. *Hydrology and Earth System Sciences*, *25*(1), 447–471. doi: 10.5194/hess-25-447-2021

Nash, E., & Sutcliffe, V. (1970). River flow forecasting through conceptual models Part I - A discussion of principles. *Journal of Hydrology*, *10*, 282–290. doi: 10.1016/0022-1694(70)90255-6

Newman, B. D., Wilcox, B. P., Archer, S. R., Breshears, D. D., Dahm, C. N., Duffy, C. J., ... Vivoni, E. R. (2006). Ecohydrology of water-limited environments: A scientific vision. *Water Resources Research*, *42*(6), 1–15. doi: 10.1029/2005WR004141

Prăvălie, R. (2016). Drylands extent and environmental issues. A global approach. *Earth-Science Reviews*, *161*, 259–278. doi: 10.1016/j.earscirev.2016.08.003

Reynolds, J. F., Stafford Smith, D. M., Lambin, E. F., Turner, B. L., Mortimore, M., Batterbury, S. P., ... Walker, B. (2007). Ecology: Global desertification: Building a science for dryland development. *Science*, *316*(5826), 847–851. doi: 10.1126/science.1131634

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Journal of Applied Physics*, *1*(5), 318–333. doi: 10.1063/1.1745010

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., ... Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. doi: 10.1111/ecog.02881

Rodriguez-Iturbe, I., Chen, Z., & Rinaldo, A. (2021). On the fractal structure of soil moisture fields. *Advances in Water Resources*, *147*(October 2020), 103826. doi: 10.1016/j.advwatres.2020.103826

Rodriguez-Iturbe, I., & Porporato, A. (2009). *Ecohydrology of water-controlled ecosystems: Soil moisture and plant dynamics.* Cambridge University Press. doi: 10.1017/CBO9780511535727

Roebroek, C. T. J., Melsen, L. A., Hoek van Dijke, A. J., Fan, Y., & Teuling, A. J. (2020). Global distribution of hydrologic controls on forest growth. *Hydrology and Earth System Sciences Discussion*, 1–22. doi: 10.5194/hess-2020-32

Rudin, C., Chen, C., Chen, Z., Huang, H., Semenova, L., & Zhong, C. (2021). Interpretable machine learning: Fundamental principles and 10 grand challenges. *CoRR*, *abs/2103.11251*. Retrieved from `https://arxiv.org/abs/2103.11251`

Saxton, K. E., & Rawls, W. J. (2006). Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions. *Soil Science Society of America Journal*, *70*(5), 1569–1578. doi: 10.2136/sssaj2005.0117

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379-423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, *2*(28), 307–317.

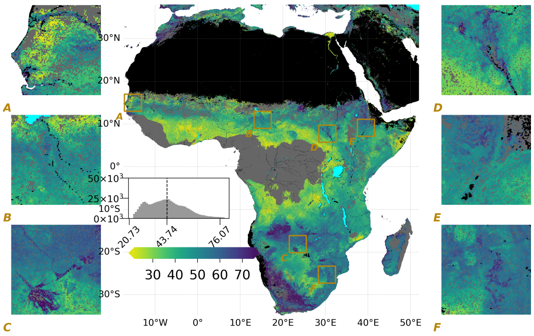
Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences*, *116*(4), 1–12. doi: 10.1029/2011JG001708

Tai, X., Anderegg, W. R., Blanken, P. D., Burns, S. P., Christensen, L., & Brooks, P. D. (2020). Hillslope Hydrology Influences the Spatial and Temporal Patterns of Remotely Sensed Ecosystem Productivity. *Water Resources Research*, *56*(11), 1–13. doi: 10.1029/2020WR027630

