

**Monitoring arid-land groundwater abstraction through  
optimization of a land surface model with remote  
sensing-based evaporation**

Dissertation by  
Oliver Miguel López Valencia

In Partial Fulfillment of the Requirements

For the Degree of  
Doctor of Philosophy

King Abdullah University of Science and Technology  
Thuwal, Kingdom of Saudi Arabia

February, 2018

## **EXAMINATION COMMITTEE PAGE**

The dissertation of Oliver Miguel López Valencia is approved by the examination committee.

Committee Chairperson: Matthew Francis McCabe

Committee Members: Himanshu Mishra, Ibrahim Hoteit, Justin Sheffield

©February, 2018

Oliver Miguel López Valencia

All Rights Reserved

## ABSTRACT

Monitoring arid-land groundwater abstraction through optimization of a land surface model with remote sensing-based evaporation

Oliver Miguel López Valencia

The increase in irrigated agriculture in Saudi Arabia is having a large impact on its limited groundwater resources. While large-scale water storage changes can be estimated using satellite data, monitoring groundwater abstraction rates is largely non-existent at either farm or regional level, so water management decisions remain ill-informed. Although determining water use from space at high spatiotemporal resolutions remains challenging, a number of approaches have shown promise, particularly in the retrieval of crop water use via evaporation. Apart from satellite-based estimates, land surface models offer a continuous spatial-temporal evolution of full land-atmosphere water and energy exchanges. In this study, we first examine recent trends in terrestrial water storage depletion within the Arabian Peninsula and explore its relation to increased agricultural activity in the region using satellite data. Next, we evaluate a number of large-scale remote sensing-based evaporation models, giving insight into the challenges of evaporation retrieval in arid environments. Finally, we present a novel method aimed to retrieve groundwater abstraction rates used in irrigated fields by constraining a land surface model with remote sensing-based evaporation observations. The approach is used to reproduce reported irrigation rates over 41 center-pivot irrigation fields presenting a range of crop dynamics over the course of one year. The results of this application are promising, with mean absolute errors below  $3 \text{ mm.day}^{-1}$ , bias of  $-1.6 \text{ mm.day}^{-1}$ , and a first rough estimate of total annual abstractions of  $65.8 \text{ Mm}^3$  (close to the estimated value using reported farm

data, 69.42  $Mm^3$ ). However, further efforts to address the overestimation of bare soil evaporation in the model are required. The uneven coverage of satellite data within the study site allowed us to evaluate its impact on the optimization, with a better match between observed and obtained irrigation rates on fields with higher frequency of available data. The inclusion of novel remote sensing sources (e.g. CubeSats) that offer higher frequencies and higher resolution can also be explored to improve the methodology, although further validation of these systems is needed. The developed framework has the potential to be used as a water management tool to monitor groundwater losses over large remote regions.

## ACKNOWLEDGEMENTS

I would like to acknowledge the invaluable help of colleagues, advisors and friends. Professor Matthew McCabe's direction and support helped shape this dissertation and helped me to clearly see the impact of this work. Rasmus Houborg had a direct involvement in many aspects of this work, and provided useful comments to improve this dissertation. Bruno Aragon, friend and colleague, provided direct help in the retrieval of high resolution thermal based evaporation, while Jorge Rosas's work in land surface temperature retrieval was an important element for these retrievals as well. Umer Altaf, research scientist at WDRC, offered direct help with the optimization aspect of this work. Hari Prasad, Ibrahim Hoteit and more generally the Earth Fluid Modeling and Prediction group, provided numerical weather prediction data, which was an important input requirement for this work. Jorge Rosas, Haleem Shah, Samir Al Masharawi and many other friends within the Hydrology, Agriculture and Land Observation group, as well as the Water Desalination and Reuse Center and in general, KAUST, helped me stay motivated throughout this time. Finally, I would like to thank my family as well: Sara and Mila Brady for the support and joy they provide me every day.

## TABLE OF CONTENTS

<b>Examination Committee Page</b>	<b>2</b>
<b>Copyright</b>	<b>3</b>
<b>Abstract</b>	<b>4</b>
<b>Acknowledgements</b>	<b>6</b>
<b>List of Figures</b>	<b>10</b>
<b>List of Tables</b>	<b>17</b>
<b>1 Introduction</b>	<b>18</b>
1.1 Remote sensing and numerical modeling opportunities for water management . . . . .	19
1.2 Agricultural development and groundwater resources in Saudi Arabia	21
1.3 Objectives . . . . .	23
<b>2 Estimating large-scale water storage variations in the Arabian Peninsula from gravity measurements</b>	<b>25</b>
2.1 Introduction: The Gravity Recovery and Climate Experiment (GRACE)	26
2.2 GRACE data processing . . . . .	27
2.2.1 Linking water movement to the Earth’s gravity field . . . . .	27
2.2.2 The averaging kernel . . . . .	28
2.2.3 The Gaussian smoothing kernel . . . . .	29
2.2.4 Removing correlated errors in the GRACE SH coefficients . .	30
2.2.5 Scaling the GRACE signal . . . . .	32
2.2.6 Other GRACE processing approaches . . . . .	34
2.3 Recent water storage variations on the Arabian Peninsula . . . . .	35
2.3.1 GRACE level 3 (gridded) product inter-comparison . . . . .	36
2.3.2 Water storage trends using SH coefficients . . . . .	37

2.3.3	Recent increase in irrigated land cover within the Saq aquifer region . . . . .	40
2.4	Conclusions . . . . .	40
<b>3</b>	<b>Measuring large-scale evaporation from space</b>	<b>43</b>
3.1	Introduction . . . . .	44
3.2	Evaporation products and auxiliary data . . . . .	48
3.2.1	MOD16 . . . . .	49
3.2.2	CSIRO-PML . . . . .	50
3.2.3	GLEAM . . . . .	51
3.2.4	Precipitation data: GPCP . . . . .	52
3.2.5	Runoff data . . . . .	53
3.3	Selection of study basins . . . . .	54
3.4	Methodology . . . . .	57
3.4.1	Spherical harmonic (SH) analysis of evaporation and precipitation data . . . . .	57
3.4.2	Regional spherical harmonic analysis . . . . .	58
3.4.3	Evaluating spatial agreement in spherical harmonics . . . . .	59
3.5	Results . . . . .	60
3.5.1	Assessing the consistency of evaporation products . . . . .	60
3.5.2	Basin-scale assessment . . . . .	61
3.5.3	Applying a phase lag to GRACE data . . . . .	68
3.6	Discussion . . . . .	71
3.6.1	Challenges to implementing hydrological consistency . . . . .	72
3.6.2	Temporal lag in terrestrial storage response . . . . .	75
3.6.3	Discriminating between satellite evaporation products . . . . .	76
3.7	Conclusions . . . . .	78
<b>4</b>	<b>High resolution modeling of water fluxes in an extreme arid environment</b>	<b>80</b>
4.1	Introduction . . . . .	81
4.2	High-resolution RS-based evaporation retrieval . . . . .	84
4.2.1	Description of the Two Source Energy Balance model (TSEB)	84
4.3	The Community Atmosphere Biosphere Land Exchange (CABLE) model	87
4.4	Meteorological inputs . . . . .	93
4.4.1	CABLE meteorological input module modification . . . . .	96
4.4.2	Evaluating the WRF-predicted rainfall . . . . .	96

4.5	Surface input data for the CABLE model . . . . .	97
4.5.1	High-resolution surface and vegetation characteristics based on Landsat 8 data . . . . .	98
4.6	Initializing the land surface model . . . . .	103
4.7	Evaporation from land surface modeling using WRF rainfall only: com- parison with remote sensing based evaporation at Tawdeehiya farm .	105
4.8	Conclusions . . . . .	109
<b>5</b>	<b>Estimating water use in irrigated agriculture: combining high-resolution remote sensing of evaporation and land surface modeling</b>	<b>113</b>
5.1	Introduction . . . . .	114
5.2	Methodology . . . . .	116
5.2.1	Optimization . . . . .	116
5.2.2	Tuning the SPSA parameters using synthetic data . . . . .	118
5.2.3	Testing the performance of SPSA using synthetic data . . . . .	123
5.2.4	Sensitivity to the initial guess . . . . .	123
5.2.5	Validation measures . . . . .	124
5.3	Results . . . . .	126
5.3.1	Comparison with reported farm data (EMI rates) . . . . .	126
5.3.2	Analysis of hydrologic fluxes . . . . .	129
5.3.3	Estimated total water consumption . . . . .	137
5.4	Discussion . . . . .	140
5.4.1	Tradeoff between selection of optimization periods and number of observations . . . . .	140
5.4.2	Accuracy of the observations and retrievals . . . . .	141
5.4.3	Optimization methodology . . . . .	142
5.4.4	Potential improvements to the methodology . . . . .	143
5.5	Summary and conclusions . . . . .	145
<b>6</b>	<b>Concluding Remarks</b>	<b>147</b>
6.1	Large-scale estimates of water depletion and evaporation from satellite observations . . . . .	148
6.2	Measuring agricultural water consumption using hyper-resolution re- mote sensing data and land surface modeling . . . . .	150
6.3	Directions and recommendations for further research . . . . .	152
	<b>References</b>	<b>155</b>

## LIST OF FIGURES

1.1	Left: aerial view of a center-pivot irrigation field. Right: Close-up showing a section of the pivot arm spraying water on the field. . . . .	21
1.2	Location of two agricultural regions with center-pivot irrigation in Saudi Arabia: Wadi As-Sirhan in the northwest, and the Tawdeehiya farm near Riyadh. Also shown is the outline of the Arabian Peninsula Aquifer System (APAS) and the Saq aquifer system. . . . .	22
2.1	Typical striping pattern found in unfiltered GRACE coefficients (GRACE CSR product for 2010/06/01 - 2010/06/30, converted to water storage anomalies using mean coefficients computed from 2002-2016) . . . . .	31
2.2	GRACE CSR product for 2010/06/01 - 2010/06/30 filtered with the de-striping filter and Gaussian smoothing with radius a) 0 <i>km</i> (no smoothing), b) 350 <i>km</i> , c) 500 <i>km</i> and d) 750 <i>km</i> . . . . .	31
2.3	Agricultural expansion in Al-Jowf, Saudi Arabia as evidenced by Landsat satellites imagery. . . . .	36
2.4	GRACE level 3 retrieval of terrestrial water storage anomalies (TWSA) averaged over 1° cells inside five regions within the MENA domain and the Amazon River basin. Note that the y-axis scales differ across each of the basins, illustrating the variability of TWSA across these regions.	37
2.5	Averaging kernels obtained for the Arabian Peninsula (left) and Saq aquifer (right) regions. . . . .	38
2.6	GRACE terrestrial water storage anomalies (TWSA) derived from spherical harmonic coefficients using averaging kernels of the five regions within the MENA domain and the Amazon River basin. . . . .	39

- 2.7 MODIS-derived land cover classification (500 *m* resolution) for five agricultural regions around the Saq aquifer. The cities of Tabuk, Hail, Buraydah and Riyadh are visible in black (urban classification). The largest agricultural region in terms of crop area is located on the North, in the region of Al-Jowf (A; see also figure 2.3). The Tawdeehiya farm, a region that will be explored in more detail in Chapters 4 and 5 is marked in red. . . . . 41
- 2.8 Crop area for the Saq aquifer and five agricultural regions estimated using the MODIS-derived product (500 *m* resolution) and Landsat data (30 *m* resolution). The increase within the Saq aquifer region is evident from 2004 – 2012, coinciding with the depletion period estimated using GRACE data. . . . . 41
- 3.1 Selected study basins used within the analysis. Criteria for the selection of basins included: predominantly arid climate (more than 50% areal coverage with any of the arid Köppen climates: BWk, BW<sub>h</sub>, BSk or BSh), size, geographical location and amplitude and trends in the water storage variations. . . . . 55
- 3.2 Average *P*, *E* and *Q* fluxes within the four study basins for the period 2003-2011. Three evaporation (*E*) data sets were used in this study, including CSIRO-PML, MOD16 and GLEAM (see table 3.1 for details). 56
- 3.3 Left: CSIRO-PML monthly evaporation (*E*) for April 2003. Right: the same data set after spherical harmonic analysis and synthesis of the evaporation data. Missing data are set to zero. The data appears smoothed because it is an approximation (equation 2.1) limited by  $l_{max}$ . There are also other effects such as ringing that are visible in regions where the data is close to zero. . . . . 58
- 3.4 Top: anomalies of the terrestrial water storage (TWSA) observed by GRACE (with 20 *mm* uncertainty bounds) and *P* – *E* using three global evaporation products over the Colorado river basin. Below: varying degree correlation measure ( $r_l$ ) with time and degree (from 25 to 40) using the three global evaporation products. The average  $r_l$  is shown as a time series (black line). The degree correlation measure can range from -1 to 1 as shown in the color scale on the secondary axis. 63

3.5	Top: anomalies of the terrestrial water storage (TWSA) observed by GRACE (with 20 mm uncertainty bounds) and $P-E$ using three global evaporation products over the Niger river basin. Below: varying degree correlation measure ( $r_l$ ) with time and degree (from 25 to 40) using the three global evaporation products. The average $r_l$ is shown as a time series (black line). The degree correlation measure can range from -1 to 1 as shown in the color scale on the secondary axis. . . . .	65
3.6	Top: anomalies of the terrestrial water storage (TWSA) observed by GRACE (with 20 mm uncertainty bounds) and $P - E$ using three global evaporation products over the Aral basin. Below: varying degree correlation measure ( $r_l$ ) with time and degree (from 25 to 40) using the three global evaporation products. The average $r_l$ is shown as a time series (black line). The degree correlation measure can range from -1 to 1 as shown in the color scale on the secondary axis. . . . .	67
3.7	Top: anomalies of the terrestrial water storage (TWSA) observed by GRACE (with 20 mm uncertainty bounds) and $P - E$ using three global evaporation products over the Lake Eyre basin. Below: varying degree correlation measure ( $r_l$ ) with time and degree (from 25 to 40) using the three global evaporation products. The average $r_l$ is shown as a time series (black line). The degree correlation measure can range from -1 to 1 as shown in the color scale on the secondary axis. . . . .	68
3.8	Top: average degree correlation statistics per study region and evaporation product. Bottom: GRACE data were shifted by two months to match the phase with $P - E$ anomalies. The boxplots show the first, second (median) and third quartiles. Outliers, defined as data outside the 1.5 inter-quartile range (IQR) whiskers below or above the first and third quartiles are shown as circles. . . . .	71
4.1	Conceptual model of TSEB including the network of resistance (adopted from <i>Norman et al. (1995)</i> and <i>Colaizzi et al. (2012)</i> ) . . . . .	86
4.2	Flow diagram of CABLE model showing the importance of the stability parameter iterative calculations. A more complete diagram can be found in <i>Kowalczyk et al. (2006)</i> . . . . .	91

4.3	Seasonality of daily average temperature over the 2011-2015 WRF simulation period within the interior (3 km resolution) domain consisting of 1522 x 1210 grid points covering the Arabian Peninsula and neighboring regions. . . . .	94
4.4	10 AM (local time) snapshots of WRF variables relevant for CABLE extracted from the location of the Tawdeehiya Farm for the period 2013 - 2015. . . . .	95
4.5	Average annual precipitation (2011 - 2015) over the WRF interior (3 km resolution) domain. . . . .	97
4.6	Daily precipitation at the Tawdeehiya site for the years 2013 - 2015. .	98
4.7	Default vegetation and soil classification types in CABLE. . . . .	98
4.8	Default global monthly LAI dataset based on MODIS data (averaged from 2002 to 2009) included in CABLE auxiliary files (a: July, b: October, c: January, d: April). . . . .	99
4.9	Left: Landsat 8 based land cover classification at Tawdeehiya Farm for the year 2015. Right: The 41 pivots numbered approximately in SW-NE stripes. The smaller pivot in the eastern part of the farm was omitted for this study, as well as the pivot located in the SE corner and the rectangular alfalfa field in the eastern part of the farm. Two pivots that did not grow vegetation for the year 2015 were omitted as well (their numbers, 15 and 17, were skipped). . . . .	101
4.10	Two scenes of Landsat 8 based LAI data from ( <i>Houborg and McCabe, 2018</i> ). The proximity of the farm to the eastern swath edge for Landsat path/row 164/43 is noticeable in the Western side of the farm . . . .	102
4.11	Landsat reflectance values upon which surface (e.g. albedo, left) and vegetation parameters (e.g. LAI, right) are based upon are affected when there is significant cloud cover. The number of observations is therefore reduced, affecting the number of evaporation retrievals as well.	103
4.12	Coverage of Landsat scenes (cloud-covered scenes excluded) per pivot and crop type for four three-month periods starting from September 2014. Pivots located on the west side of the Farm receive less coverage due to the swath edge of Landsat 8 . . . . .	104
4.13	Calculated “effective monthly irrigation” based on farm data for pivot number 13 (see figure 4.9) . . . . .	105

