

Remote sensing-based vegetation and soil moisture constraints reduce irrigation estimation uncertainty

Wanshu Nie¹, Sujay V. Kumar², Rajat Bindlish², Pang-Wei Liu^{2,3}, Shugong Wang⁴

¹Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland, USA

²Hydrological Science Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

³Science Systems and Applications Inc., Lanham, Maryland, USA

⁴Applied Research Associates Inc., Raleigh, North Carolina, USA

Corresponding author: Wanshu Nie (nwanshu1@jhu.edu)

Abstract

Understanding the human water footprint and its impact on the hydrological cycle is essential to inform water management under climate change. Despite efforts in estimating irrigation water withdrawals in earth system models, uncertainties and discrepancies exist within and across modeling systems conditioned by model structure, irrigation parameterization, and the choice of input datasets. Achieving model reliability could be much more challenging for data-sparse regions, given limited access to ground truth for parameterization and validation. Here, we demonstrate the potential of utilizing remotely sensed vegetation and soil moisture observations in constraining irrigation estimation in the Noah-MP land surface model. Results indicate that the two constraints together can effectively reduce model sensitivity to the choice of irrigation parameterization by 7% - 43%. It also improves the characterization of the spatial patterns of irrigation and its impact on evapotranspiration and surface soil moisture by correcting for vegetation conditions and irrigation timing. This study highlights the importance of utilizing remotely sensed soil moisture and vegetation measurements in detecting irrigation signals and correcting for vegetation growth. Integrating the two remote sensing datasets into the model provides an effective and less feature engineered approach to constraining the uncertainty of irrigation modeling. Such strategies can be potentially transferred to other modeling systems and applied to regions across the globe.

Keywords: irrigation, remote sensing, data assimilation, vegetation, soil moisture

Introduction

Irrigated agriculture contributes to 40% of total food produced worldwide. Freshwater withdrawals from both surface and groundwater systems allow for production intensification and crop diversification as compared to rainfed agriculture and cover about 20% of cultivated land over the globe [FAO, 2020]. The understanding of irrigation in altering the water and carbon cycles [Baker et al., 2012; Scanlon et al., 2012; Troy et al., 2015; Asoka et al., 2017; Rodell et al., 2018; Zaveri and Lobell, 2019; Jasechko et al., 2017], repartitioning the surface energy balance [Kueppers et al., 2007; Lobell et al., 2008; Pokhrel et al., 2012; Lawston et al., 2020], its sensitivity to climate variability [Schewe et al., 2014; Russo and Lall, 2017; Condon et al., 2020], and its reflection and interaction with the human water management strategies [Guan et al., 2017; Lund et al., 2018; Wang et al., 2018; Deines et al., 2019] has been advanced in recent decades. Such irrigation related processes and feedbacks have been increasingly recognized as an important component that should be included in Earth System Models (ESMs), advancing the understanding of the water demand shift under climate change, thus informing preparedness for water and food security [Döll et al., 2012; Wada et al., 2017; de Graaf et al., 2019; Condon et al., 2019].

Despite the importance of reliable irrigation estimates for management of global freshwater withdrawals, large uncertainties exist among model estimates, surveys, and reports across scales [Huang et al., 2018; Puy et al., 2021]. These uncertainties can be sourced from the choice of acquiring approach, input datasets, modeling structure, irrigation parameterization approaches, among others. Irrigated area explains more than half of the irrigation variability at the global scale; however, it remains a highly uncertain variable for some regions [Puy et al., 2021]. Climate forcing is another factor contributing to the uncertainty of modeling irrigation. Wisser et al. [2008] demonstrated that choice of the forcing input can lead to as large as 70% variability in irrigation estimation relative to mean values at the national level. Additionally, features such as crop types, irrigation efficiency, and soil hydraulic parameters can contribute to irrigation uncertainty in various ways [Puy et al., 2021].

While most irrigation uncertainty assessments have focused on comparisons across input sources and differences among models [Wisser et al., 2008; Wada et al., 2013; Scanlon et al., 2018], few studies have quantified irrigation uncertainty from the parameterizations within a single model [DeJonge et al., 2012; Lawston et al., 2015]. The model parameters to which irrigation estimation may be potentially sensitive are often determined on a strong empirical basis or calibrated to variables that may not be directly relevant to irrigation amounts. Some models have a more sophisticated representation of irrigation where processes such as prognostic vegetation, soil infiltration, and runoff generation all play important roles in irrigation estimation and its associated impact on other fluxes. For instance, in most irrigation schemes in land surface models [Haddeland et al., 2006; Ozdogan et al., 2010; Harding and Snyder, 2012; Pokhrel et al., 2012; Leng et al., 2013], the triggering of irrigation depends on vegetation condition while the irrigation water use is based on a soil moisture deficit approach. In such models, physics associated with vegetation conditions and soil moisture states are essential in determining the timing and magnitude of irrigation. Consequently, both the parameters in the irrigation schemes and the physics that are linked to irrigation may introduce uncertainties in the irrigation estimation. Though many studies have quantified parameter sensitivities on various land surface models simulating natural processes [Demaria et al., 2007; Hou et al., 2015; Cuntz et al., 2016; Arsenault et al., 2018], specific examination of the sensitivities and uncertainties related to anthropogenic water use such as irrigation has not been conducted.

Incorporating remote sensing (RS) measurements via data assimilation or parameter estimation have been proven to benefit irrigation estimation by constraining or correcting for model states such as soil moisture [Felfelani *et al.*, 2018; Lei *et al.*, 2020], vegetation, and groundwater [Nie *et al.*, 2019]. However, these studies have not investigated how effective these constraints might be given the underlying parameter sensitivities. In other words, the effectiveness of utilizing RS measurements may be parameter-dependent and a configuration which worked well for one region might show nontransferable skill for another [Nie *et al.*, 2021]. Quantifying the dependence of RS constraints on key parameters is, therefore, necessary to understand how perturbations of model parameters may affect the irrigation estimation and whether and by what means the incorporation of RS constraints may affect the irrigation uncertainty brought by parameter perturbations.

