# JASMIN: a high-resolution soil moisture analysis system for fire prediction

Vinodkumar[1,2], Imtiaz Dharssi[1,2] and Paul Fox-Hughes[1]

1. Bureau of Meteorology.

2. Bushfire and Natural Hazards CRC.

## Abstract

Soil moisture is found to be a key factor that influences fuel moisture content. Consequently, operational forest fire prediction systems typically include soil moisture as one of the inputs for fire behaviour calculations. The soil moisture input to these fire prediction models is usually provided in the form of moisture deficit. There is evidence that the current operational methods used in Australia for fire prediction perform poorly in estimating soil moisture status. A research project was initiated in partnership with the Bushfire and Natural Hazards Cooperative Research Centre to develop an advanced, state-of-the-art soil moisture analysis for Australia. Consequently, a prototype, high-resolution, land surface modelling-based soil moisture analysis called JULES based Soil Moisture Information (JASMIN) has been developed. JASMIN can provide hourly moisture estimates for four soil layers, at a spatial resolution of 5 km.

The present paper will discuss the evaluation of JASMIN carried out against observations from ground-based networks. Among the results, the mean Pearson's correlation for surface soil moisture across three in-situ networks is found to be between 0.78 and 0.85. We also focus on the research carried out to downscale the JASMIN product from 5 km to 1 km spatial resolution. The downscaling research is motivated by the desirable impact a higher resolution soil moisture product can provide for fire prediction, considering the high spatial variability in soil moisture and fuel moisture. We discuss the application of three downscaling algorithms: two regression-based methods and one with a theoretical basis. The three methods applied in the present study are based on the information derived from characterizing a two-dimensional surface temperature/vegetation index scatterplot domain obtained from thermal and optical remote sensing observations. We present an overview of the application of each method, along with an evaluation against ground-based soil moisture observations. Evaluation results indicate that the regression methods, in general, fail to capture the observed temporal variability. The theoretically based method, on the other hand, provides a temporal correlation of 0.81 and captures the skill of the parent JASMIN product.

## Introduction

In a fire prediction context, soil moisture status, usually provided in the form of moisture deficit, is a key parameter to assess the fuel availability. In Australia, there is evidence that the Keetch Byram Drought Index (KBDI) and Soil Dryness Index (SDI) methods used to estimate soil moisture deficit in operational fire prediction perform poorly (Vinodkumar and Dharssi 2017). A prototype, high resolution, land surface modelling system has been developed by the Bureau of Meteorology (Dharssi and Vinodkumar 2017) to provide soil moisture estimates with high accuracy and precision. This prototype system is based on the Joint UK Land Environment Simulator (JULES; Best et al. 2011) land surface model and is forced mainly by observation based meteorological analyses. The new system is called the JULES based Australian Soil Moisture Information (JASMIN) and estimates soil moisture at a spatial resolution of 5 km.

For applications like fire prediction, there is a requirement for soil moisture information at even higher spatial resolution than currently provided by JASMIN. A common practice to overcome such a problem is to employ downscaling methods to increase the spatial scale of the product. Recent advances in optical remote sensing have allowed researchers to use different remote sensing products that reflect soil moisture variability as ancillary information. A method based on a "universal triangle" concept is used in several studies where a relationship between soil moisture, vegetation index (VI) and surface radiant temperature (Ts) from optical remote sensing sensors is established. The universal triangle concept arises from the emergence of a triangular or trapezoidal shape when VI and Ts measures taken from heterogeneous areas are plotted in two-dimensional feature space – forming a Ts/VI scatterplot. Of the different land surface parameters, normalized difference vegetation index (NDVI) and land surface temperature (LST) are the most widely used. Theoretical and experimental studies have demonstrated the relationship between surface soil moisture, NDVI and LST for a given region under specific climatic conditions and land surface types.