4.14	Soil moisture time series (six layers) for the constant (a-b) and “effective irrigation” (c-d) irrigation schemes for the fifth consecutive run of the model (year 2014). The model was initialized with values of 0.3 (a, c) and 0.1 (b, d) $m^3.m^{-3}$ for all layers. . . . .	106
4.15	Daily water budget for the WRF-rainfall model run from March 2014 to August 2015 after the spinup period, using two different spinup scenarios: constant irrigation (a) and EMI (b). $\Delta WS$ is the change in water storage from the soil, while $\Delta WSC$ is the change in water storage from the canopy layer. . . . .	107
4.16	Estimated total evaporated water in $m^3$ from TSEB and CABLE without prescribed irrigation. . . . .	108
4.17	High-resolution daily evaporation maps obtained by TSEB. . . . .	111
4.18	High-resolution daily evaporation maps obtained by CABLE using only WRF rainfall as “irrigation”. . . . .	112
5.1	Diagram showing the general methodology for estimating irrigation rates (left) and optimization periods used in this study (right). The grayed area represents the study period. . . . .	115
5.2	Results of six tuning experiments for two selected fields (A:13, B: 25)	120
5.3	Results of six tuning experiments for two selected fields (C: 25, D: 30).	121
5.4	Results of six tuning experiments for two selected fields (A:13, B:25) using a higher initial guess . . . . .	122
5.5	Results from applying one set of SPSA parameters to all 41 fields for the first optimization period using synthetic data as observations. Circles in red are out of scale (larger than $18 mm.day^{-1}$ ). See text for additional details. . . . .	124
5.6	Testing different initial guess values (A: 6.3; B: 7.56; C: 9.24 and D: $10.5 mm.day^{-1}$ ). Optimization during the first three-month period (2014MAM) using synthetic data as observations. . . . .	125
5.7	Comparison of optimized irrigation values (right) against farm data (average of each corresponding three-month period for each field). . .	127

5.8 Comparison of optimized irrigation values against “observed” values (average of each corresponding three-month optimization period for each field). A: Data for all 41 fields during the four study periods. B: Data for all 41 fields during the four study periods, but omitting observed values with a large monthly variability ( $>5 \text{ mm.day}^{-1}$ ). C: Data for the East fields (24-43) during the four study periods. D: Same as C, but omitting observed values with high variability. . . . . 128

5.9 Comparison of optimized values and farm data (averaged over each three-month period) for each of the four study periods. E: East fields (24-43); W: West fields (1-23); R: values with more than  $5 \text{ mm.day}^{-1}$  difference between the three values considered for the average effective irrigation. . . . . 130

5.10 High-resolution daily evaporation maps obtained using the optimized model. . . . . 131

5.11 Left: scatter plot of average evaporation for each field (34 days) comparing RS-E data (i.e. from TSEB, see section 4.2.1) to the LSM results. Right: Comparisons for individual days shown in figure 5.10 . . 132

5.12 Top: LAI dynamics and observed evaporation (TSEB) response for fields 33, 34, 35, 38, 41 and 42. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (six fields). Bottom (right): location of the relevant fields. . . . . 133

5.13 Evaporation partitioning from model and relationship to LAI (field number 38). . . . . 133

5.14 Top: LAI dynamics and observed evaporation (TSEB) response for fields 36, 37, 39, and 40. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (four fields). Bottom (right): location of the relevant fields. . . . . 134

5.15 Evaporation partitioning from model and relationship to LAI (field number 39). . . . . 135

5.16 Top: LAI dynamics and observed evaporation (TSEB) response for fields 22, 24, 25, 26, and 43. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (five fields). Bottom (right): location of the relevant fields. . . . . 135

5.17	Evaporation partitioning from model and relationship to LAI (field number 26). . . . .	136
5.18	Top: LAI dynamics and observed evaporation (TSEB) response for fields 3, 4, 5, 9, and 10. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (five fields). Bottom (right): location of the relevant fields. . . . .	137
5.19	Evaporation partitioning from model and relationship to LAI (field number 5). . . . .	137
5.20	Top: LAI dynamics and observed evaporation (TSEB) response for fields 2, 6, 7, and 8. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (four fields). Bottom (right): location of the relevant fields. . . . .	138
5.21	Evaporation partitioning from model and relationship to LAI (field number 7) and change in water storage (WS). . . . .	138
5.22	Total annual groundwater abstraction (millions of cubic meters) estimated using the optimization methodology for the entire Tawdeehiya farm, and partitioning of water fluxes. . . . .	139
5.23	Comparison of total daily water evaporation for 17 fields (27-43) using 1) only meteorological rainfall (section 4.7) and 2) optimized irrigation values. . . . .	140

## LIST OF TABLES

3.1	Description of the satellite products used in this study. The temporal resolution is daily except for MOD16 (8-daily) and GRACE (monthly). The original MOD16 product is available at 1 <i>km</i> resolution in the sinusoidal projection. In this study, the product was reprojected onto a 0.05° regular grid using the MODIS Reprojection Tool (MRT). . . .	50
5.1	Statistical measures obtained for the four study periods, grouped by A: fields with irrigation values <18 <i>mm.day</i> <sup>-1</sup> , and B: same as A but excluding values with range >5 <i>mm.day</i> <sup>-1</sup> . The analysis was split for half the number of fields in the West and East sides of the Tawdeehiya farm. . . . .	129

## **Chapter 1**

### **Introduction**

---

The availability and sustainability of water resources as well as food security is a global concern. In Saudi Arabia, irrigated agriculture has increased over the last few decades, becoming a significant user of groundwater in the Kingdom. Monitoring this use is therefore critical for managing the limited water resources. This Chapter introduces the potential that remote sensing and land surface modeling have for monitoring irrigated agricultural water use. The objectives of this dissertation are defined here as well.

---

## 1.1 Remote sensing and numerical modeling opportunities for water management

Concerns for food and water security are growing worldwide due to increased population and the effects that climate change will likely have on the Earth's global water and energy cycles. In Saudi Arabia, population increase as well as life style changes continue to increase the Kingdom's water demand at a high rate (*Al-Salamah et al.*, 2011; *DeNicola et al.*, 2015; *Odhiambo*, 2017). Even though water desalination has recently become the main water source for urban water demand (particularly in coastal metropolitan cities such as Jeddah), remote agricultural areas depend heavily on groundwater aquifers (*Missimer et al.*, 2012; *Lopez et al.*, 2014; *Chowdhury and Al-Zahrani*, 2015). While climate models forecast insignificant changes in precipitation over the Arabian Peninsula, surface air temperature is expected to increase and affect evaporation rates and recharge processes (*Evans*, 2009; *AlSarmi and Washington*, 2011; *Kirtman et al.*, 2013).

Monitoring the water cycle has thus become an important task for managing the limited water resources in the region and the world. Over remote areas, traditional ground-based measurements are difficult to manage and expensive to operate. Remote sensing has become a valuable tool to retrieve information across a wide range of land surface hydrological components (*McCabe et al.*, 2017), representing the only source of information for remote areas that otherwise lack data. Space-based observations (i.e. remote sensing via satellites) offer a unique opportunity to characterize terrestrial processes in space and time, from local to global scales.

However, the development of remote sensing sensors, platforms and algorithms has been met with challenges. For example, deriving evaporation from remote sensing is particularly challenging since it can only be measured indirectly, since heat fluxes do not absorb nor emit electromagnetic signals directly (*Jiménez et al.*, 2011). Consequently, the evaluation of remote sensing-based evaporation relies on alternative

techniques (*Kalma et al.*, 2008) (Chapter 3). Satellite-based gravity measurements (*Swenson and Wahr*, 2002; *Rodell et al.*, 2009; *Famiglietti et al.*, 2011; *Voss et al.*, 2013) offer a unique insight into water movement across the Earth's surface, but have considerable limitations and particular challenges in terms of scale and accuracy (section 2.2). Furthermore, the development of satellites and sensors has usually been done separately for different components of the water and energy cycles, and thus combining them is also a challenge (*McCabe et al.*, 2017).

Land Surface Models (LSM) have also become powerful tools for studying the natural variations in the hydrological fluxes and states within the Earth's land surface (*Koster et al.*, 2000). Over the last three decades, LSMs have improved their representation of the land-atmosphere interactions from simple bucket-type schemes to complex characterizations that include biogeochemical processes and vegetation-climate feedbacks (*Pitman*, 2003; *Zhang et al.*, 2011). However, simulations over regions where there is a significant impact of anthropogenic activities (e.g. excessive groundwater pumping or intense irrigation) have been shown to deviate significantly from observations (*Voss et al.*, 2013; *Anderson et al.*, 2015; *Zeng et al.*, 2016). Therefore, there is a need for improving the physical representation of land surface models in regions where there are significant human contributions to changes in the water balance.

Advances in remote sensing techniques also contribute to the improvement of land surface model applications, enabling high-resolution simulations that can better incorporate changes in the land surface, such as agricultural development (chapter 4). Therefore, exploring novel techniques to combine the advantages of remote sensing and land surface models is of great value in local to global scale water resources (*Chen et al.*, 2011; *Choi*, 2013; *Sridhar*, 2013; *Renzullo et al.*, 2014; *Vahmani and Hogue*, 2014).



Figure 1.1: Left: aerial view of a center-pivot irrigation field. Right: Close-up showing a section of the pivot arm spraying water on the field.

## 1.2 Agricultural development and groundwater resources in Saudi Arabia

As has been witnessed over the last few decades in Saudi Arabia, growing food in an extreme arid environment is possible (*Hussain and Al-Jaloud, 1995; Qadir et al., 2003; Grindle et al., 2015; Chowdhury and Al-Zahrani, 2015; Fiaz et al., 2016*), but it requires the intensive extraction of groundwater and associated energy costs to drive the diesel pumps transporting the water from deep aquifers. In Saudi Arabia, center-pivot irrigation is the most common type of farming technique, and is used to grow cereals, grains, fruits and vegetables (*Abo-Ghobar, 1992; Al-Yemeni and Grace, 1995; Al-Rumikhani, 2002; Madugundu et al., 2017*). In these systems, a groundwater well is located in the center of the pivot, supplying hundreds of gallons per minute to a rotating arm consisting of several nozzles with varying sizes, ideally providing a uniform distribution of water to the circular field (figure 1.1). These type of systems are widespread in other arid and semi-arid regions of the world, because they can reduce labor requirements, and are more flexible in terms of management, although at the expense of higher energy requirements *Steiner et al. (1983); Izquier et al. (2015)*.

Figure 1.2 shows two regions in Saudi Arabia where the widespread use of center-

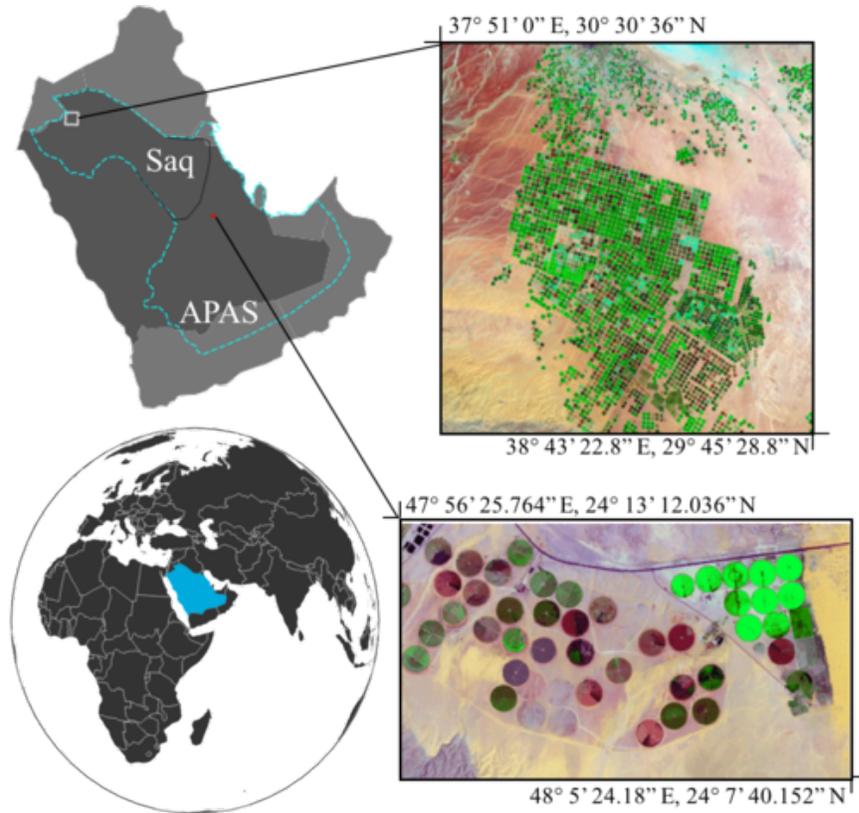


Figure 1.2: Location of two agricultural regions with center-pivot irrigation in Saudi Arabia: Wadi As-Sirhan in the northwest, and the Tawdeehiya farm near Riyadh. Also shown is the outline of the Arabian Peninsula Aquifer System (APAS) and the Saq aquifer system.

pivot irrigation is visible from space: Wadi Al-Sharhan in the Al-Jowf region (northwest), and the Tawdeehiya farm in the Al-Kharj region, near Riyadh. Most of the eastern and northwestern regions of Saudi Arabia lie within the Arabian Peninsula Aquifer System (APAS), with groundwater flow directed predominantly towards the east (*Sultan et al.*, 2008, 2014). A group of aquifers in different geologic formations (at varying depths) located on the northwest part of the APAS constitute the Saq aquifer system, and contains outcrops in the westernmost part of the system. The majority of aquifers within the APAS are located hundreds of meters below the surface (*Sultan et al.*, 2014), and groundwater elevation data from monitoring wells within this region show a historic depletion trend (*Al-Salamah et al.*, 2011).

### 1.3 Objectives

In an effort to develop a strategy towards a sustainable balance between food security and the long-term availability of water resources, a key task is to monitor the growth and impact that these agricultural fields have on the limited groundwater resources. The overall objective of this work is to study the large-scale and field-scale impacts of agricultural development in Saudi Arabia on its water resources. The first part of the study (chapters 2 and 3) is dedicated to the large-scale retrieval of hydrologic fluxes (i.e. water storage changes, evaporation and precipitation) using remote sensing. Here we evaluated different approaches for large-scale retrievals of water storage variations using gravity data. We compared depletion trends obtained by applying two different approaches over the Arabian Peninsula and nearby regions (e.g. within the Middle East - North Africa region). We then explored whether this information (gravity-based water storage retrievals) could be used to evaluate a range of global evaporation products (*Miralles et al.*, 2016; *McCabe et al.*, 2017) (chapter 3). This was done by assessing how well these independently developed products relate in terms of the expected water budget closure (*McCabe et al.*, 2008; *Sheffield et al.*, 2009). Evaluating how well these large-scale evaporation retrievals are able to capture regional scale flux variability is a key task in monitoring the impact of agricultural crop water use.

The second part of the study aimed towards developing a methodology to estimate groundwater abstraction rates used in irrigated agriculture. To do this, we first used high-resolution (30 m) remote sensing data and meteorological data from a numerical weather prediction model to drive a land surface model (LSM). The ability of this model to capture the impact from agricultural activities was examined by comparing the evaporative fluxes with a thermal-based remote sensing evaporation (RS-E) model (chapter 4). Next, we examined whether the RS-E model could be used to constrain the LSM using an optimization algorithm in order to infer the missing irrigation com-

ponent, and hence infer the groundwater abstraction rates (chapter 5). Finally, the developed framework was tested on a relatively small number of agricultural fields (Tawdeehiya farm; figure 1.2, bottom) in order to identify potential challenges and limitations of the methodology. Future work will be focused towards the retrieval of groundwater abstraction rates over a number of regions with high density of agricultural fields (e.g. figure 1.2, top). Other future efforts will include providing several improvements to the methodology (e.g. in terms of increasing the frequency of observations and improving key model parameters), as well as developing a user interface and other features (e.g. data visualization) that will be integrated as an operational water management tool.

## Chapter 2

### Estimating large-scale water storage variations in the Arabian Peninsula from gravity measurements

---

This chapter introduces the Gravity Recovery and Climate Experiment (GRACE), a modern international effort to map gravity anomalies in the Earth's surface. GRACE is relevant to hydrologic studies because of the link between water movement in the Earth's surface and the gravity changes that GRACE can detect. The chapter describes the theory behind this link and the techniques required to infer the terrestrial water storage variations (TWSA) over large regions of the land surface. These techniques are then applied to estimate recent long-term ( $>10$  years) and large-scale depletion rates over the Arabian Peninsula, which are suggested to be linked to increased agricultural activity in the region.

---

## 2.1 Introduction: The Gravity Recovery and Climate Experiment (GRACE)

In March 17, 2002, the National Aeronautics and Space Administration (NASA) launched a pair of co-orbiting satellites into a Low Earth Orbit ( $\sim 480$  km) as part of the Gravity Recovery and Climate Experiment (GRACE) mission (*Tapley et al.*, 2004a). The purpose of the mission is to routinely capture the Earth’s gravity field. Despite its planned lifetime of 5 years, the mission continues as of today (although with a reduced operation due to battery management during the past few years). In terms of its operation, when the leading satellite passes over a denser portion of the Earth’s surface, its acceleration and relative position to the other satellite are recorded. The relationship between these variables can then be used to determine the Earth’s gravity field. In a sense, it acts “as a scale in the sky” (*Famiglietti and Rodell*, 2013).