The main goal of this paper is to assess the uncertainty of irrigation estimation to key irrigation parameters in an advanced land surface model – Noah-MP [Niu *et al.*, 2011; Yang *et al.*, 2011], and how vegetation and soil moisture RS constraints may affect the model uncertainty and provide guidance towards a more generalized and robust estimation of irrigation withdrawals. First, we identify and perturb three key irrigation parameters to generate an ensemble of 25 simulations representing irrigation uncertainty introduced by parameter selection. We then conduct four experiments with or without the two RS constraints, each consisting of 25 ensemble members (with a total of 100 experiments) and evaluate the irrigation estimation against the USDA report for the year 2018. The study also evaluates the propagation of irrigation uncertainty to evapotranspiration (ET) and soil moisture, given the importance of these variables in coupled land-atmosphere systems and water resources applications.

Methods

Noah-MP model and irrigation. We use the Noah-MP land surface model within the NASA Land Information System [LIS; Kumar *et al.*, 2006], with a configuration established in a previous study [Nie *et al.*, 2021], which has been evaluated over the Contiguous United States (CONUS) with reasonable skill in demonstrating irrigation impacted water and energy budgets. The irrigation scheme was originally built into Noah-MP based on Ozdogan *et al.* [2010] and further adapted by Nie *et al.* [2018; 2021] to account for the spatiotemporal variation of irrigation area and source water partition (More details can be found in the supplementary Text S1).

In this study, we create 25 combinations of the two vegetation (A, B) and one soil moisture (C) parameters (Supplementary Text S1 and Table 1) to generate the ensemble members for each experiment as they play a vital role in determining the timing, frequency, and amount of irrigation estimation. The values of A and B are selected based on choices made by previous practices [Ozdogan *et al.*, 2010; Lawston *et al.*, 2015; Nie *et al.*, 2018], while C is bounded by 0.4-0.58 with an interval of 0.02. The combination of these parameters is selected in a way that is general enough to represent an irrigation spread from a relatively loosely to a more strictly parameter-constrained condition. We then perform four experiments with or without RS constraints to investigate how these constraints may affect parameter-induced irrigation uncertainties.

Data sources and experimental design. We select two remote sensing datasets to explore whether and by what means incorporating them into a generic irrigation rule may constrain irrigation uncertainties. One dataset is the 500 m MODIS-MCD15A2H Version 6 LAI product available at 8-day interval [Myneni *et al.*, 2015]; and the other is the 1 km daily soil moisture (referred as RS soil moisture hereinafter) disaggregated from the Soil Moisture Active Passive [Chan *et al.*, 2018; SMAP; O'Neill *et al.*, 2019] product using the Thermal Hydraulic algorithm [THySM; Liu *et al.*, 2022]. Vegetation condition is an important indicator for irrigation season while the SMAP soil moisture retrievals, that we utilize for this study, has been proved to have higher information content [Kumar *et al.*, 2018] and capability of detecting irrigation signal over the CONUS [Lawston *et al.*, 2017a; Felfelani *et al.*, 2018]. We design four experiments with each experiment including 25 ensemble members populated by the parameter combinations of A, B, and C as mentioned above:

Open-loop experiment (OL_{irr}): We perform an open-loop run in which neither MODIS LAI nor the RS soil moisture is incorporated into the model. The prognostic vegetation module is turned on to simulate the vegetation growth on its own.

MODIS-based LAI constrained experiment (LAI_{DA}): The MODIS LAI dataset is incorporated into the model by assimilating the LAI values using the Ensemble Kalman Filter (EnKF) algorithm following Kumar *et al.* [2019] (More details are provided in supplementary Text S2).

RS soil moisture constrained experiment (SSM_{TC}): Irrigation demand computed in the model should be informed by soil moisture signal *prior* to irrigation instead of soil moisture condition that is already impacted by irrigation. Therefore, rather than assimilating RS soil moisture into the model which may violate the rationale of the irrigation routine, here we introduce RS soil moisture as an additional timing constraint to the existing irrigation rules. Irrigation is triggered if RS soil moisture is larger than the model predicted surface soil moisture by a certain threshold (Supplementary Text S3 provides more details).

MODIS-based LAI and RS soil moisture constrained experiment ($LAI-SSM$): In this experiment, both LAI and soil moisture constraints are incorporated into the model concurrently.

All simulations are conducted for the year 2018 at a spatial resolution of 0.125° considering the availability of both the SMAP dataset and USDA irrigation reports. The model is forced by the combination of the Integrated Multi-satellitE Retrievals for Global Precipitation Measurement [IMERG; Huffman *et al.*, 2015] near real-time final run and the NASA's Modern Era Reanalysis for Research and Applications version 2 [MERRA-2; Gelaro *et al.*, 2017] product.

Evaluation metrics. We use the irrigation survey of the year 2018 conducted by the USDA National Agricultural Statistics Service (USDA-NASS) to evaluate the simulated irrigation amount. We first identify five water resource regions (Figure 1) that cover most of the intensively irrigated areas for analyses, including the Pacific Northwest Region (PNR), the California Region (CR), the Missouri Region (MR), the Arkansas-White-Red Region (AWRR), and the Lower Mississippi Region (LMR). Further, to investigate the agreement between simulated and USDA-reported irrigation amount in terms of spatial pattern and magnitude, we fit a linear regression to the annual state-level total irrigation amount and the linear regression slope

for the 25 ensemble members for each experiment is reported. The ensemble members are then arranged based on the ascending order of the simulated irrigation total over the CONUS for OL_{irr} . We calculate the Spearman's rank correlation for the state-level irrigation amount to further explore how different remote sensing constraints affect the spatial correlation and across the ensemble members within each experiment.

To illustrate the dependency of ET uncertainty on the irrigation uncertainty, we fit the linear slope and estimated Spearman's correlation for state-averaged annual total ET between the simulations and the gridded 5 km daily Atmosphere-Land Exchange Inverse (ALEXI; Anderson et al., 2007) product. Only actively irrigated grid cells within each state are selected to calculate the state-average ET to eliminate the impact brought by non-irrigated areas. The same analysis is also conducted for ET within effective growing seasons. However, the results and conclusions are not sensitive to the choice of annual total versus growing season total as the magnitude of ET is dominated by the growing season ET in most cases.

Since the SMAP data is only used as a timing constraint for irrigation, we use it also as a reference dataset to explore and evaluate how irrigation uncertainty propagates into surface soil moisture by computing the Kullback-Leibler divergence [KLD; *Kullback and Leibler*, 1951] between SMAP and the simulations. The KLD metric measures the statistical distance between the probability distribution of two datasets, with lower KLD values indicating a closer match between the two datasets and vice versa. The climatological seasonal cycle is removed from the daily surface soil moisture before estimating KLD. The KLD values for actively irrigated grid cells are organized by water resource regions with the median, the lower, and the upper quartile reported.