Based on the triangular space paradigm, an empirical, polynomial fitting downscaling method was proposed by Piles et al. (2011) over south-eastern Australia to retrieve soil moisture at 1 km resolution from Soil Moisture and Ocean Salinity (SMOS) mission using NDVI and LST data from Moderate Resolution Imaging Spectro-radiometer (MODIS). Merlin et al. (2008; 2012) had developed a physics based method to downscale soil moisture by exploring the direct relationship existing between soil moisture and Soil Evaporative Efficiency (SEE; the ratio of actual to potential evaporation), leading to the emergence of the "Disaggregation based on Physical And Theoretical scale Change (DisPATCh)" model (Merlin et al. 2012). The DisPATCh method was found to yield a temporal correlation of 0.7 when compared to ground-based observations over the semi-arid Murrumbidgee catchment.

The present study explores the applicability of the multiple linear regression method discussed in Piles et al. (2011) and the DisPATCh method to downscale JASMIN soil moisture from 5 km to 1 km spatial resolution using MODIS LST and NDVI data. The main reason for selecting these methods is that they have been tested and documented to derive soil moisture information at 1 km spatial resolution over Australian regions. Further, the input data used in these methods are readily available. To investigate whether the skill of the multiple linear regression method can be improved further by regularization, we implemented the Least Absolute Shrinkage and Selection Operator (LASSO; Tibshirani et al. 1996) regression using the same feature variables used in the multiple linear regression method. The downscaling algorithms are only applied to the top JASMIN soil layer (0-10 cm). One of the main factors controlling the shape of Ts-VI scatter is the surface soil moisture. Studies have shown that the combined use of optical and thermal infrared data can be used to derive moisture estimates for the top 5 cm soil layer (e.g., Sandholt et al. 2002). Even though there are mismatches in scales for the soil column each method represents, the topmost soil layer in JASMIN is a good approximation to that the Ts-VI method represents.

# Results and discussions

## Verification of JASMIN against ground observations

The skill of JASMIN is compared against that of KBDI and SDI using ground observations from the CosmOz (Hawdon et al. 2014), OzNet (Smith et al., 2012) and OzFlux (Beringer et al. 2016) networks. For direct verification, all SM products and indices are converted to soil wetness (normalized between [0, 1]) using their own maximum and minimum values from respective long time series. Pearson's product-moment correlation (R), unbiased root mean square difference (ubRMSD) and bias metrics are used here to evaluate the skill of each product against in situ observations. The scores are computed for all stations and for the whole period where comparing data overlaps. Only scores for significant correlations with p-values < 0.001 are presented. In order to calculate correlations with seasonal effects removed, we compute the anomalies for each dataset using $\hat{\Theta}_{an}=\hat{\Theta}-\hat{\Theta}_{av}$, where $\hat{\Theta}_{av}$ is the mean and is calculated over a 31 day sliding window. The results are depicted as scatter plots (Figure 1).

JASMIN generally exhibits a stronger correlation compared to the other two models (Figure 1a). This is especially true over CosmOz and OzNet networks. The median correlation for JASMIN obtained against CosmOz, OzNet and OzFlux is 0.85, 0.81 and 0.78 respectively. JASMIN consistently display a strong positive correlation over CosmOz sites where R > 0.60 at all sites. Out of the total 45 sites in OzNet network, JASMIN display R > 0.60 for all except 4 sites. For OzFlux, JASMIN captures the temporal patterns well for 19 sites out of 21, with R > 0.60 at all these sites. KBDI exhibits a larger scatter in correlations compared to SDI. KBDI shows a relatively better performance over OzFlux compared to OzNet and CosmOz. A majority of OzFlux sites are situated in high rainfall regions, and some among them are in the tropics. KBDI is known to perform well in regions with warm climates and higher annual rainfall totals. This is typical of the region (south-eastern US) for which KBDI was designed and calibrated.

The lower ubRMSD in JASMIN compared to other two models is represented by the general clustering of points below the reference line in the respective scatter plot (Figure 1b). This indicates that the amplitude of short-term variations in observations is well captured by JASMIN compared to the other two models. The closer agreement of JASMIN and observed amplitudes is reflected by the lowest median scores of ubRMSD across all networks. KBDI generally shows large deviations from observations (Figure 1b). KBDI in fact has the largest median ubRMSD values for all networks and, in general, shows a wet bias - a result reported in earlier studies as well (e.g. Vinodkumar et al. 2017).