Water movement on the Earth’s surface is one of the largest contributing factors driving changes in the gravity field. Other factors include mass redistribution in the atmosphere, gravitational tide in the solid Earth and oceans, and post-glacial rebound (*Wahr et al.*, 1998). Modeling and removing the effects of these components is necessary to retrieve the water storage variations from the cumulative gravity response. Section 2.2 describes how the solution to the gravity field is represented, its relationship to the water cycle, and the different solutions that have been developed to extract this unique water storage signal.

GRACE’s ability to detect monthly vertically integrated terrestrial water storage changes (TWSC) over large regions has been well-documented in the literature. The studies by *Swenson et al.* (2008), *Rodell et al.* (2009), *Famiglietti et al.* (2011), *Sun* (2013), *Jin and Feng* (2013) and *Voss et al.* (2013) show the potential of GRACE for groundwater monitoring across many different spatial and temporal scales. Drought monitoring using GRACE data has also been a topic of recent research (*Houborg et al.*,

2012; *Thomas et al.*, 2014), as well as numerous studies involving water balances over large regions (e.g. *Swenson and Wahr* (2009); *Ahmed et al.* (2011)), along with efforts to use retrievals to improve land surface models *Sun et al.* (2012); *Lo et al.* (2010). GRACE has also been used to evaluate spatial patterns of precipitation through a water balance approach (*Swenson*, 2010), as well as obtaining regionally averaged evaporation estimates over a large river basin (*Rodell et al.*, 2004a).

One of the principal limitations of GRACE is the trade-off between resolution and accuracy. GRACE can only provide regionally averaged estimates of TWSC over large scales (e.g. resolution of a few hundred kilometers), and is subject to contamination from random errors as well as correlated errors. Post-processing of GRACE data involves the removal of both types of errors by using a de-stripping filter (*Swenson and Wahr*, 2006) together with smoothing using a Gaussian filter. Use of these filters also modifies the true signal, usually by reducing it (*Landerer and Swenson*, 2012), and leakage contamination from neighboring areas may also occur. To account for this, *Landerer and Swenson* (2012) suggest the use of a gain factor (section 2.2.5). In the next section, processing of GRACE coefficients will be explained in more detail.

## 2.2 GRACE data processing

### 2.2.1 Linking water movement to the Earth's gravity field <sup>1</sup>

The gravity field is usually described in terms of the geoid: an equipotential surface that is defined as corresponding to the mean sea level over the oceans (*Swenson and Wahr*, 2002). The geoid is usually approximated as a linear combination of spherical harmonics, given that these represent solutions to the Laplace equation that describes the relationship between the gravitational potential and the geoid. The approximation

---

<sup>1</sup>This section is an edited version of the description of GRACE data (section 2.1) in: Oliver Lopez, Rasmus Houborg, and Matthew Francis McCabe: Evaluating the hydrological consistency of evaporation products using satellite-based gravity and rainfall data, *Hydrol. Earth Syst. Sci.*, 21, 323-343, 2017

is of the form:

$$f(\theta, \phi) \approx \sum_{l=0}^{l_{max}} \sum_{m=0}^l \tilde{P}_{lm}(\cos \theta) \{C_{lm} \cos m\phi + S_{lm} \sin m\phi\} \quad (2.1)$$

where  $\tilde{P}_{lm}(\cos \theta)$  are the normalized associated Legendre functions,  $\theta$  corresponds to the colatitude (the complementary angle to the latitude),  $\phi$  to longitude,  $C_{lm}$  and  $S_{lm}$  are the spherical harmonic (SH) coefficients of degree  $l$  and order  $m$ , and  $l_{max}$  is the truncation degree. The total number of coefficients is given by  $((l_{max}+1)^2+l_{max}+1)/2$ , while the resolution (the scale of the smallest feature of the gravity field that can be resolved using  $l_{max}$  coefficients) is approximately  $\pi a/l_{max}$  (where  $a$  is the Earth's radius). *Wahr et al.* (1998) showed how the changes in surface mass density (assuming these occur in a thin layer of the Earth's surface) can be inverted from temporal changes in the geoid coefficients:

$$\Delta\sigma(\theta, \phi) = \frac{a\rho_{ave}}{3} \sum_{l=0}^{\infty} \sum_{m=0}^l \frac{(2l+1)}{1+k_l} \tilde{P}_{lm}(\cos \theta) \{\Delta C_{lm} \cos m\phi + \Delta S_{lm} \sin m\phi\} \quad (2.2)$$

where  $\rho_{ave}$  is the average density of the Earth ( $5,517\text{kg}/\text{m}^3$ ), and  $k_l$  are the load Love numbers (related to the additional geoid contribution from the surface mass loading of the Earth itself (*Wahr et al.*, 1998)). The GRACE level 2 products consist of a set of SH coefficients ( $C_{lm}$  and  $S_{lm}$ ) provided on a monthly basis.

### 2.2.2 The averaging kernel

While equation (2.2) represents a point value of the surface mass anomaly, it would require an infinitely large number of spectral coefficients to describe the point value. To get a spatial average over a region, the averaging must instead be done in spherical harmonics. To do this, an averaging kernel is formed as an approximation of the exact basin shape, which has a value of 1 inside the basin and 0 outside the basin. In SH

coefficients, the averaging kernel  $\vartheta$  is obtained as:

$$\begin{pmatrix} \vartheta_{lm}^c \\ \vartheta_{lm}^s \end{pmatrix} = \int \vartheta(\theta, \phi) \tilde{P}_{lm} \begin{pmatrix} \cos m\phi \\ \sin m\phi \end{pmatrix} d\Omega \quad (2.3)$$

*Swenson and Wahr* (2002) showed that the degree amplitude spectrum of the basin coefficients is concentrated over smaller degrees as the basin size increases. They also showed that satellite measurement errors increase rapidly with increasing degree, so that errors from large degrees (corresponding to small scale lengths) seriously degrade the solution. Therefore, reducing the number of degrees in the solution will eliminate the contribution from high-degree errors. However, as the averaging kernel is constructed with a smaller number of maximum degree  $l$ , the basin function will be less accurate and contain contributions from regions outside of the basin. This type of additional error is called leakage contamination. Therefore, there is a trade-off between high-degree satellite measurement errors and leakage contamination.

### 2.2.3 The Gaussian smoothing kernel

In order to obtain a regional TWS signal with a balance between measurement errors and leakage errors, *Swenson and Wahr* (2002) suggested using a Gaussian smoothing kernel, which is formed by convolving the basin function with a Gaussian filter. In SH coefficients, the Gaussian smoothing kernel can be obtained as follows:

$$\begin{pmatrix} W_{lm}^c \\ W_{lm}^s \end{pmatrix} = \exp\left(-\frac{(lr/a)^2}{4\ln(2)}\right) \begin{pmatrix} \vartheta_{lm}^c \\ \vartheta_{lm}^s \end{pmatrix} \quad (2.4)$$

where  $r$  is the Gaussian half-width radius that determines the distance at which the kernel smoothly decreases to zero (from the basin boundary with a value of 1). Changing  $r$  allows one to control the relative contribution of measurement and leakage

errors. The domain-averaged TWS signal can be obtained as:

$$\tilde{\Delta}\sigma_{region} = \sum_{l,m} \frac{K_l}{\Omega_{region}} (W_{lm}^c \Delta C_{lm} + W_{lm}^s \Delta S_{lm}) \quad (2.5)$$

where

$$K_l = \frac{a\rho_e}{3} \frac{2l+1}{1+k_l} \quad (2.6)$$

#### 2.2.4 Removing correlated errors in the GRACE SH coefficients

*Swenson and Wahr* (2006) showed that satellite measurement errors in GRACE data are dominated by the presence of physically unrealistic stripes (figure 2.1) that are oriented meridionally. In the spectral domain, they discovered that coefficients of a particular order with same parity in degree were correlated, and proposed to remove this correlation by fitting a polynomial to the coefficients in a moving window. *Duan et al.* (2009) discussed a number of similar de-stripping filtering approaches and proposed an improved version of the *Swenson and Wahr* (2006) filter in which the moving window is adjusted based on patterns of the standard deviations of the SH coefficients.

Unfortunately, the de-correlation filter also has an adverse effect on the hydrological signal, adding a bias to the desired hydrological response (generally reducing the true signal). To correct for this, a last step in the processing of GRACE coefficients must be added, which aims to restore the signal to its original amplitude. Since there is no criterion to assess the impact of the filter on the true TWS signal, the effect has to be inferred through the use of a land surface model.

The TWS output of such a model is converted to SH coefficients and processed similarly to the GRACE coefficients, so that a scaling factor may be inferred. The use of a land surface model presents a problem in the sense that there may be re-

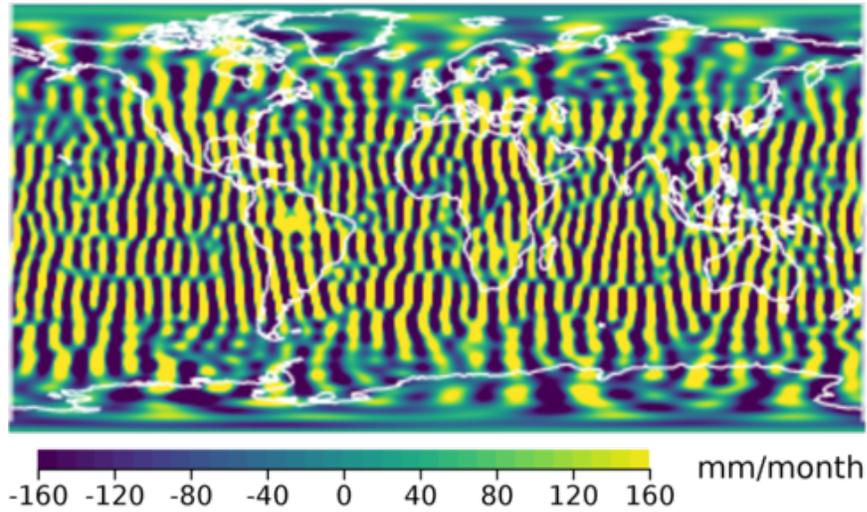


Figure 2.1: Typical striping pattern found in unfiltered GRACE coefficients (GRACE CSR product for 2010/06/01 - 2010/06/30, converted to water storage anomalies using mean coefficients computed from 2002-2016)

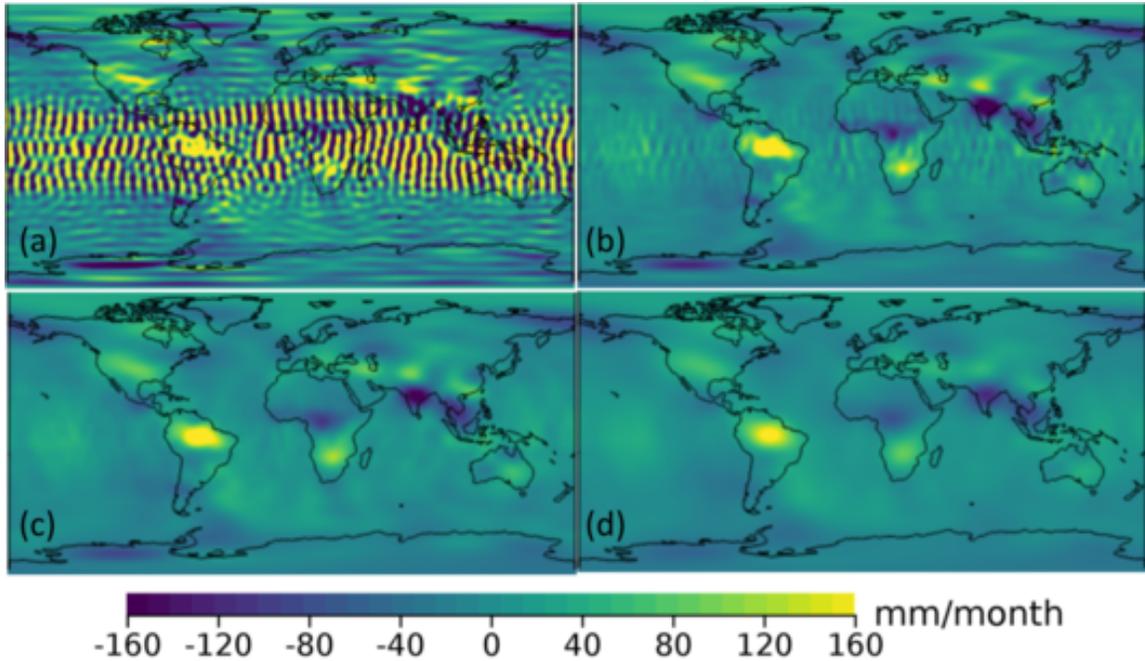


Figure 2.2: GRACE CSR product for 2010/06/01 - 2010/06/30 filtered with the de-striping filter and Gaussian smoothing with radius a) 0 *km* (no smoothing), b) 350 *km*, c) 500 *km* and d) 750 *km*

gions where the hydrological signal is misrepresented, or have external influences such as human related withdrawals or water management that ultimately have an unaccounted effect on the water storage changes. The latter is of particular concern for applications in groundwater abstractions for agriculture, since land surface models do not routinely account for such water uses.

### 2.2.5 Scaling the GRACE signal

*Long et al.* (2015) compared three methods to restore the GRACE signal. The first one (*Landerer and Swenson, 2012*) uses TWS output from a land surface model to calculate a gain factor that accounts for leakage and bias correction integrally. The gain factor  $k$  is calculated by minimizing the square of the residual between the unfiltered and filtered TWS signal:

$$M = \sum_{i=1}^T (\bar{S}_{m,i} - k\hat{S}_{m,i})^2 \quad (2.7)$$

While the gain factor was assumed to not depend explicitly on the land surface model (*Landerer and Swenson, 2012*), *Long et al.* (2015) revealed an impact on the scaling factor from the choice of land surface model, particularly in arid regions. The coefficients of variations (defined as the ratio of standard deviation and the mean) for scaling factors in arid regions were larger than 0.8 (0.3 for normal regions).

Most of the land scaling factors were larger than 1 (positive skewness), which indicates that the signal is reduced after filtration. However, there were some regions, mostly arid, where the scaling factor was less than 1, due to leakage (contamination) from other regions (e.g. dry areas adjacent to regions with higher signals). In some regions there were very large scaling factors, indicating local-scale behaviors differing from regional-scale behaviors, such as pumping from a small groundwater system. Negative values indicate out of phase behavior with surrounding areas.

The second approach (*Klees et al. (2007)*) is referred to as the additive correction approach, where an LSM is used to calculate the bias correction and leakage explicitly, which are then added to the signal:

$$S_B = \frac{1}{R_0} \int_R S_0(h - \hat{h})d\Omega \quad (2.8)$$

$$S_L = \frac{1}{R_0} \int_{\Omega-R} S_{leak}\hat{h}d\Omega \quad (2.9)$$

where the first integral calculates the difference between the unfiltered and filtered signal inside the basin, and the second calculates the filtered signal outside the basin (which should be zero with the exact basin function  $h$ , but is non-zero due to the Gibbs phenomenon, i.e. ringing due to discretization of the basin). The third approach is called the multiplicative factor and only accounts for leakage. It assumes a uniform distribution of TWS changes within a basin. The factor is calculated as:

$$k_m = \left( \frac{1}{R_0} \int_R \hat{h}d\Omega \right)^{-1} \quad (2.10)$$

$$\bar{S}_0 = \left( \hat{S}_0 - S_L \right) \cdot k_m \quad (2.11)$$

In order to enable end-users to use GRACE data without the need for expertise in spherical harmonic processing, *Landerer and Swenson (2012)*, prepared a gridded (1 degree resolution) product based on the standard processing of GRACE coefficients (Gaussian filtering, de-stripping filtering and scaling) using the GRACE level 2 solutions from three research centers: the Center for Space Research (CSR) at the University of Texas, Austin, the Jet Propulsion Laboratory (JPL) and the GFZ German Research Center for Geosciences (GeoforschungsZentrum). Although useful for easily monitoring TWS over large regions, it is best to retrieve the signal from averaging kernels better suited for specific regions than the 1-degree cells, particularly in regions vulnerable to leakage contamination such as in arid climates. Furthermore, in

order to correctly incorporate GRACE water storage changes consistently with other remote sensing data, the effects of the filters and scaling factors used in GRACE data must be taken into consideration (this will be covered in section 3.4).

### 2.2.6 Other GRACE processing approaches

Besides this brief overview of the general methodology to derive TWS from GRACE satellites, there are some variants that are worth mentioning. While the Gaussian smoothing kernel is easy to implement, *Swenson and Wahr* (2002) showed that another type of filter can be obtained by optimizing the averaging kernel with respect to the total error. Implementing this filter is, however, complicated by the fact that it requires *a priori* information on the errors. Another method, based on Lagrange multipliers, is to create an averaging kernel that minimizes the leakage error subject to a constraint on the value of measurement errors. Although more complicated to implement than the Gaussian filter, it was shown that it decreased the leakage errors considerably more when averaging over larger regions. Another type of filter, called the optimal anisotropic filter (ANS) was described by *Klees et al.* (2008), which is based on minimizing the global mean of the mean-square error (MSE) between the unfiltered and filtered signals. However, it needs *a priori* information on the error variance.