Results and discussion

Irrigation spread reduced by applying the constraints

Five water resource regions are selected for analyses, which covered most of the intensively irrigated areas over the CONUS. Figure 1 demonstrates the simulated annual total irrigation water amount in 2018 for each experiment compared against the USDA report. Results suggest that irrigation estimation is sensitive to the choice of irrigation parameters in OL_{irr} for all of the selected water resource regions. The spread of estimated irrigation amount for Pacific Northwest, Missouri, Arkansas-White-Red, and Lower Mississippi encompassed the USDA reported amount, implying that with adjustments within the range of the perturbed parameters, one could possibly estimate the magnitude of the irrigation water demand comparable to the reported values. The tendency to underestimate for California region is, however, less likely to be overcome by adjusting these parameters, as the variability in irrigation estimation that adjusting parameters can produce is much smaller than its drift from the observations.

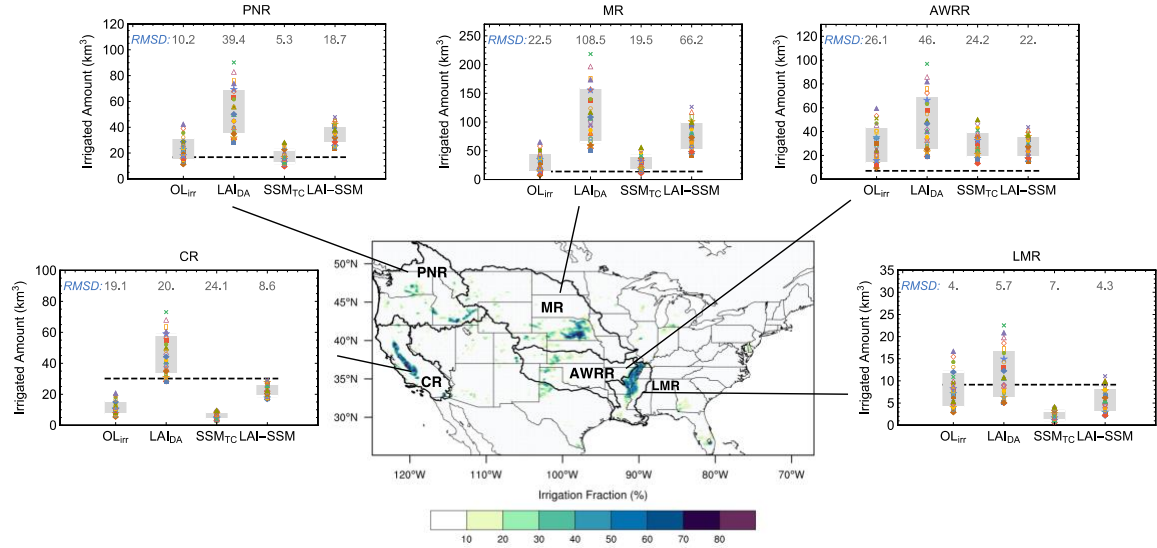


Figure 1 Comparison between simulated and USDA-reported total irrigation water amount in 2018 for the five water resource regions, the boundaries of which are highlighted on top of the irrigation fraction map over the CONUS. Colored symbols for each experiment set denote the amount simulated by the 25 simulation members with only the irrigation parameters being different. Grey bars represent the standard deviation within each experiment set centered on the mean. Black dashed lines denote the USDA-reported irrigation water amount for the corresponding region. Also shown are RMSD values with respect to the USDA report for each experiment.

Assimilating LAI (LAI_{DA}) increased the spread of the simulated amount among the members, implying that the sensitivity of the model to the selection of irrigation parameters would be increased by introducing vegetation information as a constraint for irrigation estimation. Moreover, LAI_{DA} also leads to larger estimated irrigation amounts, although this might be beneficial for California. The estimated amount for the other four regions, however, are overestimated when compared to the USDA report. Applying RS soil moisture as a timing constraint (SSM_{TC}) reduced the irrigation spread by limiting the occurrence of irrigation events within effective growing seasons. This brings the simulated amount in closer agreement with the USDA report for Pacific Northwest and Missouri; however, SSM_{TC} greatly underestimates the amount for California and Lower Mississippi, likely due to the misrepresentation of the coherence between vegetation and soil moisture states. Reduced spread and better estimation in amount are found especially for California and Arkansas-White-Red region when incorporating both constraints into the model ($LAI-SSM$), the RMSD of which has been reduced by 55%, 16%, respectively (Figure 1).

Generally, the impact of RS constraints on irrigation amount and spread indicates an increase by LAI_{DA} and a decrease by SSM_{TC} for all five regions. However, when combining the two constraints, irrigation amount and spread for Arkansas-White-Red region are further reduced while the changes are estimated in between LAI_{DA} and SSM_{TC} for the other four regions. Overall, our results reveal a degree of uncertainty in modeling irrigation amount from perturbing key irrigation parameters. However, conclusions regarding the accuracy of the estimation in magnitude relative to the USDA report are difficult to make as diverse sources (some subjective) contribute to the uncertainty of

the report. Moreover, the limited temporal scale of the available report limits the evaluation of the seasonality of irrigation patterns. The main finding of these results is that the vegetation and soil moisture inputs play relatively opposite roles in constraining the irrigation uncertainties stemming from the choice of irrigation parameters.

The role of LAI and soil moisture differ but are both important

To understand the role of LAI and soil moisture constraints in altering irrigation estimation, Figure 2 shows the seasonal profile of the estimated irrigation amount along with the greenness vegetation fraction (bottom row) for each experiment. We find that the contribution of LAI_{DA} is mainly towards modifying the seasonality of irrigation. Assimilating LAI shifts the peak and the magnitude of GVF, with an earlier peak growing season and much larger peak values for Pacific Northwest, California, and Missouri region as compared to OL_{irr}, implying that the default prognostic vegetation process within Noah-MP may have delayed the vegetation growth and underestimated the carbon assimilation amount. This deficiency has been reported in the literature [Ma *et al.*, 2017; Niu *et al.*, 2020], partially due to the simplified parameterization of the effect of soil water stress on transpiration and carbon assimilation, and root dynamics. For Arkansas-White-Red and Lower Mississippi region, LAI_{DA} largely modifies the evolution of vegetation growth. In particular, the summer peak and the following fast decay of irrigation amount are simulated in LAI_{DA}, which is represented differently in OL_{irr}. The difference in the seasonal pattern of GVF largely explains the seasonal difference in irrigation estimation in OL_{irr} and LAI_{DA}. However, this correction in the vegetation dynamics, especially in the magnitude, results in a greater spread among the ensemble members, as the irrigation amount estimation is sensitive to the rooting depth, which is parameterized by GVF.