The ability of each model to capture the short-term fluctuations in observations is quantified by the anomaly correlation metric (Figure 1d). It is worth noting here that all models considered in the present study have the same resolution and are driven by the same precipitation analysis. Hence differences in fluctuations characterized by each model cannot be due to the difference in rainfall amounts for an event in the driving data. These differences, however, can be due to how each model represents surface energy and water balance processes. JASMIN is found to have a relatively higher anomaly correlation when compared to KBDI and SDI.

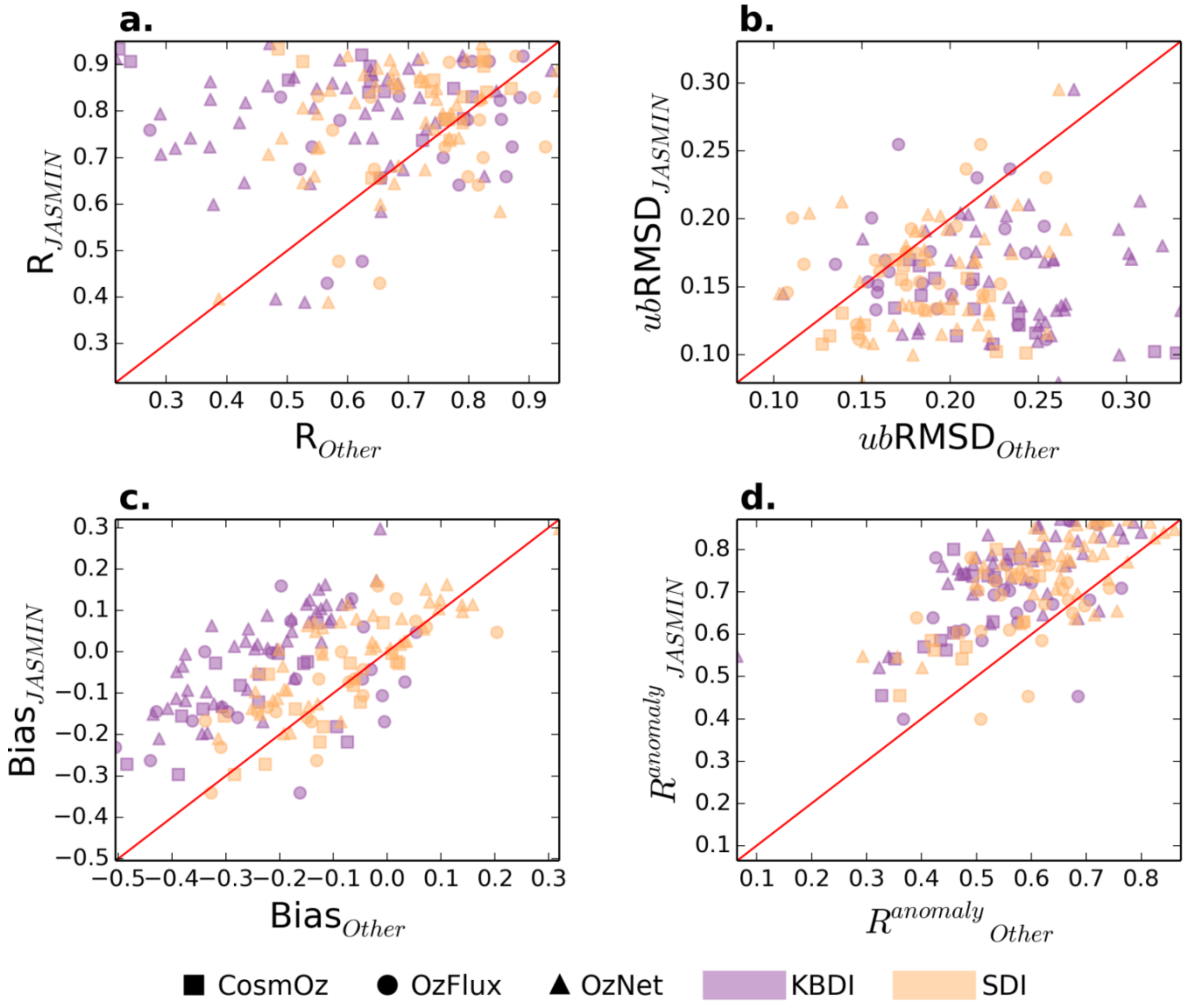


**Figure 1: Scatter plots depicting a) correlation, b) unbiased RMSD, c) bias and, d) anomaly correlation. The y-axis shows the skill scores of JASMIN against in-situ observation. The x-axis corresponds to the skill scores of the other two models (KBDI and SDI) against in-situ observations. Each colour represents a model type depicted on x-axis (i.e., KBDI and SDI). Each symbol represents an observation network type. The red line indicates equal skill between two products.**