*Frappart et al.* (2010) presented a method to optimally combine the information from three GRACE solutions (CSR, JPL and GFZ). The strategy assumes that these solutions present joint information on the true signal and that an independent component analysis (ICA) may be used to extract signals that are statistically independent, and that these correspond to the land hydrology, ocean mass anomalies, and the correlated errors or meridionally-oriented stripes. An ICA will therefore separate the land hydrology signal from the other two, and no further processing will be needed. This approach may be regarded as another type of filter to reveal the true geophysical

signal from GRACE data. In another study by *Frappart et al.* (2011), the ICA-based signals over different basins were compared to the GRACE CSR, JPL, and GFZ products, and showed better agreement with *in situ* data.

Alternative GRACE products available as global SH coefficients include the CNES/-GRGS (Centre National d'Etudes Spatiales/Groupe de Recherches de Géodésie Spatiale) product (*Lemoine et al.*, 2007; *Bruinsma et al.*, 2010), the GRACE DEOS Mass Transport (DMT) released by the Delft Institute of Earth Observation and Space Systems (*Liu et al.*, 2010), the use of mass concentration (mascon) solutions (*Rowlands et al.*, 2005, 2010), among others. These differ from the official GRACE products (CSR, JPL and GFZ) in the approaches used to recover the gravity field. Further description of the individual processing details of gravity field retrieval for these particular methods are beyond the scope of this work.

### **2.3 Recent water storage variations on the Arabian Peninsula**

In this section, we apply the methodology described above to retrieve the large-scale water storage variations over the Arabian Peninsula and compare them to changes in other areas of the world, including the larger Middle East region. Of particular interest within the Arabian Peninsula is the Saq aquifer (figure 1.2), an important aquifer that has played a key role in the agricultural development of the region, as observed by the dramatic changes in the land cover over the last few decades (figure 2.3). Groundwater measurements indicate excessive withdrawals from the Saq aquifer over the last 30 years (*Al-Salamah et al.*, 2011). In our analysis, water storage estimates were first retrieved from the 1-degree gridded product developed by *Landerer and Swenson* (2012) (section 2.2.5), and then with the full SH methodology in order to emphasize its importance within this region.

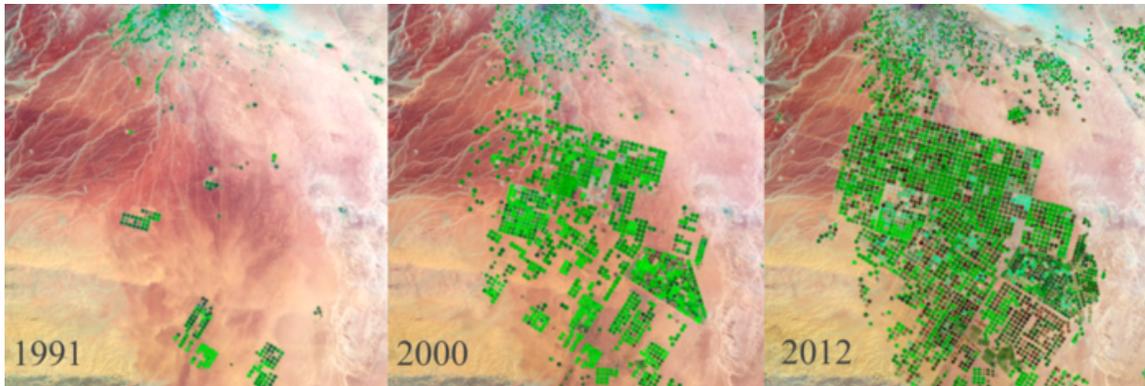


Figure 2.3: Agricultural expansion in Al-Jowf, Saudi Arabia as evidenced by Landsat satellites imagery.

### 2.3.1 GRACE level 3 (gridded) product inter-comparison

Figure 2.4 shows the terrestrial water storage anomalies (TWSA) averaged over the 1-degree cells from five defined regions within the Middle East - North Africa (MENA) domain. The Amazon River basin is included as a sixth region for comparison. The selected MENA regions are: (a) the Arabian Peninsula (figure 2.5, left), (b) the Saq aquifer within the Arabian Peninsula, (c) the Niger River basin, (d) the Nile River basin, and (e) a region encompassing the Tigris-Euphrates River basin and Western Iran (TEWI, as defined in *Voss et al.* (2013)).

While there is agreement between the three products in these regions, the long-term trends in water storage for the Arabian Peninsula, Saq aquifer and the TEWI regions do not agree with values found in the literature using averaging kernels. *Voss et al.* (2013) found a depletion trend of about  $27 \text{ mm.yr}^{-1}$  from 2003 to 2009, while the trend for that same period using the gridded product (figure 2.4-e) is only about  $9 \text{ mm.yr}^{-1}$ . A study using an averaging kernel for the Saq aquifer showed trends larger than  $10 \text{ mm.yr}^{-1}$  from 2003 to 2012 *Sultan et al.* (2014), with water storage anomalies significantly higher than those shown in figure 2.4-b. These large discrepancies highlight the need to carefully select the proper level of processing in some regions, particularly in smaller basins and/or within arid environments. However, in

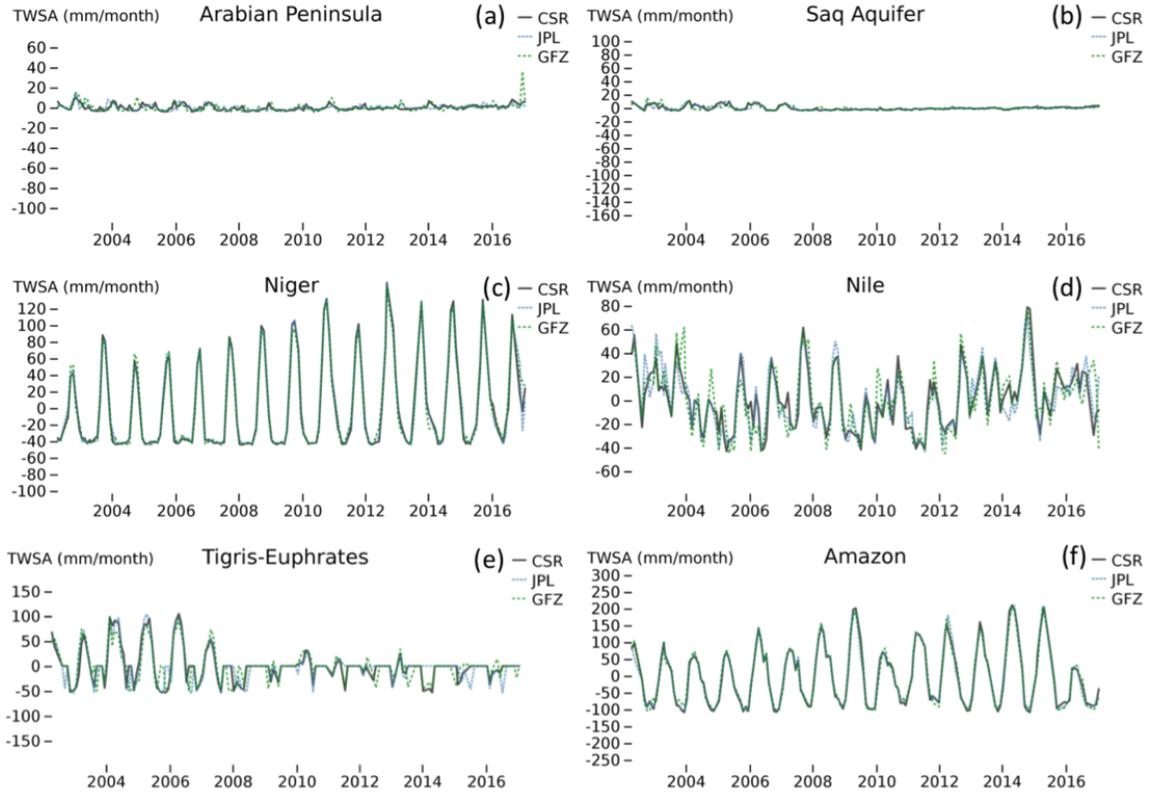


Figure 2.4: GRACE level 3 retrieval of terrestrial water storage anomalies (TWSA) averaged over  $1^\circ$  cells inside five regions within the MENA domain and the Amazon River basin. Note that the y-axis scales differ across each of the basins, illustrating the variability of TWSA across these regions.

other regions such as the Amazon river basin, the trends obtained by this product (about  $4 \text{ mm.yr}^{-1}$  from 2003 to 2014), are closer to the value ( $5 \text{ mm.yr}^{-1}$ ) found in the literature (*Chen et al.*, 2010; *Scanlon et al.*, 2018).

### 2.3.2 Water storage trends using SH coefficients

To correctly estimate water storage variations and trends in the Arabian Peninsula and the Saq aquifer within this region, two averaging kernels were prepared by means of spherical harmonic analysis (equation 2.3). This was achieved using a FORTRAN program developed by *Wang et al.* (2006) that takes a regular grid of values as input, in this case a rasterized mask of each region (with values of 1 inside and 0 outside). The program also outputs the spherical harmonic synthesis, i.e. the visual

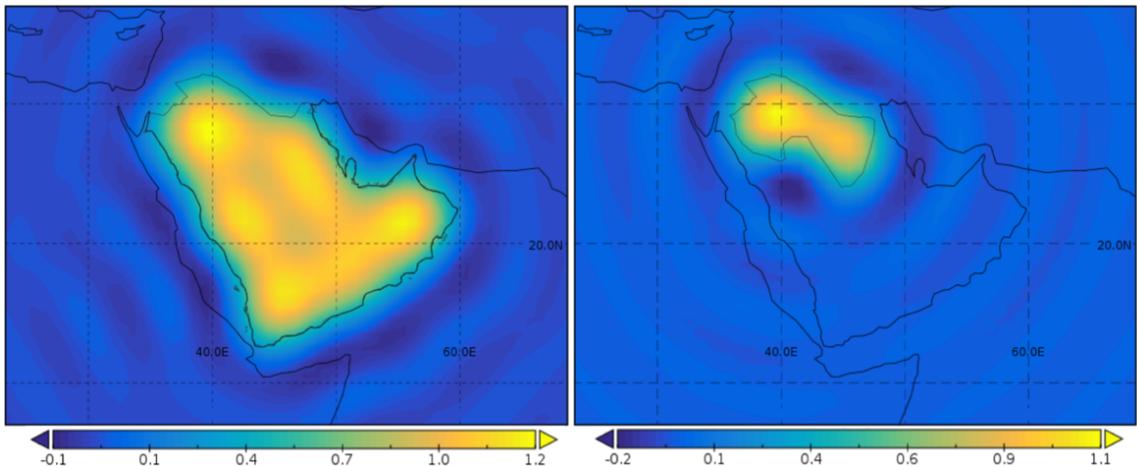


Figure 2.5: Averaging kernels obtained for the Arabian Peninsula (left) and Saq aquifer (right) regions.

representation of the coefficients back into the spatial domain. Figure 2.5 shows the resulting averaging kernels for the Arabian Peninsula (a) and the Saq aquifer (b). Values larger than 1 and even negative values are present in the SH synthesis due to the exclusion of higher degree terms (the cutoff was  $l_{max} = 60$ ). This effect is similar to the ringing phenomenon in Fourier series expansion.

The GRACE coefficients for the period 2002 - 2016 (a total of 155 months) were processed by (1) removing the mean from that period, (2) filtered using a de-stripping filter (section 2.2.4; *Swenson and Wahr (2006)*) and a Gaussian filter (section 2.2.3; *Swenson and Wahr (2002)*) with a 300 km radius. Then, the averaging kernels were used to obtain the TWSA values. Scaling factors used for the Amazon (1.11), Tigris-Euphrates (1.38), Niger (0.82) and Nile (0.89) River basins were obtained from *Long et al. (2015)*. These factors were obtained using the WaterGAP Global Hydrological Model (WGHM; *Alcamo et al. (2003)*), a land surface model that aims to incorporate anthropogenic sources of groundwater variability. For reference, the range of values obtained for these basins are (*Long et al., 2015*): 1.11 to 1.23 (Amazon), 1.22 to 1.60 (Tigris-Euphrates), 0.81 to 0.94 (Niger), and 0.55 to 1.21 (Nile). For the Saq aquifer and Arabian Peninsula, where no value is available in the *Long et al. (2015)* study,

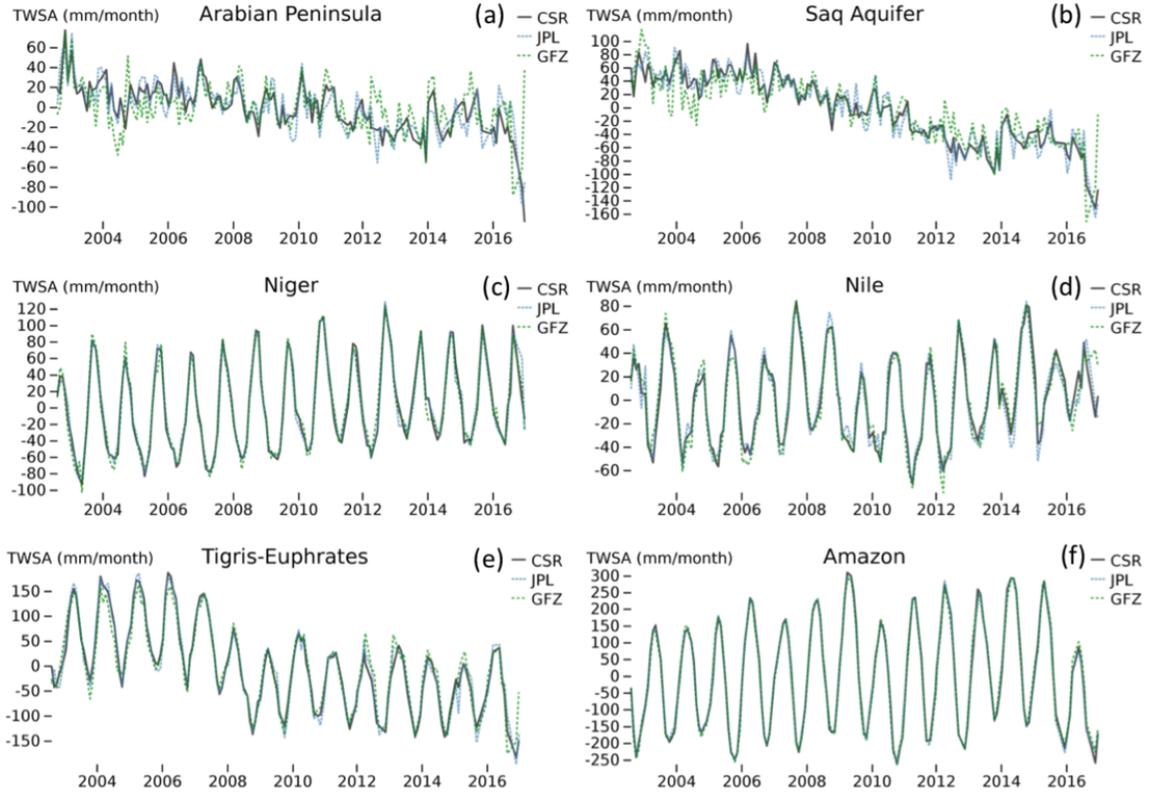


Figure 2.6: GRACE terrestrial water storage anomalies (TWSA) derived from spherical harmonic coefficients using averaging kernels of the five regions within the MENA domain and the Amazon River basin.

we used the result (1.6) obtained by *Sultan et al.* (2014), which is close to the one available for the Tigris-Euphrates basin.

Comparing the results for figure 2.4(a-b) and figure 2.6(a-b), it is evident that the gridded (level 3) product is inadequate to reveal the large depletion trends in the Arabian Peninsula (and Saq aquifer) region. Averaging the trends for the three GRACE level 2 products for the period 2006-01-01 to 2013-12-31, a trend of  $-4.44 \text{ mm.yr}^{-1}$  is obtained for the whole Arabian Peninsula and  $-16.63 \text{ mm.yr}^{-1}$  is obtained for the Saq aquifer alone. Furthermore, the differences in both figures for these regions are much larger even after considering that a scaling factor of 1.6 was applied to the data in figure 2.6. The trend obtained for the Saq aquifer for example (figure 2.6b), is closer to the value obtained by *Sultan et al.* (2014) ( $-13 \text{ mm.yr}^{-1}$ ; from 2002 to 2012)

than the value obtained with the gridded product (figure 2.4).

### **2.3.3 Recent increase in irrigated land cover within the Saq aquifer region**

As evidenced in figure 2.3, the Saq aquifer region has seen dramatic changes in terms of irrigated land expansion during the last few decades. To attempt to quantify this increase (at least in the large scale), the MODIS land cover type product (MCD12; *Friedl et al.* (2010)) was used to calculate the irrigated crop area within the Saq aquifer. This product consists of five land cover classification schemes and is available globally as MODIS sinusoidal grid tiles at 500 *m* resolution. Figure 2.7 shows the University of Maryland (UMD) classification data around the Saq aquifer region. After accumulating the information for pixels inside the Saq aquifer and five domains with visible agricultural impact, a rough large-scale estimation of the crop area was obtained (figure 2.8). Landsat data at 30 *m* resolution (e.g. figure 2.3) was also used to estimate the crop area within these regions. At least qualitatively, the increase in crop area within the Saq aquifer from 2004-2012 (more than 80% increase) corresponds to a decrease in the total water storage in the region.