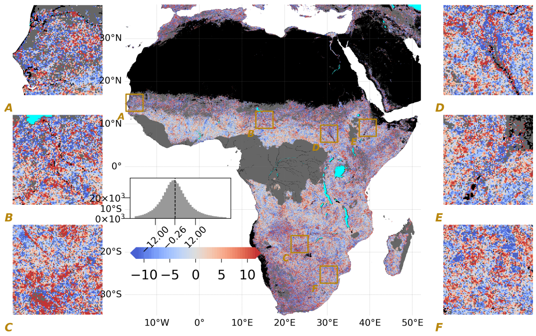
Tootchi, A., Jost, A., & Ducharne, A. (2019). Multi-source global wetland maps combining surface water imagery and groundwater constraints. *Earth System Science Data*, *892657*, 189–220. doi: 10.5194/essd-11-189-2019

Van Dijk, A. I., Schellekens, J., Yebra, M., Beck, H. E., Renzullo, L. J., Weerts, A., & Donchyts, G. (2018). Global 5 km resolution estimates of secondary evaporation including irrigation through satellite data assimilation - DISCUSSION. *Hydrology and Earth System Sciences*, *22*(9), 4959–4980. doi: 10.5194/hess-22-4959-2018

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M. (2019). MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset. *Water Resources Research*, *55*(6), 5053–5073. doi: 10.1029/2019WR024873



(a)



(b)

Figure 1: Maps of (a) model output (λ_m), in days, where larger values of λ (blue) indicate slower decay (b) residual of the model ($\lambda - \lambda_m$), in days, where positive values (red) indicate underestimation.

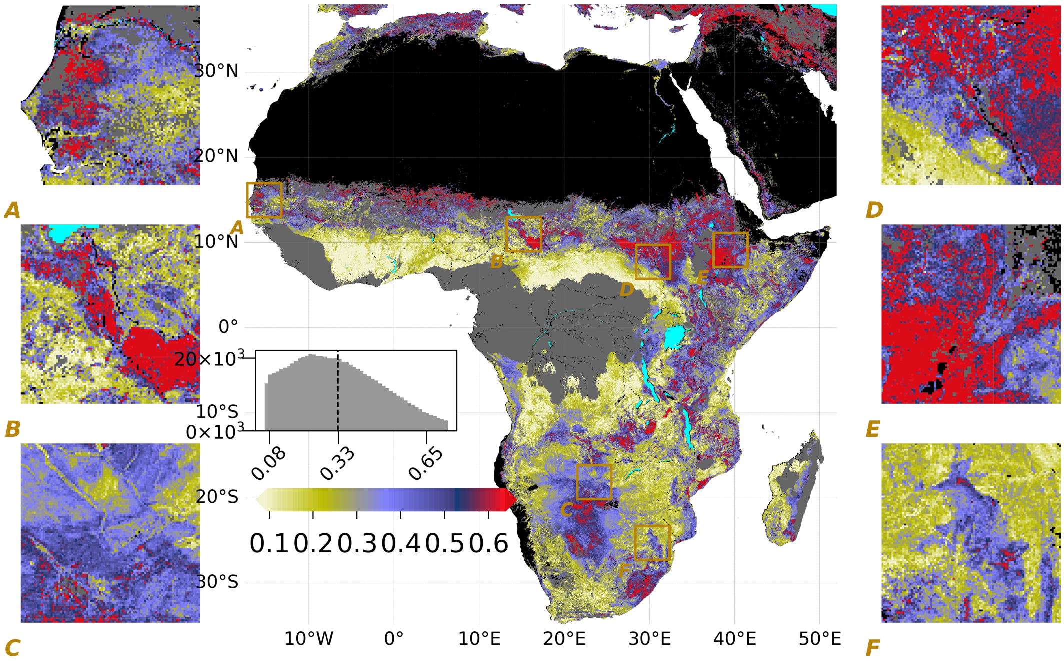


Figure 2: Spatial variations of the normalised importance of land on λ ($\Phi_{land-total}$) as output of Eq. 2 where larger (blue to red) values indicate higher importance of land parameters