Applying the RS soil moisture constraint has little impact on GVF. Minor impacts are found for California and Lower Mississippi region, in response to the interaction between irrigation impacted soil moisture stress and vegetation growth. For these regions, irrigation amount has been greatly reduced by SSM_{TC}, leading to less water availability, decreased leaf carbon assimilation, and thus smaller GVF. In fact, SSM_{TC} is more effective in constraining the irrigation spread, with the standard deviation among the 25 ensemble members being reduced by 22%-72% across the five water resource regions. The constraining in irrigation timing, by design, has also reduced the irrigation amount, which is controlled at a daily scale by LAI alone.

By incorporating both constraints into the model, we find although LAI-SSM does not constrain the spread as effective as SSM_{TC}, the standard deviation is still reduced by 7%-43% across regions except for Missouri. With the shift in the seasonal irrigation signal, LAI-SSM yields improvements, in particular, for California and Arkansas-White-Red region. These findings highlight that improved vegetation and soil moisture information is essential for irrigation estimation, and that the two remote sensing datasets have the potential to constrain irrigation uncertainty at different scales, with less parameterized formulations.

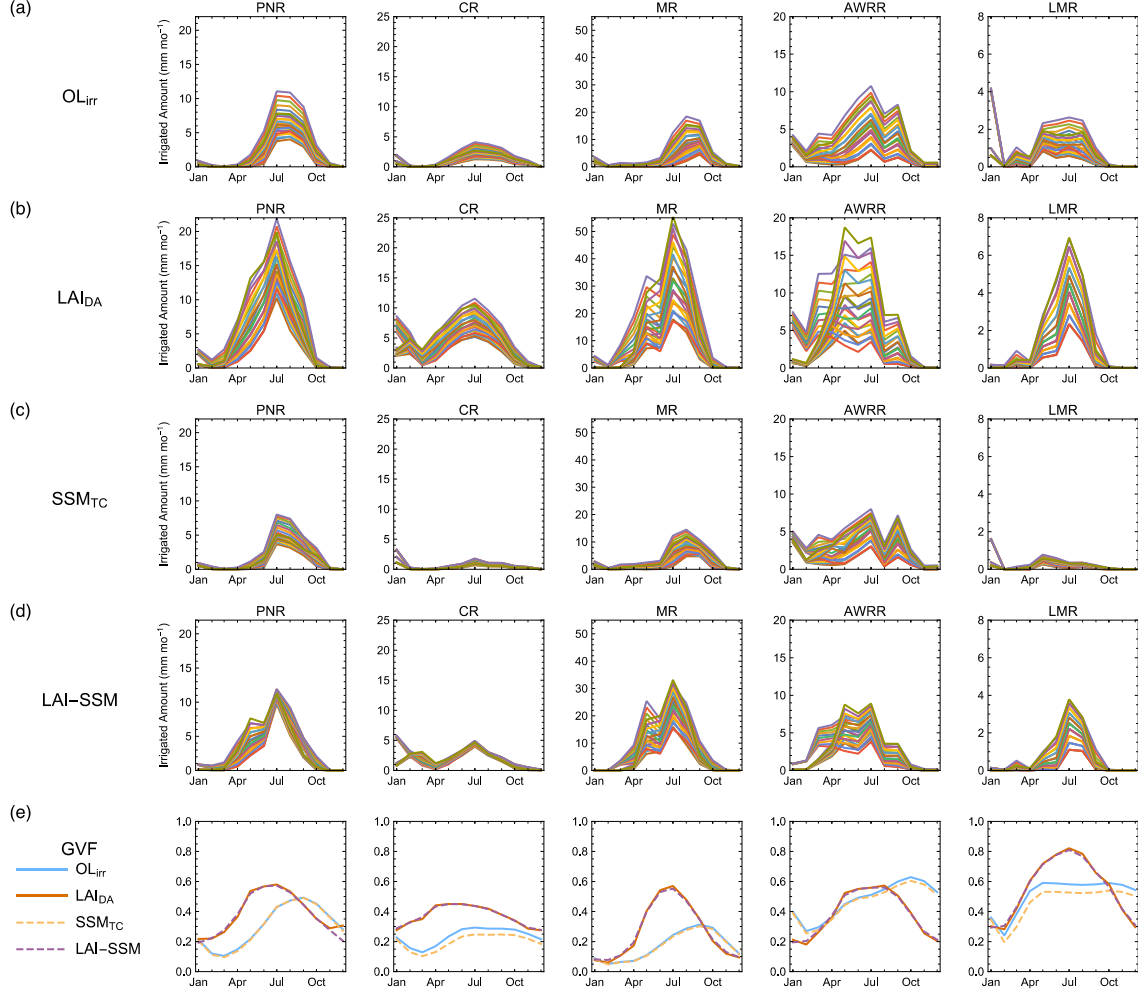


Figure 2 Evolution of monthly irrigation water amount averaged over the five water resource regions within the 25 members (colored lines) for (a) OL_{irr} , (b) LAI_{DA} , (c) SSM_{TC} , and (d) $LAI-SSM$. (e) shows the corresponding mean monthly GVF across the 25 members for each experiment set.