## **2.4 Conclusions**

Despite the limitations and challenges of retrieving water storage variations from GRACE mission, particularly in arid regions (where contamination from nearby areas might affect the retrieval), large-scale trends can be accurately estimated within the Arabian Peninsula. However, a readily available gridded product aimed for end-users lacking expertise in processing GRACE spherical harmonic coefficients is not adequate for application for all global areas. Although this product might be useful in large regions with a clearly defined seasonality (e.g. Amazon, Nile and Niger basins), it was observed that this product misses some important trends in other smaller regions

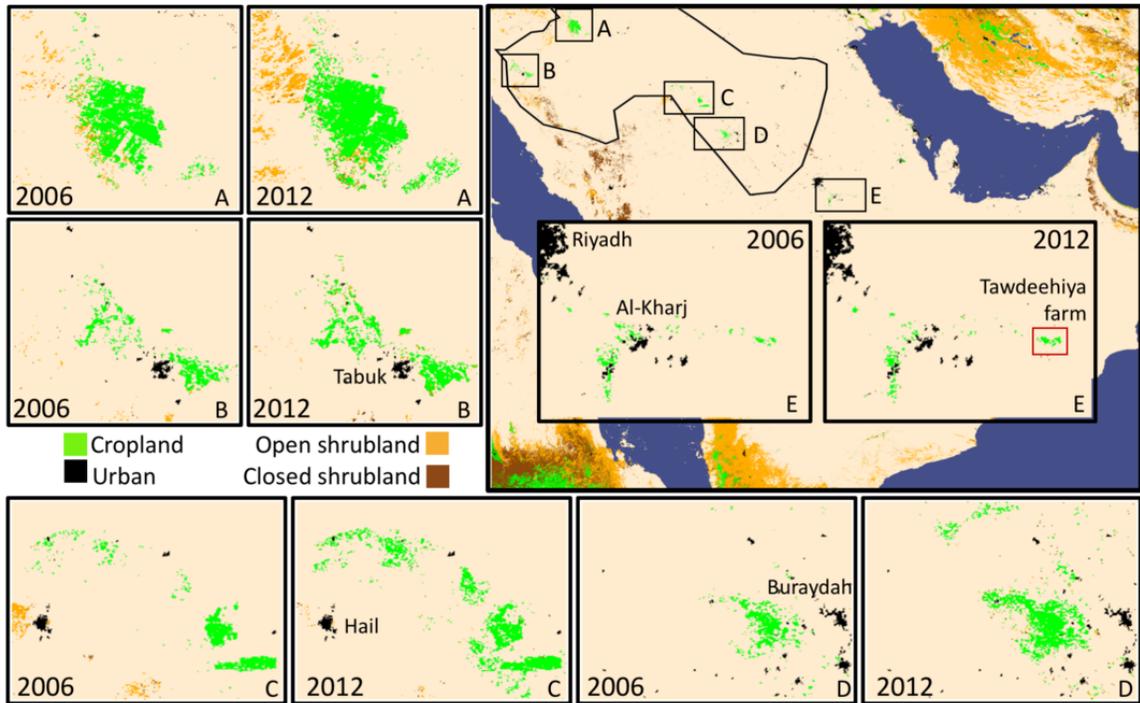


Figure 2.7: MODIS-derived land cover classification (500 m resolution) for five agricultural regions around the Saq aquifer. The cities of Tabuk, Hail, Buraydah and Riyadh are visible in black (urban classification). The largest agricultural region in terms of crop area is located on the North, in the region of Al-Jowf (A; see also figure 2.3). The Tawdeehiya farm, a region that will be explored in more detail in Chapters 4 and 5 is marked in red.

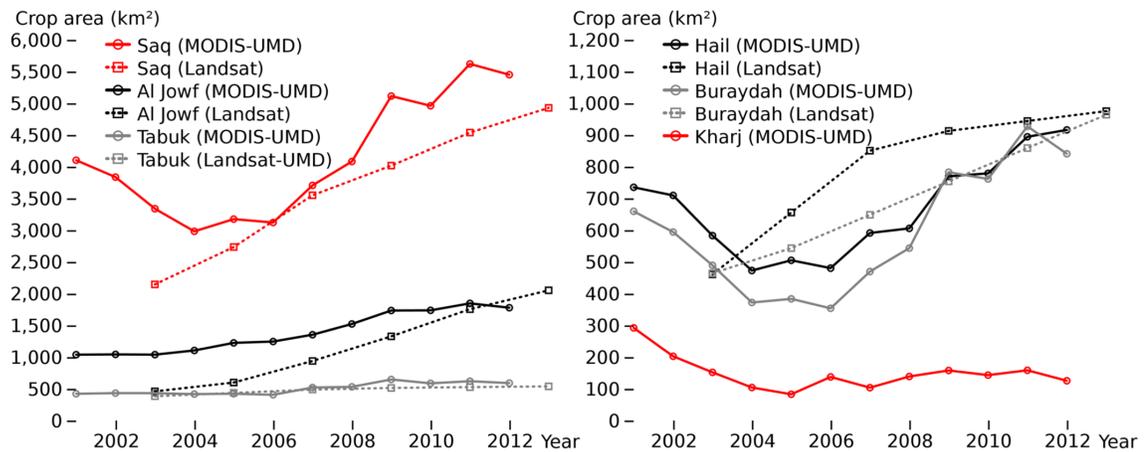


Figure 2.8: Crop area for the Saq aquifer and five agricultural regions estimated using the MODIS-derived product (500 m resolution) and Landsat data (30 m resolution). The increase within the Saq aquifer region is evident from 2004 – 2012, coinciding with the depletion period estimated using GRACE data.

that are subject to leakage contamination errors. The use of an averaging kernel and commonly used filters for error removal is recommended for retrieving TWSA values within the Middle East - North Africa (MENA) region. GRACE estimates showed a depletion trend within the Arabian Peninsula of about  $4 \text{ mm.yr}^{-1}$ , with a more significant depletion trend inside the Saq aquifer therein ( $\sim 16 \text{ mm.yr}^{-1}$ ). The depletion trend in the latter is most likely linked to the increase in agricultural activities within that region, as has been verified by Landsat and MODIS imagery. However, in order to definitively establish this link, other approaches aiming to more accurately determine the water used for irrigation are needed. This will be the topic of Chapters 4 and 5, in which a modeling framework combining high-resolution remote sensing data, numerical weather prediction data and a land surface model will be used to infer groundwater abstraction rates within a small agricultural region.

## Chapter 3

### Measuring large-scale evaporation from space <sup>1</sup>

---

This chapter introduces a number of remote sensing methods that are currently used for large-scale estimation of evaporation, and outlines a methodology to determine which (if any) of these products are suited to accurately estimate water losses in arid regions with intense agricultural presence. The methodology relies on the concept of “hydrological consistency”, which aims to reconcile independent hydrologic data obtained through different physical approaches and with data from different platforms. The methodology uses satellite-based gravity data to estimate changes in water storage, as described in Chapter 2, and precipitation estimates from a merged multi-satellite and ground-based precipitation analysis product. The methodology is applied to four large-scale basins in arid and semi-arid environments, assuming their relatively simple large-scale water budgets would enhance the applicability of the method. The study concluded that in these conditions, currently available large-scale evaporation products are not hydrologically consistent with large-scale water storage changes detected using GRACE data (chapter 2), likely due to the lack of evaporation modeling from localized sources and/or a mismatch between large- and small-scale land cover classification.

---

<sup>1</sup>Portions of this chapter derive from an edited version of: Lopez O, Houborg R and McCabe MF (2017): Evaluating the hydrological consistency of evaporation products using satellite-based gravity and rainfall data, *Hydrol. Earth Syst. Sci.*, 21, 323-343, doi: 10.5194/hess-21-323-2017

### 3.1 Introduction

Progress in satellite-based observation of the Earth system has enabled the characterization of land surface hydrological components and an improved representation of terrestrial processes (Famiglietti et al., 2015). Dedicated space missions such as the Gravity Recovery and Climate Experiment (GRACE; see section 2.1)(*Tapley et al.*, 2004b), the Global Precipitation Measurement Mission (GPM) (*Hou et al.*, 2014), and a suite of microwave-based soil moisture platforms (*Liu et al.*, 2012) represent important efforts that have contributed to these advances. Considering the spatial advantage that space-based observations have over ground-based measurements, there has been a proliferation of regional- to global-scale data products, providing knowledge on the multi-scale behavior and patterns of hydrological states and fluxes useful for enhanced process description (*Stisen et al.*, 2011). However, one of the challenges of space-based remote sensing is how to characterize the degree to which these products represent realistic estimates of the underlying variables they attempt to retrieve.

Terrestrial evaporation ( $E$ ), comprising the sources of soil and canopy evaporation together with plant transpiration, plays a key role in the water cycle as a linking mechanism between the surface and the atmosphere (*Mueller et al.*, 2011). Unlike microwave or radiative emissions from the surface or atmosphere, which can be used to inform upon soil moisture, surface temperature, or rainfall, evaporative fluxes provide no directly observable trace that can be detected from satellites, and are instead estimated through interpretive or empirical models (*Jiménez et al.*, 2011; *Ershadi et al.*, 2014). Recently, several of these models have been used to develop global-scale evaporation products by combining satellite observations of surface variables with meteorological and other ancillary data (*McCabe et al.*, 2016; *Miralles et al.*, 2016). When ground-based flux observations are available, they can be used for calibration and evaluation (*McCabe et al.*, 2005a), but large-scale assessment is inevitably

constrained by the lack of distributed and representative *in situ* networks to comprehensively assess simulations, as well as the inherent uncertainty associated with these observations (*Jana et al.*, 2016). Some recent evaluation efforts have sought to estimate the uncertainty of satellite-based evaporation products as well as those from land surface model and reanalysis data, in terms of the variance amongst the products (*Mueller et al.*, 2011; *Jiménez et al.*, 2011; *Long et al.*, 2014). These and related attempts have shown that no single evaporation model consistently outperforms any other, whether applied at local (*Ershadi et al.*, 2014) or global scales (*Miralles et al.*, 2016), even in cases where the same input data have been used. Considering the issue of spatial mismatch and model variability, it seems inappropriate to assess these large-scale products via direct comparison to *in situ* data alone. Moreover, the quality of any satellite-based product should not be judged solely on its agreement with potentially unrepresentative point-scale approaches. Central to this challenge is the issue of scale, a consequence of both a lack of abundant high-quality *in situ* data and the fact that there is an inevitable scale mismatch between ground- and satellite-based observations (*McCabe et al.*, 2005b). To compensate for this, it is important that a range of methods be used to evaluate the large-scale implementation of evaporation models.

Beyond the assessment of evaporation models, a limited number of studies have sought to quantify large-scale water budgets using either satellite observations alone (*Sheffield et al.*, 2009) or through a combination of satellite observations and data assimilation (*Pan and Wood*, 2006; *Pan et al.*, 2008, 2012; *Sahoo et al.*, 2011). While some of these studies (*Sheffield et al.*, 2009; *Gao et al.*, 2010) evaluate water budget closure by comparing the residual of the water budget (i.e. inferred runoff) with measured runoff, others aim to provide merged or observation-constrained estimates of the water cycle components, with estimates of uncertainty given in terms of the variability among the products (e.g. *Long et al.* (2014)). The results of these studies

have generally illustrated large water budget closure errors, focusing on the temporal scale and invoking the use of a hydrological model to guide analysis or force closure, rather than being solely observation-driven assessments. Observation-only studies are important, as they not only provide an unbiased perspective on hydrological closure, but also allow for a first-order examination of the underlying agreement between component variables. However, rather than just comparing the uncertainties between evaporation products and other hydrological components (which are poorly defined), there is still a need for alternative assessment techniques that exploit the inherent connection between hydrological variables at both temporal and spatial scales. One approach to determining this is to evaluate the hydrological consistency between observed products (*McCabe et al.*, 2008).

The term “hydrological consistency” refers to the spatial and temporal match that should exist between independent observations of hydrological states and fluxes, based upon physical considerations. It is a concept that encompasses the expectation of water cycle behavior and mass balance: that is, changes in one term should be reflected in related variables, both spatially and temporally. For instance, a rainfall event should result in an observable change in soil water storage and a consequent increase in evaporative flux, which in turn should reduce the available soil moisture. This relatively simple concept has been explored in the recent past, including in efforts to improve precipitation events by employing cloud detection methodologies (*Milewski et al.*, 2009); using soil moisture changes to infer precipitation amounts (*Brocca et al.*, 2014); examining the connection between soil moisture state and changes in atmospheric variables such as humidity and sensible heat flux (*McCabe et al.*, 2008); as well as in assessments of land-atmosphere coupling between observations and reanalysis data (*Ferguson and Wood*, 2011).

In considering these earlier contributions, there remains a need to determine whether the basic idea of hydrological consistency can be realistically extended to

explore the agreement between independent global-scale satellite-based hydrological products. To examine this question, it makes sense to focus on catchments that have relatively simple hydrological interactions, as they represent natural laboratories within which the evaluation of large-scale products and the concept of hydrological consistency can be reasonably undertaken. For example, *Wang et al.* (2014) evaluated the level of agreement between three satellite-based hydrological cycle variables over arid regions in Australia, where surface and sub-surface runoff were minimal. With a sufficiently low runoff component, a lack of snow accumulation, and a relatively strong coupling of precipitation and evaporation components, arid and semi-arid environments represent potential candidates within which to undertake such process assessments. Recognizing the need to advance a more comprehensive evaluation strategy for remote sensing retrievals, this study seeks to explore the hydrological consistency within a number of basins where hydrological processes are relatively simple, i.e. reflecting the conditions described above. Our analysis constitutes a framework for assessing the utility of hydrological consistency to evaluate remotely sensed hydrological products. This framework involves comparing first the spatial and temporal agreement of these evaporation products with precipitation and water storage variations, and then identifying differences in the spatial agreement among the solutions obtained when using the different evaporation data.

Constrained by the large size requirement in GRACE data, identifying potential basins with a relatively simple hydrological system over different regions in the world was challenging. In this study, we undertake the analysis over four large river basins within arid and semi-arid environments distributed across the globe, with study regions comprising the Colorado River basin in North America, the Niger River basin in Africa, the Aral Sea basin in Asia, and the Lake Eyre basin in Australia. Although the Colorado River basin and Aral Sea basin do contain potential sources of delayed runoff (e.g. snowmelt) that can affect the analysis, these were still included for com-

parison. In contrast, the Lake Eyre basin has no runoff component into the ocean, so it represents one ideal region of study in which to expect hydrological consistency.

The overall objective of this study is to evaluate the hydrological consistency of three global-scale satellite-based evaporation products against remotely sensed retrievals of precipitation and terrestrial water storage across a selection of basins that exhibit relatively well-defined hydrological interactions. Throughout this analysis we aim to determine whether the hydrological consistency concept can expand the range of evaluation metrics used to assess large-scale hydrological data sets such as evaporation, and enable some differentiation of relative product quality to be made.

### **3.2 Evaporation products and auxiliary data**

Several satellite-based evaporation products have been developed over the last decade, based on a range of modelling schemes (*Mu et al.*, 2011; *Leuning et al.*, 2008; *Miralles et al.*, 2011a) and global-scale input data. Given the importance of evaporation within studies of the global energy and water cycle, considerable effort has been directed towards accurately reproducing its spatial and temporal variability, with comprehensive reviews of various approaches to do this provided by *Kalma et al.* (2008) and *Wang and Dickinson* (2012). Here we employ a range of global evaporation data sets, which are briefly described in the following paragraphs and summarized in table 3.1. Two of these products (MOD16 and CSIRO) are largely based on the Penman-Monteith equation, with key differences in the modeling of canopy conductance. Another product (GLEAM) relies on calculating potential evaporation using the Priestley-Taylor equation and then correcting the estimate using a water stress factor. Soil moisture observations are assimilated in GLEAM and incorporated into the calculation of the stress factor. GLEAM also incorporates observed precipitation in a rainfall interception module. Another evaporation model which is not included in this study due to the lack of a global product (because of the higher number of input requirements) is

discussed in section 4.2.1. To ensure consistency with the GRACE data, the evaporation products were aggregated from daily (or 8-daily in the case of the MODIS Global Evapotranspiration product - MOD16) to monthly estimates using the same date bounds specified in the GRACE monthly gravity field solutions. In the aggregation from daily to monthly data, pixels that presented missing data for more than 20% days in a given month were not included in the calculation. This only occurred for isolated pixels representing less than 1% of the total data. Furthermore, because there is a smoothing filter applied to the data (see section 3.4), the effect from removing single pixels of data becomes negligible. Finally, given the main objective of this work (to evaluate evaporation products), we fixed the precipitation estimates to one dataset only. However, the analysis was repeated using an alternative precipitation product, finding no significant differences in the results when using this alternate product.