Given the complexity of physical processes that govern surface soil moisture dynamics, these results indicate a robust modelling approach in JULES Land Surface Model. The governing complex physical processes also explain the low skill in KBDI and SDI. For example, neither of these models consider many physical factors including soil type, vegetation type, or terrain aspect which affect soil moisture. Further, no information on atmospheric controls of evapotranspiration such as net radiation, wind speed, or relative humidity is used.

## Verification of downscaled JASMIN soil moisture

This section discusses the temporal skill of each downscaling product against ground-based observations. The scores are computed using the same methodology discussed in the previous section. An evaluation of each model's skill over different land use / land cover (LULC) is presented in Figure 2. The LULC classification is made based on the land cover types over which the observation sites are located. We broadly classify the land cover types into forests, woodlands, grasslands and croplands. Of the 60 sites in total across three networks, 12 are classified as croplands, 12 as forests, 9 under woodlands, and the remaining 27 under grasslands.

The temporal skill is reduced when JASMIN is downscaled using the two regression-based methods. For example, the median values obtained by the LASSO method over woodlands for correlation, ubRMSD, bias and anomaly correlation are 0.41, 0.08, -0.08, and 0.26 respectively. For the multiple linear regression method, the above scores are 0.37, 0.11, -0.1 and 0.32 respectively. The LASSO method produces a higher skill than the multiple linear regression, highlighting the fact that there was some overfitting in the multiple linear regression method which is reduced in the LASSO method. Because of its safeguarding against noise, LASSO has a higher correlation and lower ubRMSD than the multiple linear regression method. This is demonstrated through the timeseries plot over the

Weany Creek site in the northern Queensland, which is part of the CosmOz network (Figure 3). This site is in a grazed open woodland with grassy and shrubby understory. The multiple linear regression method is found to have larger temporal variability than the LASSO method and the other two products (JASMIN and DisPATCh). This is particularly noticeable during the dry seasons where the multiple linear regression method shows large variability compared to the observations. A possible reason for this is the large sensitivity of estimated soil moisture in multiple linear regression to noise in the LST data. The uncertainties involved in the thermal infrared based LST retrievals are found to be about 2 K (Li et al., 2014). By applying regularization through LASSO, this sensitivity is reduced to some extent, but not to a point where the LASSO estimates match the temporal skill of the JASMIN product at 5km (Figure 2c).

In the case of DisPATCh, the temporal skill is similar to the JASMIN 5 km product and better than the other two downscaling methods. The average correlation of DisPATCh over the three networks is 0.81, identical to JASMIN. In the woodland, cropland and grassland cases, disaggregation either marginally improves or retains the mean R, bias and ubRMSD.

The similar skill of DisPATCh and JASMIN can be appreciated from the box and whiskers provided in Figure 2 Specifically, DisPATCh shows an increase in R and reduction in bias over the woodland sites. The good performance of the DisPATCh over woodlands is re-affirmed by the timeseries plot at the Weany Creek, which is an open woodland site (Figure 3d). The DisPATCh shows similar temporal variability to the observations and does not produce the large variability observed in the other two downscaling methods.

However, it is observed that DisPATCh has lower skill than the JASMIN product over forested sites, possibly due to the increase of random uncertainties attributable to the models and data used by DisPATCh. Studies have shown that DisPATCh performs better over low-density vegetated areas in semi-arid environments (Merlin et al. 2012). A possible reason for this behaviour is the weaker coupling between evaporation and surface soil-moisture in temperate (where most forested sites are located) than in semi-arid climates. Further, the presence of dense vegetation poses a challenge in the retrieval of the soil temperature from thermal infrared data. The vegetation water stress may increase the remotely sensed land surface temperature independent of near-surface soil moisture.