Improved spatial distribution in irrigation and ET

To examine how the two constraints affect the spatial distribution of irrigation, we apply 1) a standard linear regression and 2) Spearman's rank correlation to the state-averaged annual total irrigation amount against the USDA state-level report over the CONUS. The same metrics are also applied to simulated annual ET against the ALEXI ET dataset to understand how the impact on irrigation can propagate into energy fluxes (Figure 3). For the irrigation amount, the linear regression slope in OL_{irr} ranges between 0.4 -1.5 among the ensemble members. The strong influence of LAI data assimilation on vegetation conditions makes the model very sensitive to selected irrigation parameters. Depending on parameter combinations, the fitted slope can differ by up to four orders of magnitude, indicating a strong tendency to overestimate the amount as well as the spatial variability (Figure 3 (a)). However, although LAI_{DA} increases the irrigation

uncertainty and overestimated the amount, the benefit of correcting the vegetation growth toward improving the monotonic spatial distribution is robust, with a much higher Spearman's rank correlation than that of OL_{irr} (Figure 3 (b)). Applying the soil moisture constraint substantially stabilizes the fitted slope for irrigation, but its impact on rank correlation is negligible. By combining the two constraints, LAI-SSM narrows down the range of the fitted slope closer to 1 attributed mainly by the soil moisture constraint and maintains the high rank correlation benefited from the LAI constraint.

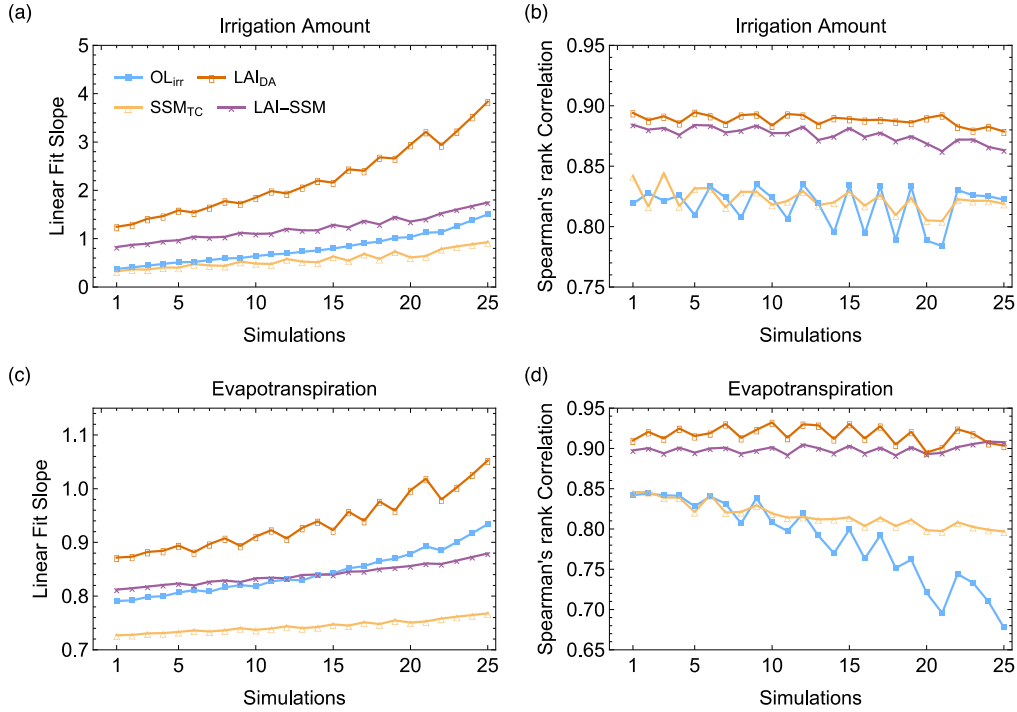


Figure 3 The linear fit slope and Spearman's rank correlation for state-averaged irrigation amount (a, b) compared against USDA-reported and ET (c, d) compared against ALEXI dataset. The simulation members are displayed in ascending order according to the total amount simulated by OL_{irr} .

The relative impact on ET in terms of the fitted slope is similar to that of irrigation amount, except that the magnitude of ET seems to be better represented by LAI_{DA} . Note that the ALEXI estimates are also subject to uncertainties and biases in representing the response of ET to irrigation and that the reliability of ALEXI ET is mainly in describing the spatial pattern and seasonality [Hain *et al.*, 2015]. Therefore, the comparison of ET magnitudes shown in Figure 3c is less reliable. Unlike the simulated irrigation, the rank correlation comparison is relatively insensitive to parameter choices. Figure 3d shows the degradation in the model's ability to capture the monotonic spatial distribution in ET. The correlation coefficient decreased from 0.84 to 0.68 as the simulation index increased. In this case, the LAI constraint (LAI_{DA} and $LAI-SSM$) not only improves the correlation of ET but also decouples the dependence of ET on the choice of parameters.

Impact on soil moisture distribution

Although SMAP data is utilized in the SSM_{TC} and LAI-SSM runs, the way it is incorporated into the model could be regarded as a loose constraint, as sub-daily and seasonal scale of soil moisture variation still heavily relies on the model physics itself. Therefore, we regard the RS soil moisture here as a reference dataset to examine whether and how irrigation affects surface soil moisture distribution. We use the KLD metric to measure the probability distribution difference between the simulated and RS soil moisture anomalies. Simulations with lower KLD values indicate a closer match with the RS soil moisture anomaly distribution. Figure 4 shows the median, along with the lower and the higher quartile of KLD for actively irrigated areas stratified by each water resource region.

For Pacific Northwest, Missouri, and Lower Mississippi, there is a subtle difference among the experiments and across the ensemble members. One possible explanation is that the irrigation signal is not strong enough to make a difference in the soil moisture distribution as these regions are relatively wet and the soil moisture anomaly distribution might be dominated by natural processes. The possible overlap of the growing season for irrigation with the rainy season in these regions causes soil moisture change to be dominated by the precipitation signal. In such instances, irrigation plays a secondary role in influencing soil moisture. The situation is different for California and Arkansas-White-Red regions where the overestimation of irrigation has a clear impact on the KLD pattern. LAI_{DA} results in a much greater increase in KLD as the simulation index gets larger, with less agreement to SMAP. On the other hand, by applying the soil moisture constraints, SSM_{TC} and LAI-SSM lead to smaller KLD values across the ensemble members, with the difference between the median being comparable and larger than the range of the quartiles within each experiment, especially for the California region. This indicates that the irrigation impact on surface soil moisture might be more obvious for dry regions such as the central Great Plains, and for regions like California, where the main growing season is mostly rain-free. Our results align with Lawston et al. [2017] showing that the irrigation signal detected by SMAP is strongest for the California central valley as the irrigation signal is consistent and extensive in the rain-free summer.