### 3.2.1 MOD16

*Cleugh et al.* (2007) developed an algorithm for large-scale evaporation monitoring based on the Penman-Monteith (PM) equation, using meteorological forcing data and a surface resistance linearly modeled through remotely sensed leaf area index (LAI), as measured by the MODerate resolution Imaging Spectroradiometer (MODIS). Improvements to this approach *Mu et al.* (2007, 2011) led to the development of the MODIS Global Evapotranspiration product (MOD16), a three-source scheme used for terrestrial land flux estimation. In MOD16, the linearization of the surface resistance is specified for each biome separately via a look-up table, with the evaporation calculated for daytime and nighttime conditions. Other adjustments incorporated into MOD16 include soil heat flux calculation, distinction of dry and wet canopy, as well as moist and wet soil, and improvements to the aerodynamic resistance. The MOD16 product comprises transpiration, evaporation from the soil and wet canopy, as well

Table 3.1: Description of the satellite products used in this study. The temporal resolution is daily except for MOD16 (8-daily) and GRACE (monthly). The original MOD16 product is available at 1 *km* resolution in the sinusoidal projection. In this study, the product was reprojected onto a 0.05° regular grid using the MODIS Reprojection Tool (MRT).

Product name	Spatial resolution	Time span	Reference
<i>Evaporation</i>			
MOD16 (A2)	0.05°	2000 - 2013	<i>Mu et al. (2011)</i>
CSIRO-PML	0.25°	1981 - 2011	<i>Zhang et al. (2012)</i>
GLEAM (v2A)	0.25°	1980 - 2011	<i>Miralles et al. (2011a,b)</i>
<i>Water storage</i>			
GRACE (CSR RL05)	333km ( $l_{max} = 60$ )	2003 - present	<i>Tapley et al. (2004a)</i>
<i>Precipitation</i>			
GPCP (1DD v1.2)	1°	1996 - present	<i>Huffman et al. (2001)</i>

as total evaporation calculated as the sum of these three components. Each component is weighted-based on the fractional vegetation cover, relative surface wetness and available energy. Inputs to the model include net radiation ( $R_N$ ), air temperature and humidity, as well as LAI and vegetation phenology. Importantly, it does not require wind speed, precipitation or soil moisture data, making it a relatively parsimonious model in terms of input requirements. In this study, we used the actual evaporation (AET) product from MOD16 *Mu et al. (2011)* with 8-day temporal resolution and 1 *km* resolution in the sinusoidal projection. The product was reprojected onto a 0.05° regular grid using the MODIS Reprojection Tool (MRT) before transformation into spherical harmonics (section 2.2). Further details on the modeling basis behind the MOD16 product can be found in *Mu et al. (2013)*, *Ershadi et al. (2014)* and *Michel et al. (2016)*.

### 3.2.2 CSIRO-PML

In parallel to the PM-Mu model, *Leuning et al. (2008)* introduced improvements to the *Cleugh et al. (2007)* algorithm, resulting in the two-source Penman-Monteith-Leuning (PML) model. An important new feature of the PML approach was a biophysical algorithm for the calculation of the surface resistance, which was previously calculated

as LAI multiplied by a constant  $c_L$  (Cleugh *et al.*, 2007). The new parameterization of the surface resistance in the PML model was optimized using data from 15 globally distributed flux station sites, with two key parameters identified: the maximum stomatal conductance ( $g_{sx}$ ) and the ratio of actual to potential evaporation at the soil surface. Zhang *et al.* (2010) developed a method to further optimize the spatial variability of these two parameters (i.e. at each grid pixel) using gridded meteorological data and a simple Budyko-curve hydrometeorological model developed by Fu (1981) that includes precipitation and available energy as inputs. Mean annual evaporation for each grid pixel is calculated using the Fu model and gridded meteorological data. The value of  $g_{sx}$  is optimized using a non-linear least square regression based on the difference between the PML and the Fu model. Interestingly, the Fu model is calibrated by comparing the output evaporation with the residual of precipitation and runoff, i.e. by assuming negligible annual water storage changes and groundwater inflow and outflow. Zhang *et al.* (2012) used this approach to develop a global gridded terrestrial evaporation product (hereafter referred to as CSIRO-PML; Zhang, 2014, personal communication) with a  $0.25^\circ$  resolution (in this study, we used the actual evaporation product). They used gridded meteorological data from diverse sources, including vapor pressure and temperature from the Climate Research Unit (New *et al.*, 2000), LAI and land cover type from Boston University (Ganguly *et al.*, 2008), precipitation from the Global Precipitation Climatology Centre (GPCC, version 4; Schneider *et al.* (2011), and radiation data from the Global Energy and Water Cycle Exchanges (GEWEX) Surface Radiation Budget (Gupta *et al.*, 2006).

### 3.2.3 GLEAM

The Global Land Evaporation: the Amsterdam Methodology (GLEAM) (Miralles *et al.*, 2011a) is a satellite-based model developed to estimate evaporation at a global scale. In this approach, rainfall interception loss is evaluated using an analytical

model (*Gash*, 1979) as a first step. GLEAM then employs the Priestley-Taylor equation to calculate the potential evaporation of bare soil and vegetation components (both short and tall canopy), with values constrained to actual evaporation via application of a stress factor. The stress factor is calculated using vegetation optical depth from a combination of different satellite passive microwave observations using the Land Parameter Retrieval Model (*Liu et al.*, 2013). GLEAM also has the capacity to explicitly calculate sublimation of snow-covered surfaces (*Takala et al.*, 2011) as well as open water evaporation. Satellite observations of surface soil moisture can be assimilated using a Kalman filter assimilation approach to estimate the moisture profile over several soil layers. Here we employ version 2A of GLEAM, which uses a combination of satellite, ground, and reanalysis input data. Precipitation is obtained from the Climate Prediction Center Unified data set, consisting of data from over 30,000 stations (CPC-Unified, *Joyce et al.* (2004)). The radiation product used in this version of GLEAM is the European Center for Medium-Range Weather Forecasts (ECMWF) ERA-Interim meteorological reanalysis product (*Dee et al.*, 2011). In this version of GLEAM, surface soil moisture data from the Water Cycle observation Multi-mission Strategy Climate Change Initiative (WACMOS-CCI) merged product (from a combination of several passive and active microwave products) are assimilated (*Liu et al.*, 2012), while air temperature is derived from both the International Satellite Cloud Climatology Project (ISCCP) and the Atmospheric Infrared Sounder (AIRS) (*Rossow and Dueñas*, 2004). Further details of the model can be found in *Miralles et al.* (2010, 2011a,b).

### 3.2.4 Precipitation data: GPCP

Global daily precipitation ( $P$ ) estimates derived from multi-satellite observations for the period 2003-2011 were obtained from the Global Precipitation Climatology Project (GPCP) (*Huffman et al.*, 2001), the official World Climate Research Program

(WCRP) Global Energy and Water Cycle Exchanges (GEWEX) product. The GPCP product is a merged precipitation analysis combining information from microwave, infrared, and sounder data observed by a constellation of international precipitation-related satellites (*Huffman et al.*, 1997, 2001; *Adler et al.*, 2003). The estimates from microwave and infrared data are based on the Threshold-Matched Precipitation Index (TMPI). The combined satellite-based product is corrected by rain gauge analysis where data are available (i.e. the intermediate product from the Global Precipitation Climatology Center; *Schneider et al.* (2011)). Over many areas of the world, the GPCP product represents one of the best available sources of precipitation data and has been previously used in soil-moisture- and evaporation-based analyses (*Crow*, 2007; *Miralles et al.*, 2011a). In this study, we used the daily product (GPCP 1DD) and converted daily values to monthly estimates, centered on the dates provided in GRACE monthly gravity field solutions. Pixels were assigned as missing data when more than 20% of the month was missing (on a per-pixel basis).

### 3.2.5 Runoff data

Because of the constrain in size from GRACE data, finding large basins with the ideal conditions for this study (e.g. low or no sources of runoff) was challenging. To put into perspective the relative contribution of total runoff to water storage changes, we used runoff data to compare the amplitudes with precipitation and evaporation data. Surface runoff, sub-surface runoff, and snowmelt were derived from the NOAH land surface model included in the Global Land Data Assimilation System (GLDAS) (*Rodell et al.*, 2004b). GLDAS uses global satellite and ground-based observational products to obtain optimal estimates of land surface states and fluxes from land surface models using data assimilation techniques. Although these values were not constrained with ground estimates and thus may contain biases, as noted, runoff values were only used to provide an assessment of runoff against the observed

precipitation and evaporation data. The version of the product used in this study (GLDAS-2.0) is forced with meteorological data from the Princeton University forcing data set (*Sheffield et al.*, 2006) and is available at 1° resolution from 1948 to 2010.

### 3.3 Selection of study basins

The study basins were targeted primarily on their climate classification, with river basins in regions with a predominantly arid or semi-arid climate preferentially selected. This criterion was established in order to seek a relatively simple hydrological system (i.e. constrain the range of possible hydrological interactions), thereby maximizing the conditions under which hydrological consistency between evaporation and precipitation and water storage changes might be achieved. A Köppen classification map, generated using data sets from the Climatic Research Unit and the Global Precipitation Climatology Centre up to 2006 (*Kottek et al.*, 2006), was used to identify arid and semi-arid regions. The basins were selected from a set of 405 globally distributed river basins provided by the Global Runoff Data Centre (GRDC) and derived from flow direction data of the HYDRO1k Elevation Derivative Database, developed at the U.S. Geological Survey (USGS). A threshold of 50% areal extent containing any of the arid Köppen climates (BWk, BWh, BSk or BSh) was used to select potential basins. Secondary criteria for basin selection from the GRDC data set focused on size, geographical distribution and amplitude and trends in the water storage variations. In terms of the size of the basin, a smaller size would more likely satisfy the assumption of a relatively simple hydrological system. However, due to the coarse resolution of GRACE data (see section 2.1), this requirement had to be compromised. Given these considerations, four basins were selected as focus regions of study: the Colorado River basin (CRB) in North America, the Niger River basin (NRB) in Africa, the Aral Sea basin (ASB) in Asia and the Lake Eyre basin (LEB) in Australia (figure 3.1).

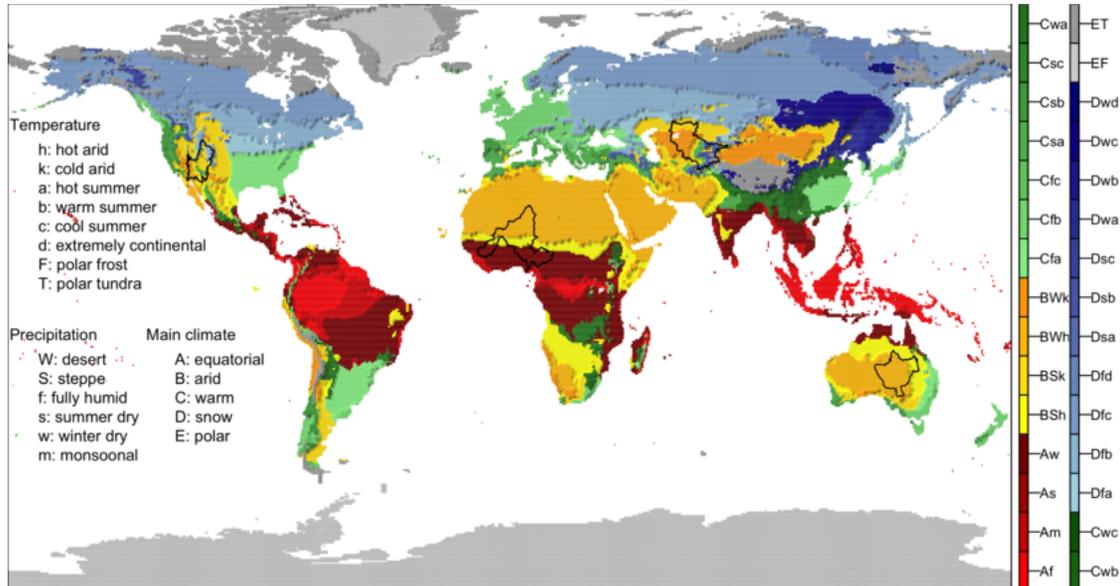


Figure 3.1: Selected study basins used within the analysis. Criteria for the selection of basins included: predominantly arid climate (more than 50% areal coverage with any of the arid Köppen climates: BWk, BWh, BSk or BSh), size, geographical location and amplitude and trends in the water storage variations.

Figure 3.2 shows the spatially averaged hydrological fluxes over the study basins, including the sum of surface, subsurface and snowmelt runoff ( $Q$ ) derived from the GLDAS NOAH version 2 monthly product (*Rodell et al., 2004b*).  $Q$  is included in these figures to establish the extent to which a major assumption of the study (i.e. a simple water budget) is met across each of the study regions. Although predominantly arid, with a combination of hot arid desert and cold arid steppe climate classifications, both the Colorado River basin and the Aral Sea basin contain a snow component. Snowmelt in these two regions plays an important role in the water cycle, particularly in the delivery and redistribution of water to other areas of the basin. Therefore, hydrological consistency might not be satisfied completely in these regions using our simple water budget assumption for some periods. Likewise, the Niger River basin also has a runoff component that is close in magnitude to evaporation, but we assume that it will not affect the spatial distribution of water storage anomalies. In the Lake Eyre basin we expect that the limited and sporadic runoff component will not have

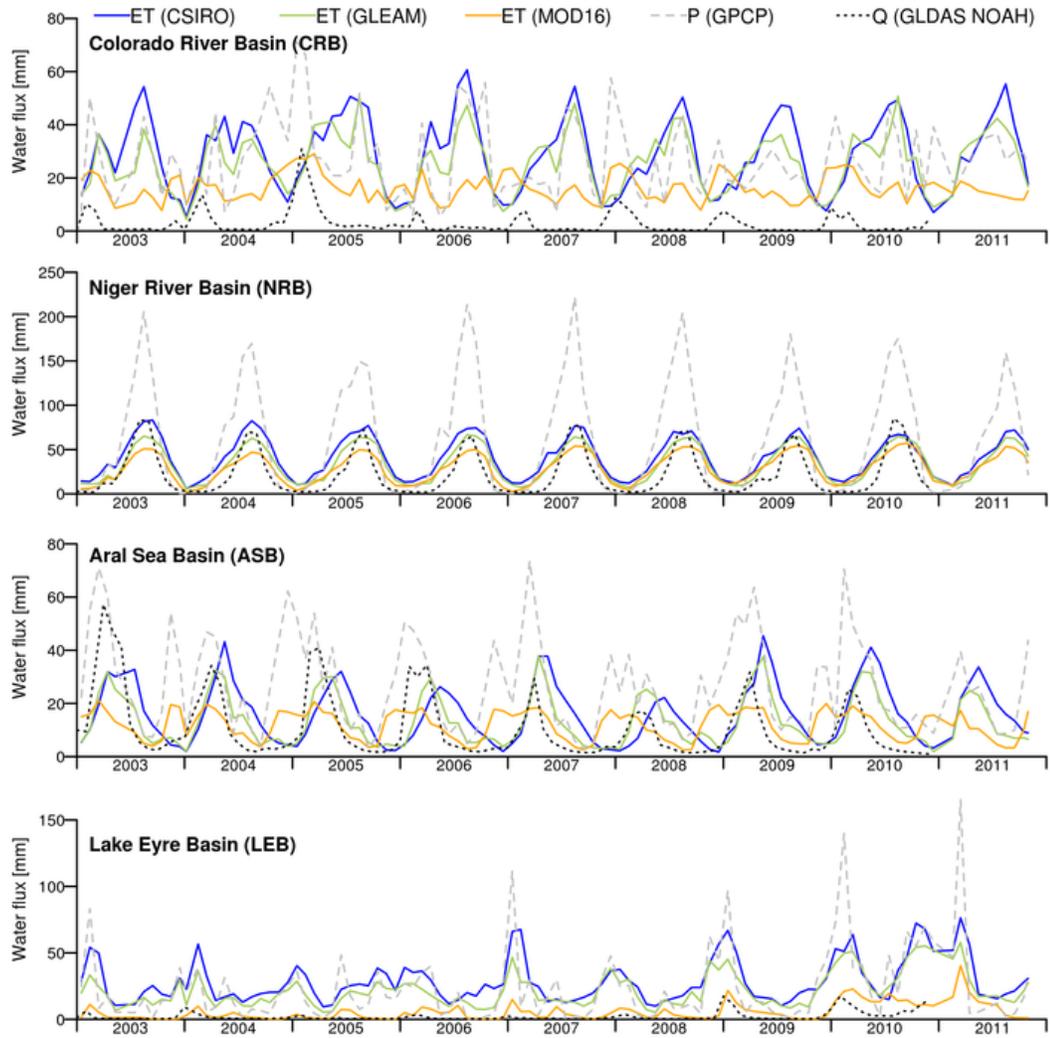


Figure 3.2: Average  $P$ ,  $E$  and  $Q$  fluxes within the four study basins for the period 2003-2011. Three evaporation ( $E$ ) data sets were used in this study, including CSIRO-PML, MOD16 and GLEAM (see table 3.1 for details).

a significant effect on the hydrological consistency analysis undertaken there.