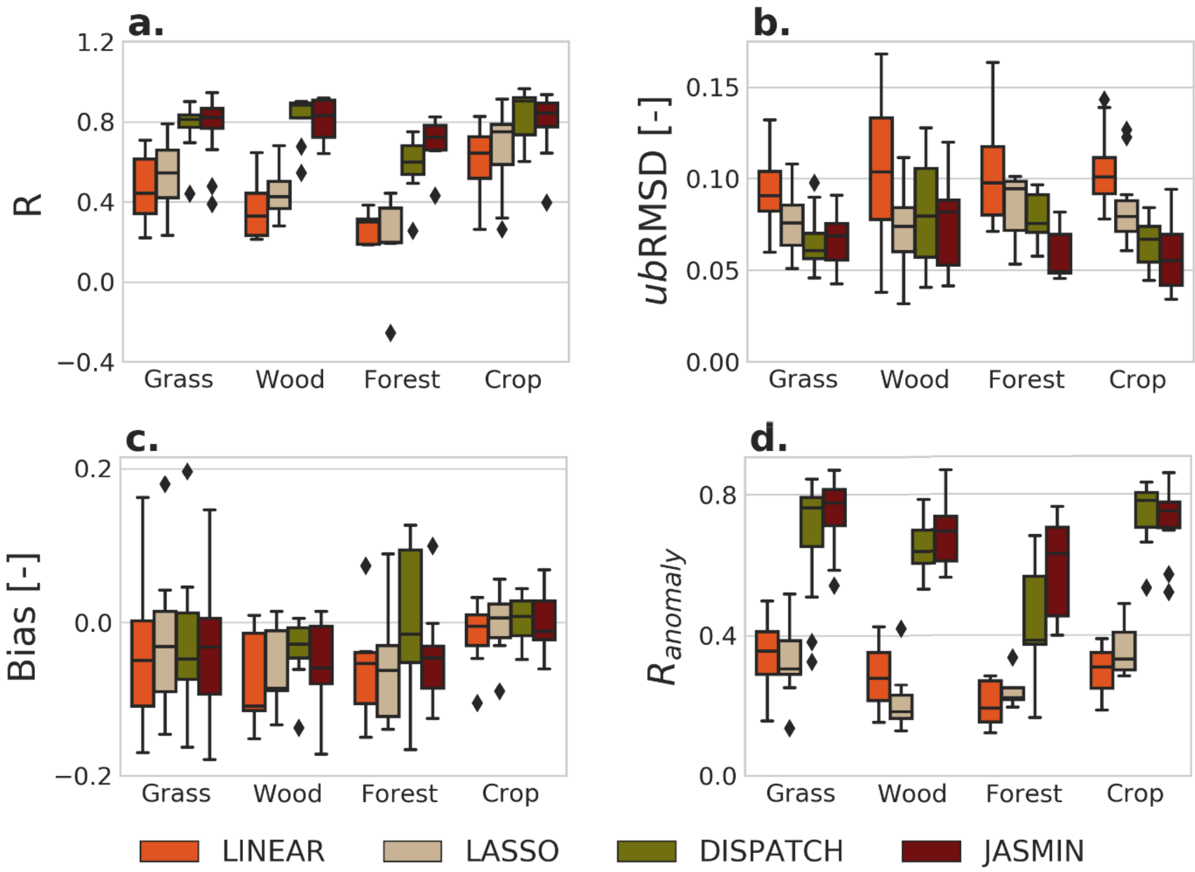


**Figure 2: Skill of soil wetness products over various land cover types: a) Pearson's correlation, b) unbiased RMSD, c) bias, and d) anomaly correlation. The grouping is done based on the land cover type of the observing site. The outliers are marked as diamonds. The orange boxes represent multiple linear regression method, light khaki colour represents LASSO method, the green boxes represent DisPATCh and the magenta coloured boxes represent the original JASMIN product at 5 km resolution.**