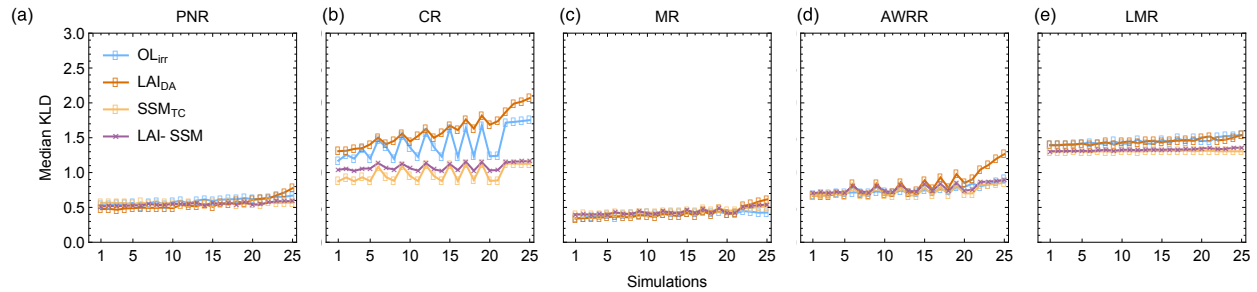


Figure 4 Kullback-Leibler divergence (KLD) estimates for all simulation members stratified by the five water resource regions. Solid lines represent the median and the corresponding shades represent the upper and lower quartile of the KLD.

Summary and conclusions

The results of the remotely sensed vegetation and soil moisture constraints impact on irrigation estimation, its uncertainty, and the associated fluxes indicate that both soil moisture and vegetation states are important for model realism in irrigation estimation in Noah-MP. Assimilating LAI

enables corrections to the vegetation growth seasonality while imposing the RS soil moisture constraint further informed the irrigation occurrences at daily scales. Combining the two constraints are important as applying LAI constraint alone may increase the model sensitivity to irrigation parameters, subjecting to potentially greater bias if a suitable set of parameters are not provided, while applying soil moisture constraint alone may misrepresent seasonal irrigation signal given the model weakness in prognostic vegetation process. In addition, soil moisture constraints also improve anomaly conditions, particularly in dry regions, and vegetation constraint leads to a more consistent agreement in the spatial pattern of ET.

Satellite remote sensing of vegetation conditions is relatively consistent and robust while it has been a challenge to date for passive soil moisture retrievals to detect irrigation. SMAP soil moisture retrievals have been proved to have higher information content [Kumar *et al.*, 2018] and capability of detecting irrigation signal over the CONUS [Lawston *et al.*, 2017; Felfelani *et al.*, 2018]. From this study, we conclude that the approach of incorporating both remotely sensed vegetation and soil moisture constraints to a generic soil moisture deficit-based irrigation routine can improve the spatial distribution of irrigation estimation and reduce its dependence on specific parameterizations. Our approach highlights the importance of correctly representing vegetation seasonality in irrigation estimation and provides a simple but more reliable way to extract information from remotely sensed soil moisture measurements for irrigation detection. Integrating the two constraints provides an effective and less parameter-dependent approach to reduce the uncertainty of irrigation modeling. With proper adaptation, the approach can be employed in other modeling systems and applied over different regions, to improve the quantification of anthropogenic water use for irrigation.

Data availability statement

The data that supports the findings of this study are openly available via the Johns Hopkins University Data Archive: <https://doi.org/10.7281/T1/0EYOMQ>.

Acknowledgements

This work was supported by the NASA sponsored Earth Information System (EIS) pilot project. Computational resources were provided by the NASA's Center for Climate Simulation (NCCS).

Conflict of interest

We declare no conflicts of interest.

References

Arsenault, K. R., G. S. Nearing, S. Wang, S. Yatheendradas, and C. D. Peters-Lidard (2018), Parameter Sensitivity of the Noah-MP Land Surface Model with Dynamic Vegetation, *J. Hydrometeor.*, 19(5), 815–830, doi:10.1175/jhm-d-17-0205.1.

- Asoka, A., T. Gleeson, Y. Wada, and V. Mishra (2017), Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India, *Nature Geosci*, 10(2), 109–117.
- Baker, J. M., T. J. Griffis, and T. E. Ochsner (2012), Coupling landscape water storage and supplemental irrigation to increase productivity and improve environmental stewardship in the U.S. Midwest, *Water Resour. Res.*, 48(5), 247, doi:10.1029/2011WR011780.
- Chan, S. K. et al. (2018), Development and assessment of the SMAP enhanced passive soil moisture product, *Remote Sensing of Environment*, 204, 931–941.
- Condon, L. E., A. L. Atchley, and R. M. Maxwell (2020), Evapotranspiration depletes groundwater under warming over the contiguous United States, *Nature Communications*, 11(1), 1–8.
- Condon, L. E., and R. M. Maxwell (2019), Simulating the sensitivity of evapotranspiration and streamflow to large-scale groundwater depletion, *Science Advances* 5.6, doi:10.1126/sciadv.aav4574.
- Cuntz, M., J. Mai, L. Samaniego, M. Clark, V. Wulfmeyer, O. Branch, S. Attinger, and S. Thober (2016), The impact of standard and hard-coded parameters on the hydrologic fluxes in the Noah-MP land surface model, *J. Geophys. Res. Atmos.*, 121(18), 10,676–10,700, doi:10.1002/2016JD025097.
- de Graaf, I. E. M., T. Gleeson, L. P. H. R. van Beek, E. H. Sutanudjaja, and M. F. P. Bierkens (2019), Environmental flow limits to global groundwater pumping, *Nature*, 574(7776), 90–94, doi:10.1038/s41586-019-1594-4.
- Deines, J. M., A. D. Kendall, J. J. Butler, and D. W. Hyndman (2019), Quantifying irrigation adaptation strategies in response to stakeholder-driven groundwater management in the US High Plains Aquifer, *Environ. Res. Lett.*, 14(4), 044014.
- DeJonge, K. C., J. C. Ascough, M. Ahmadi, A. A. Andales, and M. Arabi (2012), Global sensitivity and uncertainty analysis of a dynamic agroecosystem model under different irrigation treatments, *Ecological Modelling*, 231, 113–125.
- Demaria, E. M., B. Nijssen, and T. Wagener (2007), Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model, *J. Geophys. Res.*, 112(D11), doi:10.1029/2006JD007534.
- Döll, P., H. Hoffmann-Dobrev, F. T. Portmann, S. Siebert, A. Eicker, M. Rodell, G. Strassberg, and B. R. Scanlon (2012), Impact of water withdrawals from groundwater and surface water on continental water storage variations, *Journal of Geodynamics*, 59, 143–156.
- FAO (2020), AQUASTAT, FAOs Global Information System on water and agriculture.