Even though these four basins were preselected based upon their location within dryland systems (*Wang et al., 2012*), they reflect a range of trends in water storage and precipitation. For example, the Colorado River basin experienced intervals of wet and dry periods, while the Niger River basin exhibits a small but steady increase in water storage with a clear seasonal variability in both water storage and precipitation. Meanwhile, the Aral Sea basin experienced a significant loss of water during the

study period ( $-8.2 \text{ mm.yr}^{-1}$ , calculated from GRACE water storage variations; see figure 3.6), in line with the historical depletion of this inland sea in response to increased agricultural productivity (*Zmijewski and Becker, 2014*). The Lake Eyre basin showed a marked increase in precipitation during the end of the study period (2009-2011), with a corresponding increase in water storage (figure 3.7) during the following years, reflecting the larger scale hydrometeorological conditions affecting that region (*Boening et al., 2012*).

### 3.4 Methodology

In order to provide a meaningful spatial evaluation of the hydrological consistency between the data sets (i.e. at sub-basin scale) and to ensure that a fair comparison between GRACE data and satellite products could be undertaken, the analysis was carried out in spherical harmonics. The effects of the de-stripping filter (see section 2.2.4) are incorporated into the analysis directly instead of relying on a land surface model, the choice of which can severely impact the results of our analysis in arid regions (*Long et al., 2015*). In this section, we present a detailed account of how the transformation was carried out, as well as how the actual evaluation of hydrological consistency is performed in spherical harmonics.

#### 3.4.1 Spherical harmonic (SH) analysis of evaporation and precipitation data

The spherical harmonic analysis refers to the process of solving equation 2.1 for a set of coefficients  $C_{lm}$  and  $S_{lm}$  up to an approximation of degree  $l_{max}$ . Several computational packages are available to perform this type of analysis. Here we used a FORTRAN program developed by *Wang et al. (2006)*, which is suited for regularly gridded regional and/or global non-smooth data sets. The program can also perform spherical harmonic synthesis, which is the inverse transformation (i.e. from coefficients

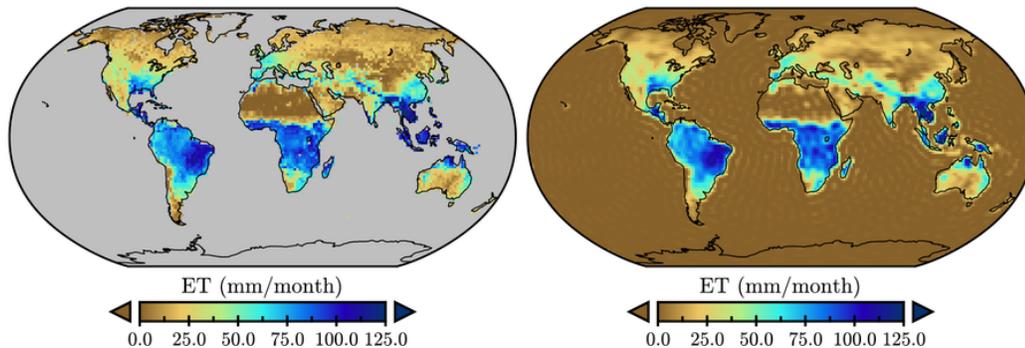


Figure 3.3: Left: CSIRO-PML monthly evaporation ( $E$ ) for April 2003. Right: the same data set after spherical harmonic analysis and synthesis of the evaporation data. Missing data are set to zero. The data appears smoothed because it is an approximation (equation 2.1) limited by  $l_{max}$ . There are also other effects such as ringing that are visible in regions where the data is close to zero.

to spatial data). Figure 3.3 presents an example of the transformation based on the gridded CSIRO-PML data for April 2003. Because all data sets are evaluated up to the same degree  $l_{max}$ , any differences due to the mismatch in the resolution of the products are eliminated after spherical harmonic analysis and synthesis. After this process, we generated three  $P - E$  anomaly data sets, i.e. one for each  $E$  product. Next, we applied the de-stripping and Gaussian filters to account for the effect that these have on GRACE TWSA data (see section 2.2.4).

### 3.4.2 Regional spherical harmonic analysis

In the analysis so far, the computed spherical harmonic coefficients are undertaken at the global scale (e.g. figure 3.3). In order to evaluate the hydrological consistency of the study regions (figure 3.1), the data need to be masked for the particular study basins. In *Swenson and Wahr* (2002), an exact averaging kernel is defined as a function with a value of 1 inside the boundaries of a region and 0 outside. To isolate the GRACE signal, an approximated averaging kernel was computed in spherical harmonics and convolved with the Gaussian filter in order to obtain a spatially averaged value of the TWSA (at the basin scale). In this study, we instead compute the spher-

ical harmonic from the product of the global data sets (e.g. TWSA or  $P - E$ ) with the averaging kernel:

$$f^b(\theta, \phi) = f^g(\theta, \phi) \vartheta(\theta, \phi), \quad (3.1)$$

where  $f^b(\theta, \phi)$  is the isolated regional data,  $f^g(\theta, \phi)$  is the global data set and  $\vartheta(\theta, \phi)$  is the averaging function. The relation in spherical harmonics is given by:

$$f^b = \sum_{j_1, m_1} \sum_{j_2, m_2} f_{j_1, m_1}^g \vartheta_{j_2, m_2} Q_{j_1 m_1 j_2 m_2}^{j m}, \quad (3.2)$$

$$Q_{j_1 m_1 j_2 m_2}^{j m} = \sqrt{\frac{(2j_1+1)(2j_2+1)}{4\pi(2j+1)}} C_{j_1 0 j_2 0}^{j 0} C_{j_1 m_1 j_2 m_2}^{j m}, \quad (3.3)$$

where  $C_{j_1 m_1 j_2 m_2}^{j m}$  are the Clebsch-Gordan coefficients (*Martinec*, 1989). We used the program developed by *Martinec* (1989) to mask the three global  $P - E$  data sets (as well as GRACE data) over the four study regions (figure 3.1).

### 3.4.3 Evaluating spatial agreement in spherical harmonics

The spatial agreement between two data sets can be evaluated using spherical harmonic coefficients by computing the degree correlation measure (*Arkani-Hamed*, 1998; *Tapley et al.*, 2004b):

$$r_l = \frac{1}{\sigma_l^{(A)} \sigma_l^{(B)}} \sum_{m=0}^l (C_{lm}^{(A)} C_{lm}^{(B)} + S_{lm}^{(A)} S_{lm}^{(B)}), \quad (3.4)$$

where  $\sigma_l^2$  is the degree variance given by:

$$\sigma_l^2 = \sum_{m=0}^l (C_{lm}^2 + S_{lm}^2) \quad (3.5)$$

The degree correlation measure is computed for every degree( $l$ ), and therefore we can in principle evaluate the hydrological consistency at different length scales (i.e. sub-basin variability). As noted in section 2.2, GRACE data is limited in resolution by  $l_{max} = 60$ , or to approximately 330  $km$ . In practice however, the de-stripping filter removes all coefficients larger than 40 and as such we are limited to length scales of about 500  $km$  and larger. The smallest basin in this study is the Colorado River, covering an area of about 640,000  $km^2$ . Based on this area, we can set a limit for the approximate largest spatial scale relevant to our study as 800  $km$ , corresponding approximately to degree 25.

## 3.5 Results

### 3.5.1 Assessing the consistency of evaporation products

An examination of the evaporation data sets indicates that there are evident differences across the various products in each of the studied basins (see figure 3.2). In general, MOD16 simulates lower flux estimates when compared against both CSIRO-PML and GLEAM, a feature that has been noted in a number of recent global inter-comparison studies (*McCabe et al.*, 2016; *Michel et al.*, 2016; *Miralles et al.*, 2016). There are also clear differences in terms of the variability in the temporal response of the models, although CSIRO-PML and GLEAM show a greater level of agreement in terms of amplitude and timing, if not in absolute values. For example, during the wet period of 2004–2005 in the Colorado River basin, the response to precipitation reflected in MOD16 was far more rapid than either CSIRO-PML or GLEAM displayed. Of some concern is that CSIRO-PML is larger than precipitation during much of the study period in both the Colorado River and Lake Eyre basins, immediately negating any type of hydrological consistency analysis. In the Niger River basin, there is more consistent agreement between the evaporation products, indicating greater confidence in the retrievals of evaporation in this region. For the Aral Sea basin, the

discrepancies in  $E$  are similar to those obtained for the Colorado River basin, with an obvious phase shift in CSIRO-PML and GLEAM observed relative to MOD16. This may reflect complexities in evaporation modelling due to the intermixed climate zones in the region caused by differences in land surface parameters. In the Lake Eyre basin, there are differences in amplitude but not in the temporal behavior of  $E$ .

Overall, even from a qualitative perspective, there are clear challenges in developing a hydrological consistency approach over these comparatively simple basins. Indeed, this has been demonstrated in other studies using either satellite data alone or a combination of satellite and ground data. While it is not the intention of the current work to explore the error characterization of these different evaporation models based on hydrological closure, the techniques being used to evaluate product consistency should provide some insight into retrieval quality, at least relative to the other hydrological products (precipitation and gravity-based water storage changes) that the evaporation is being compared against. These ideas are explored more quantitatively in the following sections.

### 3.5.2 Basin-scale assessment

In this section, we examine the spatial and temporal patterns of the degree correlations between water storage variations (TWSA) and  $P - E$  anomalies. Figures 3.4 to 3.7 present the results of this assessment across each of the four large-scale basins. For each of the figures, time series of the spatial average TWSA and  $P - E$  anomalies are shown in order to compare their trends with the temporal behavior of the degree correlation ( $r_l$ ). In these figures, the degree correlation is a measure of the spatial agreement between the two fields being compared (i.e. TWSA and  $P - E$  anomalies), assuming that other outflow components such as surface runoff (assumed to be minimal in these basins) do not directly affect these spatial patterns. This comparison is helpful in determining whether the cause of trends in water stor-

age variations (either natural or anthropogenic) influence the analysis of hydrological consistency e.g. do the degree correlations behave differently during wet or dry periods, or when storage changes are driven by natural or anthropogenic causes? In these figures, the response of the degree correlations is shown in time across the x-axis and in the spectral domain along the y-axis, for each of the three evaporation products.

### 3.5.2.1 Colorado River basin

The start of the study period (2003) coincided with the end of an intense multi-year drought in the Colorado River basin (*Scanlon et al.*, 2015). During the wet period of 2004-2005 (see figure 3.2), the basin showed a corresponding increase in TWSA (see figure 3.4), although with a delay in time of two to three months. During this time of increase in TWSA, there was a corresponding increase in  $r_l$  (up to 0.9 for  $l = 25$  and 0.8 for  $l = 40$ ) until TWSA reached its peak value (November 2004 - February 2005), after which  $r_l$  decreased and showed negative values (similar, but negative, i.e. -0.9 for  $l = 25$ ; -0.8 for  $l = 40$ ) during the TWSA decrease. During the dry period (2008-2009), TWSA is correspondingly lower, but oscillating out of phase with  $P - E$  anomalies (with about 2 months of lag). There seems to be a stronger relationship between the oscillation of TWSA and degree correlations during the wet and dry periods than for other catchment conditions, where the variations of  $r_l$  appear random. In general, the degree correlations for small degrees have larger amplitudes than those for large degrees. This spatial disagreement in correlation might be related to the spatial and temporal distribution of runoff in the basin, since a large portion of the runoff comes from snowmelt originating in the upper portions of the basin (*Scanlon et al.*, 2015). Differences in absolute values and in the temporal distribution of  $E$  (especially with the MOD16 product) were evident in the degree correlation images in figure 3.4. However, they did not have a significant impact on the analysis in the sense of demonstrating any product advantage or disadvantage

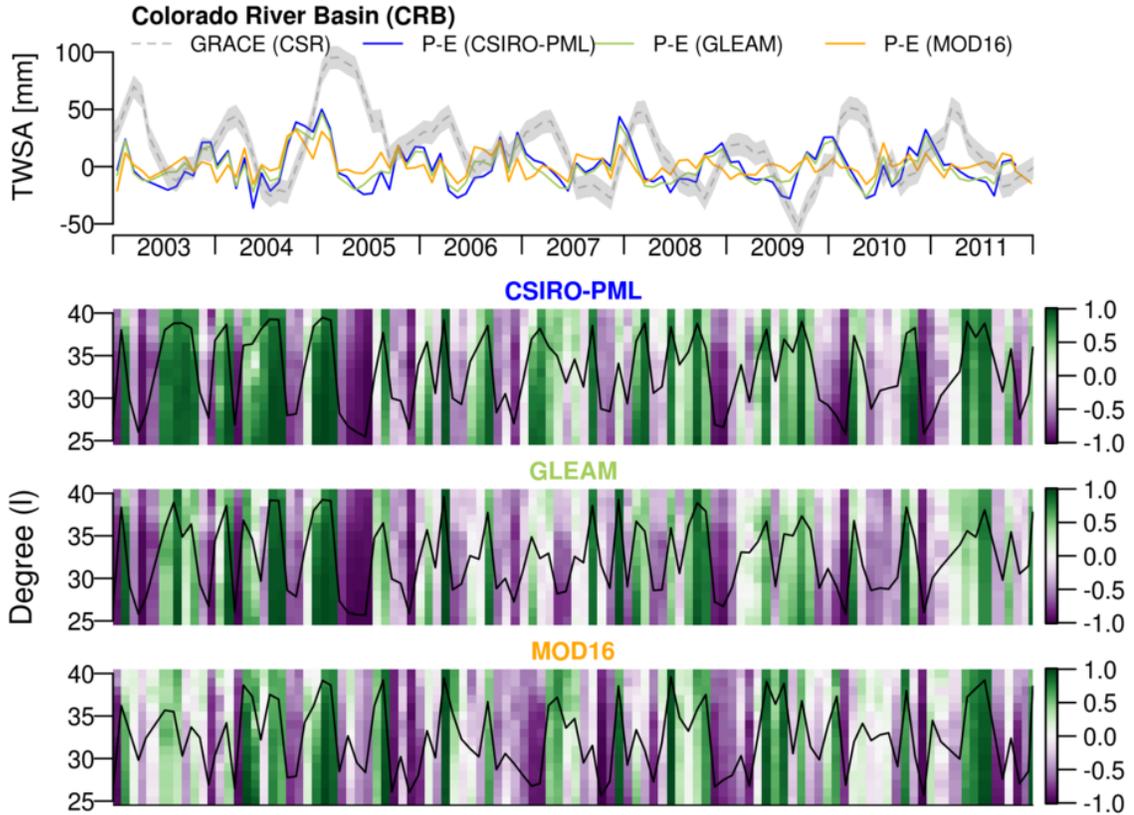


Figure 3.4: Top: anomalies of the terrestrial water storage (TWSA) observed by GRACE (with  $20 \text{ mm}$  uncertainty bounds) and  $P - E$  using three global evaporation products over the Colorado river basin. Below: varying degree correlation measure ( $r_l$ ) with time and degree (from 25 to 40) using the three global evaporation products. The average  $r_l$  is shown as a time series (black line). The degree correlation measure can range from -1 to 1 as shown in the color scale on the secondary axis.

relative to the other evaporation products, at least in terms of their hydrological consistency.

### 3.5.2.2 Niger River basin

The TWSA in this Niger River basin was characterized by an overall steady increase ( $5.79 \text{ mm.yr}^{-1}$ ) with clear seasonal variability (see figure 3.5). Over the study period, precipitation peaks between July and September, while TWSA peaks between September and November. *Ahmed et al.* (2014) attributed the observed increase in TWSA to an increase in precipitation in the region caused by warmer Atlantic Ocean

temperatures, with the trend in  $P$  validated using multiple precipitation sources, including satellite products and rain gauges. While the GPCP data set used here did not show any increase in precipitation, neither did a recent study using rainfall estimates from the Tropical Rainfall Measuring Mission (TRMM) (*Hassan and Jin, 2016*), so the true cause of this trend remains somewhat unresolved. During the first two years of the GRACE observing period (2003-2004), the basin experienced a downward trend in TWSA. During this time, the  $r_l$  values increased at the same time as TWSA decreased towards its minimum value. Then, while TWSA values were recovering, the correlation quickly decreased and became negative. This is similar to what was observed in the Colorado River basin during the dry period. During some wet periods (e.g. July/August 2006, 2007 and 2008), when TWSA increased towards its highest value,  $r_l$  increased and was positive, but then decreased after TWSA peaked. More generally, there seems to be a connection between  $r_l$  and the water cycle variations in this region: both high TWSA and low TWSA produced positive correlations. The transitions from positive to negative values make sense considering that when TWSA values approach zero, the observations are more uncertain, as they are affected by noise (i.e. signal to noise ratio). However, the relation might also be influenced by the lag in phase between GRACE observations of TWSA and  $P - E$ . Interestingly, there seems to be less inter-degree variability compared to the other basins studied here. This may be related to the simpler water budget in this basin compared to that of the Colorado River and Aral Sea basins, but requires further investigation.