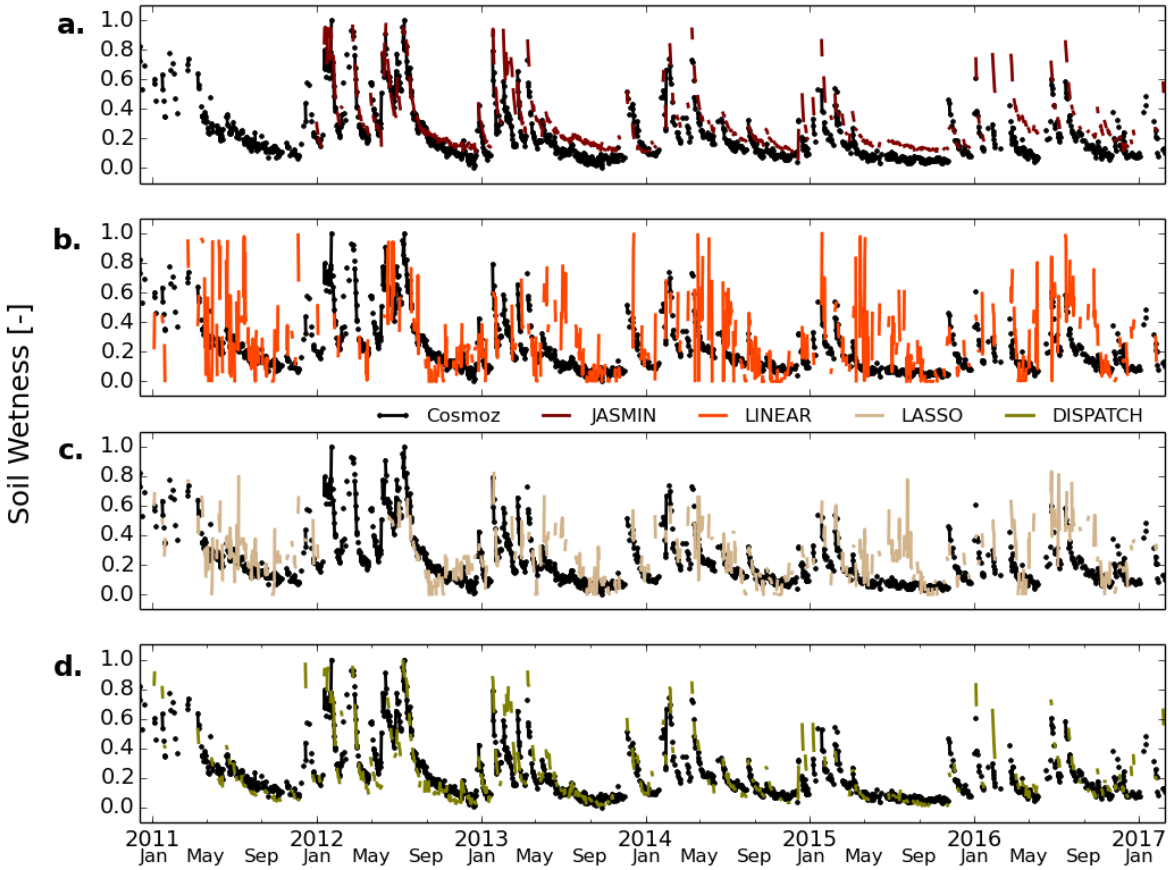


**Figure 3: Soil wetness time-series at the Weany Creek site in Queensland, part of the CosmOz network. The brown lines show JASMIN analyses at 5 km resolution, orange lines represent multiple linear regression method, light khaki depict LASSO method and the green line represents DisPATCh. The black dotted lines show the in-situ observations.**

## Summary

The present study underlines some of the limitations of traditional soil dryness indices in producing accurate soil moisture estimates, particularly for a shallow soil layer. One limitation of the traditional indices is that they use a single soil horizon to represent variations in both surface and root zone layers. The new JASMIN system can address gaps in the present operational methods by providing accurate soil moisture information in different layers. JASMIN has been shown to provide good skill in estimating soil moisture at both surface and root zone layers.

Results from the downscaling study indicate that it is feasible to improve the spatial resolution of JASMIN using all three disaggregating algorithms and preserve the general large-scale spatial structure seen in JASMIN soil moisture estimates. However, the seasonal means obtained at 1 km shows that each product displays characteristic soil moisture spatial variability at fine scales. Results from comparison with ground-based soil moisture measurements indicate that the regression methods degrade the temporal correlations and the ubRMSD scores. The DisPATCh method produces the best skill among the three algorithms tested here, and the skill scores from DisPATCh are comparable to those of the original JASMIN timeseries.