- Felfelani, F., Y. Pokhrel, K. Guan, and D. M. Lawrence (2018), Utilizing SMAP soil moisture data to constrain irrigation in the Community Land Model, *Geophys. Res. Lett.*, 45(23), 12–892–12–902.
- Gelaro, R. et al. (2017), The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), *J. Climate*, 30(14), 5419–5454.
- Guan, K., B. Sultan, M. Biasutti, C. Baron, and D. B. Lobell (2017), Assessing climate adaptation options and uncertainties for cereal systems in West Africa, *Agricultural and Forest Meteorology*, 232, 291–305, doi:10.1016/j.agrformet.2016.07.021.
- Haddeland, I., D. P. Lettenmaier, and T. Skaugen (2006), Effects of irrigation on the water and energy balances of the Colorado and Mekong river basins, *Journal of Hydrology*, 324(1), 210–223.
- Hain, C. R., W. T. Crow, M. C. Anderson, and M. T. Yilmaz (2015), Diagnosing neglected soil moisture source–sink processes via a thermal infrared–based two-source energy balance model, *J. Hydrometeor.*, 16(3), 1070–1086.
- Harding, K. J., and P. K. Snyder (2012), Modeling the Atmospheric Response to Irrigation in the Great Plains. Part I: General Impacts on Precipitation and the Energy Budget, *J. Hydrometeor.*, 13(6), 1667–1686.
- Hou, T., Y. Zhu, H. Lü, E. Sudicky, Z. Yu, and F. Ouyang (2015), Parameter sensitivity analysis and optimization of Noah land surface model with field measurements from Huaihe River Basin, China, *Stochastic Environmental Research and Risk Assessment*, 29(5), 1383–1401.
- Huang, Z., M. Hejazi, X. Li, Q. Tang, C. Vernon, G. Leng, Y. Liu, P. Döll, S. Eisner, and D. Gerten (2018), Reconstruction of global gridded monthly sectoral water withdrawals for 1971–2010 and analysis of their spatiotemporal patterns, *Hydrology and Earth System Sciences Discussions*, 22(4), 2117–2133.
- Huffman, G. J., D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, P. Xie, and S.-H. Yoo (2015), NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG), Algorithm Theoretical Basis Document (ATBD) Version, 4, 26.
- Jasechko, S. et al. (2017), Global aquifers dominated by fossil groundwaters but wells vulnerable to modern contamination, *Nature Geoscience*, 10(6), 425–429, doi:10.1038/ngeo2943.
- Kueppers, L. M., M. A. Snyder, and L. C. Sloan (2007), Irrigation cooling effect: Regional climate forcing by land-use change, *Geophys. Res. Lett.*, 34(3).
- Kullback, S., and R. A. Leibler (1951), On information and sufficiency, *The annals of mathematical statistics*, 22(1), 79–86.
- Kumar, S. V., C. D. Peters-Lidard, Y. Tian, P. R. Houser, J. Geiger, S. Olden, L. Lighty, J. L. Eastman, B. Doty, and P. Dirmeyer (2006), Land information system: An interoperable

- framework for high resolution land surface modeling, *Environmental modelling & software*, 21(10), 1402–1415.
- Kumar, S. V., M. Jasinski, D. M. Mocko, M. Rodell, J. Borak, B. Li, H. K. Beaudoin, and C. D. Peters-Lidard (2019), NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment, *J. Hydrometeor.*, 20(8), 1571–1593.
- Kumar, S. V., P. A. Dirmeyer, C. D. Peters-Lidard, R. Bindlish, and J. Bolten (2018), Information theoretic evaluation of satellite soil moisture retrievals, *Remote Sensing of Environment*, 204, 392–400.
- Lawston, P. M., J. A. Santanello Jr, and S. V. Kumar (2017), Irrigation Signals Detected From SMAP Soil Moisture Retrievals, *Geophys. Res. Lett.*, 44(23), 11,860–11,867, doi:10.1002/2017GL075733.
- Lawston, P. M., J. A. Santanello Jr, B. F. Zaitchik, and M. Rodell (2015), Impact of irrigation methods on land surface model spinup and initialization of WRF forecasts, *J. Hydrometeor.*, 16(3), 1135–1154.
- Lawston, P. M., J. A. Santanello Jr, B. Hanson, and K. Arsensault (2020), Impacts of Irrigation on Summertime Temperatures in the Pacific Northwest, *Earth Interact.*, 24(1), 1–26, doi:10.1175/EI-D-19-0015.1.
- Lei, F. et al. (2020), Data assimilation of high-resolution thermal and radar remote sensing retrievals for soil moisture monitoring in a drip-irrigated vineyard, *Remote Sensing of Environment*, 239, 111622–16, doi:10.1016/j.rse.2019.111622.
- Leng, G., M. Huang, Q. Tang, W. J. Sacks, H. Lei, and L. R. Leung (2013), Modeling the effects of irrigation on land surface fluxes and states over the conterminous United States: Sensitivity to input data and model parameters, *J. Geophys. Res. Atmos.*, 118(17), 9789–9803.
- Liu, P.-W., R. Bindlish, P. E. O'Neill, B. Fang, V. Lakshmi, Z. Yang, M. Cosh, T. Bongiovanni, C. H. Collins, P. Starks, J. Prueger, D. D. Bosch, M. Seyfried, and M. R. Williams (2022), Thermal hydraulic disaggregation of SMAP soil moisture over the continental United States, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 4072–4092, doi: 10.1109/JSTARS.2022.3165644.
- Lobell, D. B., C. J. Bonfils, L. M. Kueppers, and M. A. Snyder (2008), Irrigation cooling effect on temperature and heat index extremes, *Geophys. Res. Lett.*, 35(9).
- Lund, J., J. Medellin-Azuara, J. Durand, and K. Stone (2018), Lessons from California's 2012–2016 Drought, *J. Water Resour. Plann. Manage.*, 144(10), 04018067–13, doi:10.1061/(ASCE)WR.1943-5452.0000984.