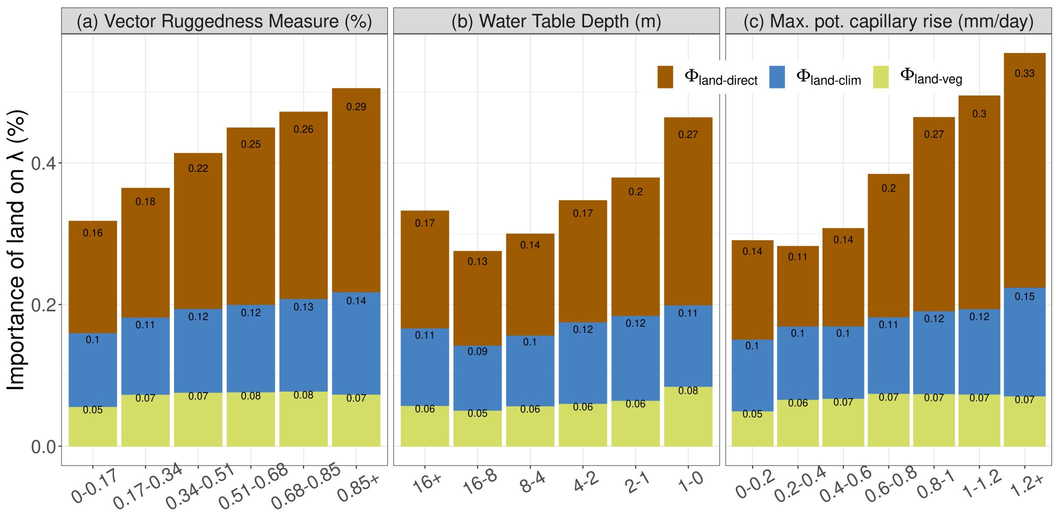


Figure 3: Normalised importance of land (same as Fig. 2) with change in Vector Ruggedness Measure (VRM), Water Table Depth (WTD), and maximum potential upwards capillary flux 1 meter above water table depth (I_{cap}). Y-axis shows the total land effects ($\Phi_{land-total}$) even though bars are coloured and annotated to show its components as direct effects ($\Phi_{land-direct}$) and interaction effects with climate ($\Phi_{land-clim}$) and vegetation ($\Phi_{land-veg}$), using Eq. 2.

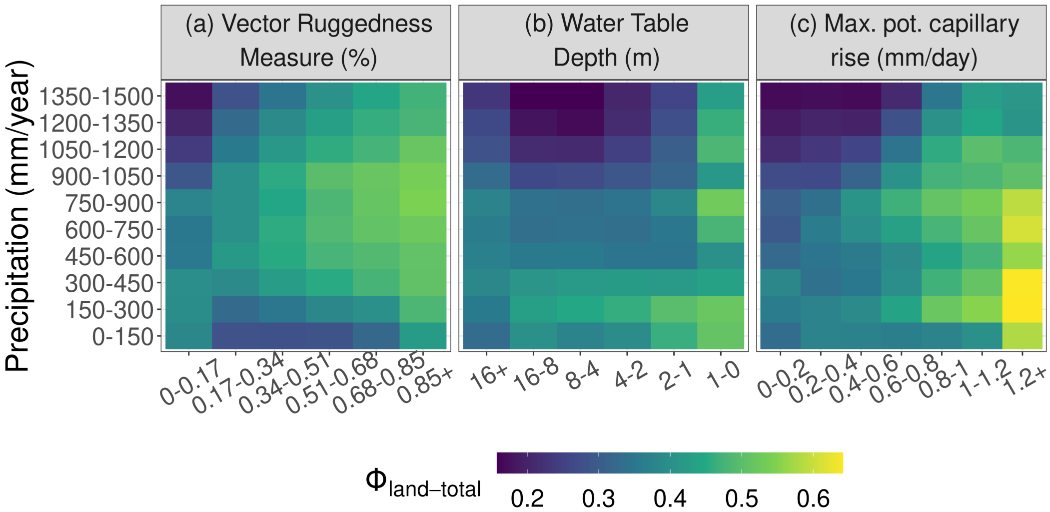


Figure 4: Effects of aridity on the importance of land parameters (see Eq. 2) with change in Vector Ruggedness Measure (VRM), Water Table Depth (WTD), and maximum potential upwards capillary flux 1 meter above water table depth (I_{cap}).