The low skill observed in regression methods possibly resulted from the large random errors attributable to the methods or uncertainties in the feature variables. It is worth noting that even the minimum and maximum limits applied to calculate the normalized LST and NDVI datasets (feature variables in the regression method) can introduce uncertainties in the downscaled soil moisture output. Further research is required to identify and minimize some of the uncertainties associated with both MODIS LST and NDVI datasets and to provide robust quality control.

Uncertainties in the MODIS input datasets have an important influence on the DisPATCh results as well, in addition to the uncertainties arising from the model assumptions and calibrations. It is found that calibration has a significant influence on the DisPATCh model behaviour. One aspect of DisPATCh that needs to be revisited is the modelling of soil moisture sensitivity to the soil evaporative efficiency. It is important to note that the DisPATCh algorithm is evolving and will continue to do so. Further work is required to test and evaluate the new ideas that will be developed in relation to DisPATCh and will be a focus of future research.

# References


Best, MJ, Pryor, M, Clark, DB, Rooney, GG, Essery, RLH, Menard, CB, Edwards, JM, Hendry, MA, Porson, A, Gedney, N, Mercado, LM, Sitch, S, Blyth, E, Boucher, O, Cox, PM, Grimmond, CSB & Harding, RJ 2011, 'The Joint UK Land Environment Simulator (JULES), model description - Part 1: Energy and water fluxes', *Geoscientific Model Development*, vol. 4, no. 3, pp. 677–699.

Beringer, J, Hutley, LB, McHugh, I, Arndt, SK, Campbell, D, Cleugh, HA, Cleverly, J, Resco de Dios, V, Eamus, D, Evans, B & Ewenz, C, 2016, 'An introduction to the Australian and New Zealand flux tower network–OzFlux', *Biogeosciences*, vol. 13, pp. 5895–5916.

Dharssi, I & Vinodkumar, 2017, 'A prototype high resolution soil moisture analysis system for Australia', *Bureau of Meteorology Research Report*, no. 026.

Hawdon, A, McJannet, D & Wallace, J, 2014, 'Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia', *Water Resources Research*, vol. 50, pp. 5029–5043.

Merlin, O, Rudiger, C, Al Bitar, A, Richaume, P, Walker, J P & Kerr, YH 2012, 'Disaggregation of SMOS Soil Moisture in South-eastern Australia', *IEEE Transactions in Geosciences and Remote Sensing*, vol. 50, no. 5, pp. 1556–1571.

Merlin, O, Walker, JP, Chehbouni, A & Kerr, Y 2008, 'Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency', *Remote Sensing of the Environment*, vol. 112, no. 10, pp. 3935–3946.

Piles, M, Camps, A, Vall-Llossera, M, Corbella, I, Panciera, R, Rudiger, C, Kerr, YH & Walker, J 2011, 'Downscaling SMOS-derived soil moisture using MODIS visible/infrared data', *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 3156–3166.

Sandholt, I, Rasmussen, K & Andersen, J 2002, 'A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status', *Remote Sensing of the Environment*, vol. 79, no. 2-3, pp. 213–224.

Smith, A, Walker, J, Western, A, Young, R, Ellett, K, Pipunic, R, Grayson, R, Siriwardena, L, Chiew, F & Richter, H 2012, 'The Murrumbidgee soil moisture monitoring network data set', *Water Resources Research*, vol. 48.

Tibshirani, R 1996, 'Regression shrinkage and selection via the lasso', *Journal of Royal Statistical Society*, vol. 58, pp. 267–288.

Vinodkumar & Dharssi, I 2017, 'Evaluation of daily soil moisture deficit used in Australian forest fire danger rating system', *Bureau of Meteorology Research Reports*, no. 022.

Vinodkumar & Dharssi, I, Bally, J, Steinle, P, McJannet, D & Walker, J 2017, 'Comparison of soil wetness from multiple models over Australia with observations, *Water Resources Research*, pp. 633–646.