- Ma, N., G. Y. Niu, Y. Xia, X. Cai, Y. Zhang, Y. Ma, and Y. Fang (2017), A systematic evaluation of Noah-MP in simulating land-atmosphere energy, water, and carbon exchanges over the continental United States, *J. Geophys. Res. Atmos.*, 122(22), 12–245–12–268.
- Myneni, R., Y. Knyazikhin, and T. Park (2015), MCD15A2H MODIS/Terra+Aqua Leaf Area Index/FPAR 8-day L4 Global 500m SIN Grid V006 [Data set], NASA EOSDIS Land Processes DAAC.
- Nie, W. et al. (2021), Towards Effective Drought Monitoring in the Middle East and North Africa (MENA) Region: Implications from Assimilating Leaf Area Index and Soil Moisture into the Noah-MP Land Surface Model for Morocco, *Hydrology and Earth System Sciences*, 26(9), 2365–2386.
- Nie, W., B. F. Zaitchik, M. Rodell, S. V. Kumar, K. R. Arsenault, and H. S. Badr (2021), Irrigation water demand sensitivity to climate variability across the Contiguous United States, *Water Resour. Res.*, 57(3), e2020WR027738.
- Nie, W., B. F. Zaitchik, M. Rodell, S. V. Kumar, K. R. Arsenault, B. Li, and A. Getirana (2019), Assimilating GRACE Into a Land Surface Model in the Presence of an Irrigation-Induced Groundwater Trend, *Water Resour. Res.*, 55(12), 11274–11294, doi:10.1029/2019WR025363.
- Nie, W., B. F. Zaitchik, M. Rodell, S. V. Kumar, M. C. Anderson, and C. Hain (2018), Groundwater Withdrawals Under Drought: Reconciling GRACE and Land Surface Models in the United States High Plains Aquifer, *Water Resour. Res.*, 54(8), 5282–5299, doi:10.1029/2017WR022178.
- Niu, G. Y. et al. (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *J. Geophys. Res.*, 116(D12), 1381–19, doi:10.1029/2010JD015139.
- Niu, G. Y., Y. H. Fang, L. L. Chang, J. Jin, H. Yuan, and X. Zeng (2020), Enhancing the Noah-MP Ecosystem Response to Droughts with an Explicit Representation of Plant Water Storage Supplied by Dynamic Root Water Uptake, *J. Adv. Model. Earth Syst.*, 1–29, doi:10.1029/2020MS002062.
- O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, R. Bindlish, and J. Chaubell (2019), SMAP Enhanced L2 Radiometer Half-Orbit 9 km EASE-Grid Soil Moisture, Version 3, Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, doi:https://doi.org/10.5067/017XZSKMLTT2.
- Ozdogan, M., M. Rodell, H. K. Beaudoin, and D. L. Toll (2010), Simulating the effects of irrigation over the United States in a land surface model based on satellite-derived agricultural data, *Journal of Hydrometeorology*, 11(1), 171–184.
- Pokhrel, Y., N. Hanasaki, S. Koirala, J. Cho, P. J.-F. Yeh, H. Kim, S. Kanae, and T. Oki (2012), Incorporating anthropogenic water regulation modules into a land surface model, *J. Hydrometeor.*, 13(1), 255–269.

- Puy, A., E. Borgonovo, S. Lo Piano, S. A. Levin, and A. Saltelli (2021), Irrigated areas drive irrigation water withdrawals, *Nature Communications*, 12(1), 4525.
- Rodell, M., J. S. Famiglietti, D. N. Wiese, J. T. Reager, H. K. Beaudoin, F. W. Landerer, and M.-H. Lo (2018), Emerging trends in global freshwater availability, *Nature*, 557(7707), 651–659.
- Russo, T. A., and U. Lall (2017), Depletion and response of deep groundwater to climate-induced pumping variability, *Nature Geosci*, 10(2), 105–108.
- Scanlon, B. R., C. C. Faunt, L. Longuevergne, R. C. Reedy, W. M. Alley, V. L. McGuire, and P. B. McMahon (2012), Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley, *Proc Natl Acad Sci USA*, 109(24), 9320–9325.
- Scanlon, B. R., Z. Zhang, H. Save, A. Y. Sun, H. M. Schmied, L. P. Van Beek, D. N. Wiese, Y. Wada, D. Long, and R. C. Reedy (2018), Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data, *Proc Natl Acad Sci USA*, 115(6), E1080–E1089.
- Schewe, J., J. Heinke, D. Gerten, I. Haddeland, N. W. Arnell, D. B. Clark, R. Dankers, S. Eisner, B. M. Fekete, and F. J. Colón-González (2014), Multimodel assessment of water scarcity under climate change, *Proc Natl Acad Sci USA*, 111(9), 3245–3250.
- Troy, T. J., C. Kipgen, and I. Pal (2015), The impact of climate extremes and irrigation on US crop yields, *Environ. Res. Lett.*, 10(5), 054013.
- Wada, Y. et al. (2013), Multimodel projections and uncertainties of irrigation water demand under climate change, *Geophys. Res. Lett.*, 40(17), 4626–4632, doi:10.1002/grl.50686.
- Wada, Y., M. F. Bierkens, A. De Roo, P. A. Dirmeyer, J. S. Famiglietti, N. Hanasaki, M. Konar, J. Liu, H. M. Schmied, and T. Oki (2017), Human-water interface in hydrological modelling: current status and future directions, *Hydrol. Earth Syst. Sci.*, 21(8), 4169–4193.
- Wang, J., C. Song, J. T. Reager, F. Yao, J. S. Famiglietti, Y. Sheng, G. M. MacDonald, F. Brun, H. M. Schmied, and R. A. Marston (2018), Recent global decline in endorheic basin water storages, *Nature Geosci*, 11(12), 926–932.
- Wisser, D., S. Frolking, E. M. Douglas, B. M. Fekete, C. J. Vörösmarty, and A. H. Schumann (2008), Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets, *Geophys. Res. Lett.*, 35(24).
- Yang, Z. L., G. Y. Niu, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, L. Longuevergne, K. Manning, D. Niyogi, and M. Tewari (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins, *J. Geophys. Res.*, 116(D12).
- Zaveri, E., and D. B. Lobell (2019), The role of irrigation in changing wheat yields and heat sensitivity in India, *Nature Communications*, 10(1), 1–7.