### 3.5.2.3 Aral Sea basin

The endorheic Aral Sea basin reflected a historical trend of water loss during the study period, most likely caused by anthropogenic consumption related to agricultural activities (*Zmijewski and Becker, 2014*). Although there were short intense precipitation events during much of the study period (figure 3.2), the total annual

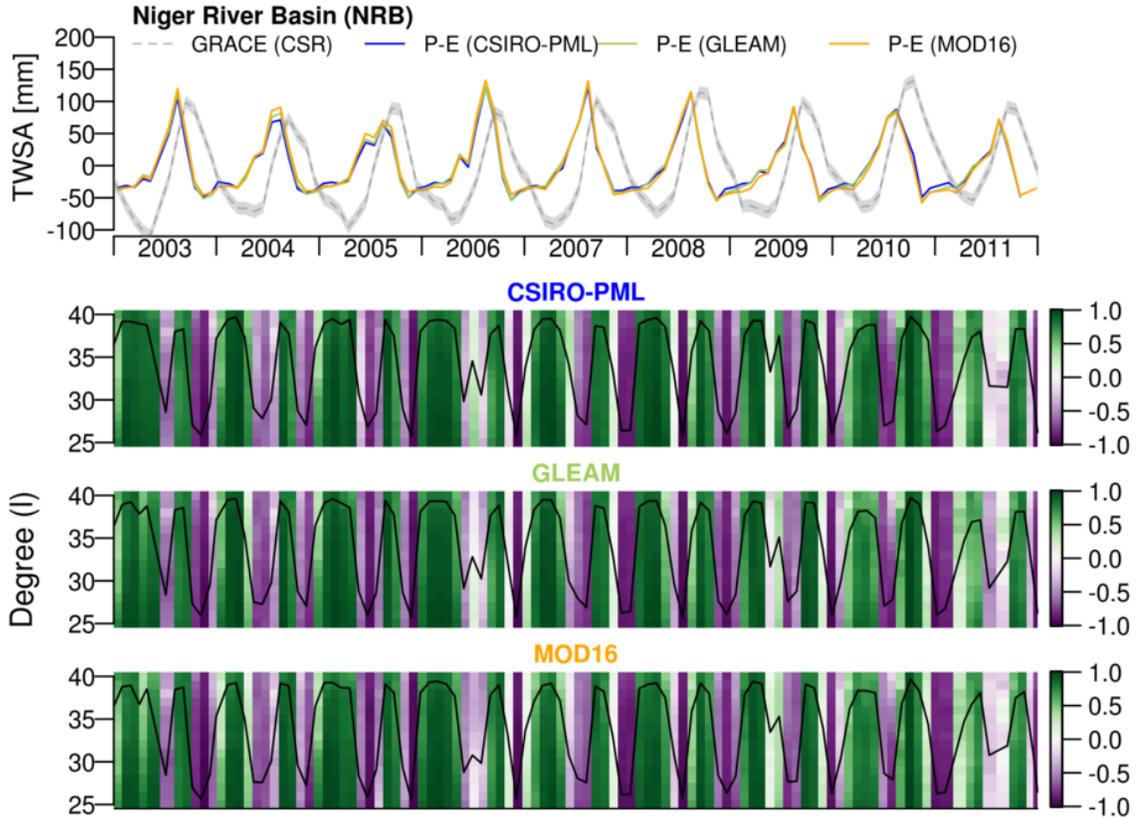


Figure 3.5: Top: anomalies of the terrestrial water storage (TWSA) observed by GRACE (with 20 mm uncertainty bounds) and  $P - E$  using three global evaporation products over the Niger river basin. Below: varying degree correlation measure ( $r_l$ ) with time and degree (from 25 to 40) using the three global evaporation products. The average  $r_l$  is shown as a time series (black line). The degree correlation measure can range from -1 to 1 as shown in the color scale on the secondary axis.

precipitation showed a negative trend of  $-31 \text{ mm.yr}^{-1}$  from 2003-2008. However, water storage values increased in 2005 as a result of the construction of a dam between the north and south portions of the Aral Sea (*Shi et al.*, 2014). During most of the study period, the  $r_l$  values oscillated in a similar way as for the Colorado River basin: that is, a weak connection between high  $r_l$  values and increasing or decreasing TWSA, before reaching the local maxima or minima, respectively (see figure 3.6). Some examples of this behavior include June-August 2008 and July-October 2009, before TWSA reaches its lowest value. In general, the  $r_l$  values decreased with increasing degree. However, inter-degree variability was more complicated in this basin during several

months. Although the Aral Sea basin is predominantly arid, the south-east portion of the basin includes a mixture of warm and cold climates and is where most precipitation occurs. Due to the mismatch in resolution and/or different land cover inputs, the evaporation products may represent these intermixed regions differently. Furthermore, glacial and snowmelt runoff present further complications to the hydrological description. These complications are reflected in the higher inter-degree variability (compared to the other basins). Differences in degree correlation due to the use of the three evaporation products were minimal i.e. no single evaporation product resulted in a significantly higher (or lower) hydrological consistency with precipitation and water storage anomalies.

#### 3.5.2.4 Lake Eyre basin

Another endorheic basin examined here was the Lake Eyre basin, which experienced a marked increase in precipitation during the rainy seasons of 2009-2011 (figure 3.2), resulting in an increase in water storage anomalies of about  $40 \text{ mm.yr}^{-1}$  (calculated from September 2009 to December 2011). The times in which TWSA and  $P - E$  were negatively correlated (i.e. negative  $r_l$  values) increased during this rainy period (see figure 3.7). Total annual precipitation decreased from 2003-2006 ( $-23 \text{ mm.yr}^{-1}$ ), with a corresponding secular decreasing trend in TWSA of  $-8.26 \text{ mm.yr}^{-1}$ . In the same period however, the degree correlations did not reveal any structure or indicate any connection with either  $P - E$  or TWSA. A short but intense precipitation event during the winter of 2006-2007 (figure 3.2) did not seem to affect the variations in  $r_l$ , relative to the earlier years. The  $r_l$  variations did show improvements during most of 2008 (when precipitation was low, i.e.  $P < 50 \text{ mm}$ ), particularly with the MOD16 evaporation product (i.e. the retrieval representing the lowest evaporation values). Overall, the  $r_l$  values generally decreased with decreasing length scales. Differences in absolute  $r_l$  values were visible between the evaporation products, but not in the overall

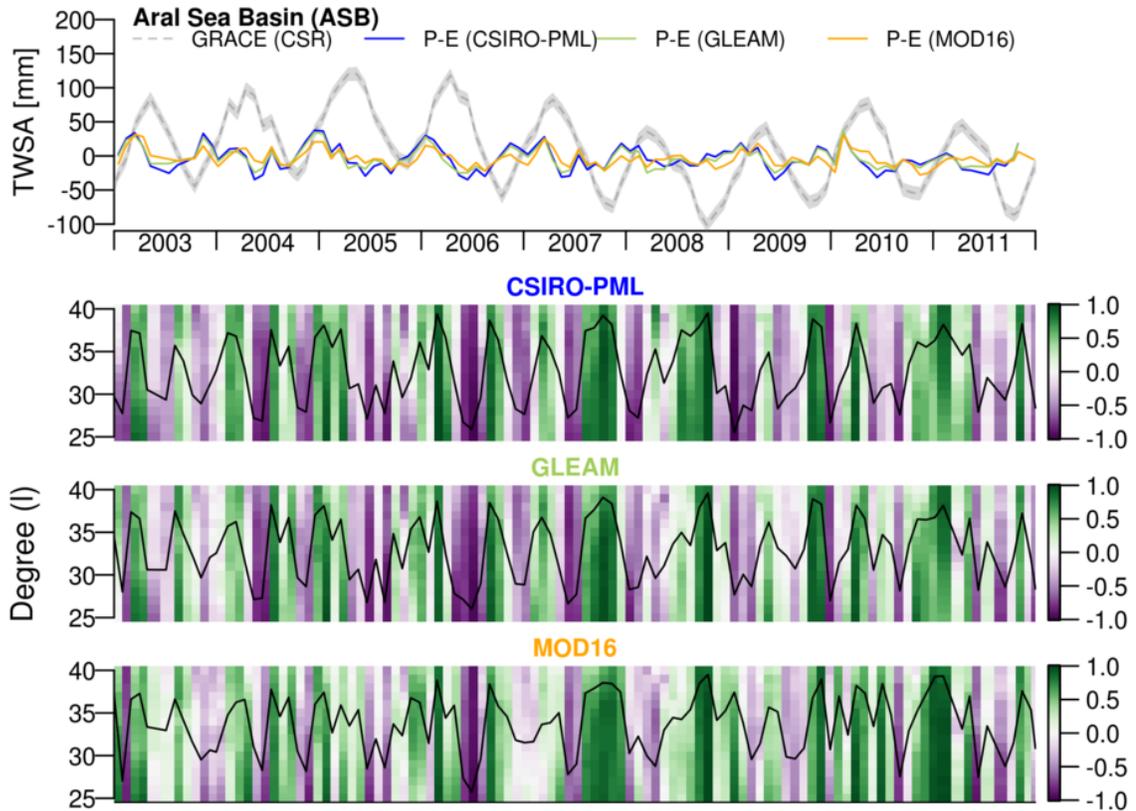


Figure 3.6: Top: anomalies of the terrestrial water storage (TWSA) observed by GRACE (with 20 mm uncertainty bounds) and  $P - E$  using three global evaporation products over the Aral basin. Below: varying degree correlation measure ( $r_l$ ) with time and degree (from 25 to 40) using the three global evaporation products. The average  $r_l$  is shown as a time series (black line). The degree correlation measure can range from -1 to 1 as shown in the color scale on the secondary axis.

spatial and temporal patterns. More importantly, none of the evaporation products showed a significant (and persistent) advantage in terms of hydrological consistency over the others. *Wang et al.* (2014) also studied the hydrological consistency of satellite products (TRMM-based  $P$ , MOD16-based  $E$  and GRACE TWSA) over this basin, as well as other predominantly arid regions of the Australian continent. At the monthly scale, their study also found poor agreement between TWSA and  $P - E$ .

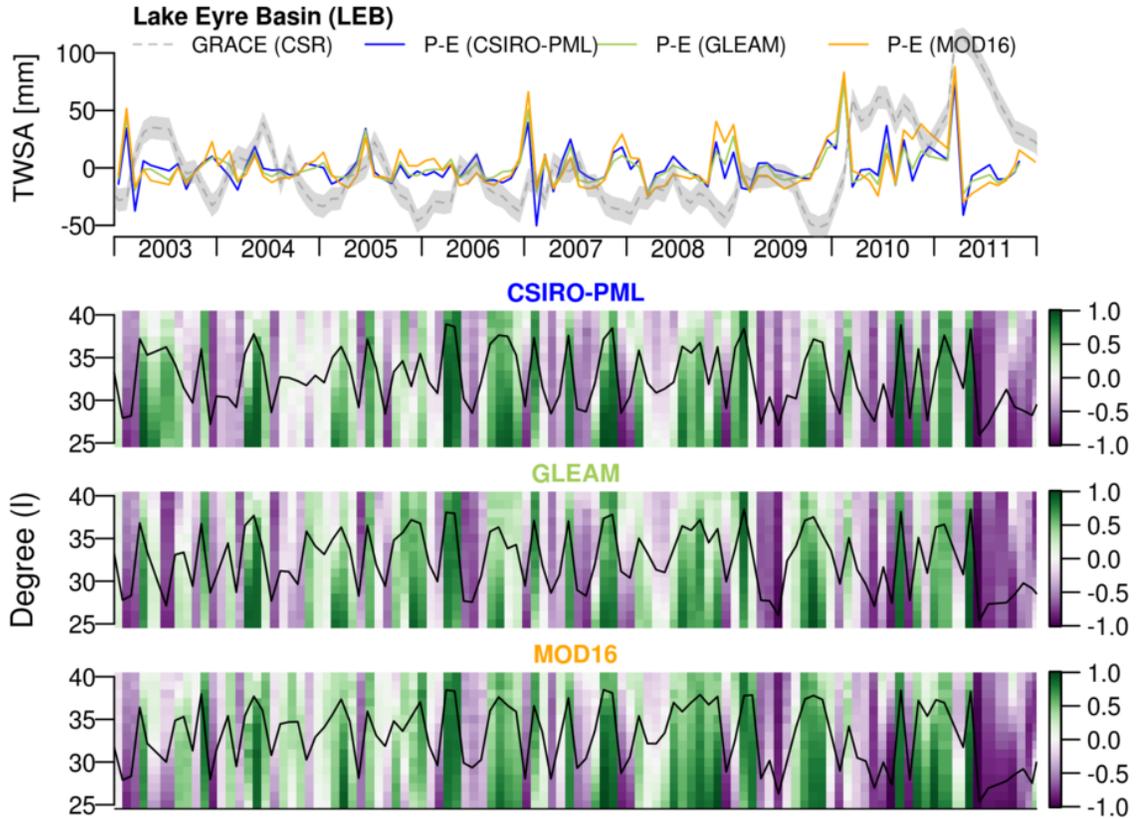


Figure 3.7: Top: anomalies of the terrestrial water storage (TWSA) observed by GRACE (with 20 mm uncertainty bounds) and  $P - E$  using three global evaporation products over the Lake Eyre basin. Below: varying degree correlation measure ( $r_l$ ) with time and degree (from 25 to 40) using the three global evaporation products. The average  $r_l$  is shown as a time series (black line). The degree correlation measure can range from -1 to 1 as shown in the color scale on the secondary axis.

### 3.5.3 Applying a phase lag to GRACE data

For GRACE to identify a water storage increase, the water mass resulting from precipitation needs to accumulate within the catchment beyond a detectable threshold. This accumulation process may take up to several months, during which time the spatially distributed rainfall drains via sub-surface processes or collects in rivers after traveling from different source areas within a basin (*Rieser et al., 2010*). The apparent lag that GRACE data illustrates relative to faster hydrological processes such as precipitation events has been observed in African basins (*Ahmed et al., 2011; Hassan and Jin, 2016*) as well as in Australia (*Rieser et al., 2010; Wang et al., 2014*).

The clearest example from amongst the basins studied here is shown in the Niger River basin (figure 3.5), where a lag of two months is evident throughout the study period. In other regions, such as the Colorado River basin and Lake Eyre basin, the time needed to detect water storage changes after precipitation events tends to vary, perhaps due to changing spatial and temporal patterns in precipitation as well as geomorphological characteristics *Ahmed et al.* (2011); *Wang et al.* (2014). Because of their large extent and geographical features, the Colorado River and Aral Sea basins include regions where snow storage plays an important role as a source of delayed runoff. The combination of snowmelt, groundwater flow and other sources of delayed flow are defined as baseflow (*Beck et al.*, 2013).

To examine this temporal component, at least in a simplified manner, a lag of one, two and three months was considered for all basins and assumed to remain constant throughout the study period. In terms of changes to the degree correlation, for the Niger River basin it was clear that a two months lag produced an improved temporal match between TWSA and  $P - E$ . For the other basins however, due to the changing dynamics in precipitation and TWSA, a temporal match could not be satisfied at all times by using an arbitrarily constant lag in GRACE. Regardless, it was found that a persistent lag of two months provided a better fit compared to all alternatives (including zero lag). *Beck et al.* (2013) developed global estimates of the Base Flow Index (BFI): a measure of the ratio of the long-term baseflow to the total runoff, using a large global data set of runoff and a regionalization procedure to transfer these and other characteristics of runoff from gauged to ungauged basins. Since we did not model any of the physical processes contributing to baseflow, the BFI was examined to assist in explaining part of the delay in observed water storage changes relative to the  $P - E$  term (although not dynamically, since the index is a long-term average in time). The spatial average of the BFI in the four study basins is, not surprisingly, within the same range: between 0.4 and 0.6. This is not unexpected, as

various climate characteristics were used as predictors of BFI. Indeed, the fact that they are similar is in agreement with our finding of similar GRACE lag times among the study basins. Further investigation is required to determine the nature of the elements affecting the lag in water storage, not limited to those found in baseflow.

Figure 3.8 presents a statistical summary of the mean degree correlation values over the study period, comparing the original analysis and using a constant lag of two months. The results are presented as boxplots, where the median is indicated as a bold black line inside a box confined by the first and third quartiles (bottom and top of the box). The whiskers below and above the first and third quartiles show a threshold of 1.5 times the inter-quartile range (IQR), defining a number of outliers outside this range. As already noted, the Niger River basin showed a significant improvement in  $r_l$  after considering the delay, not only in terms of the median  $r_l$  value, but also in terms of the variability in the results (i.e. a smaller IQR). This outcome was similar irrespective of the evaporation product used. For the Colorado River basin, the degree correlations did improve when using the CSIRO-PML and GLEAM products (median improved from 0.17 to 0.67, and from -0.01 to 0.64, respectively) but to a lesser extent for the MOD16 product (-0.03 to 0.29). The IQR was also reduced significantly with the CSIRO-PML product, moderately reduced with GLEAM, and did not change with the MOD16 product. The degree correlation in the Aral Sea basin also benefited from an imposed lag in GRACE data, although there remained considerable variability in the results. The Lake Eyre basin showed only a marginal increase in the amplitude of  $r_l$  and a minor reduction in the temporal variability (-0.06 to 0.14, 0.08 to 0.20 and 0.13 to 0.29 with CSIRO-PML, GLEAM and MOD16 respectively).

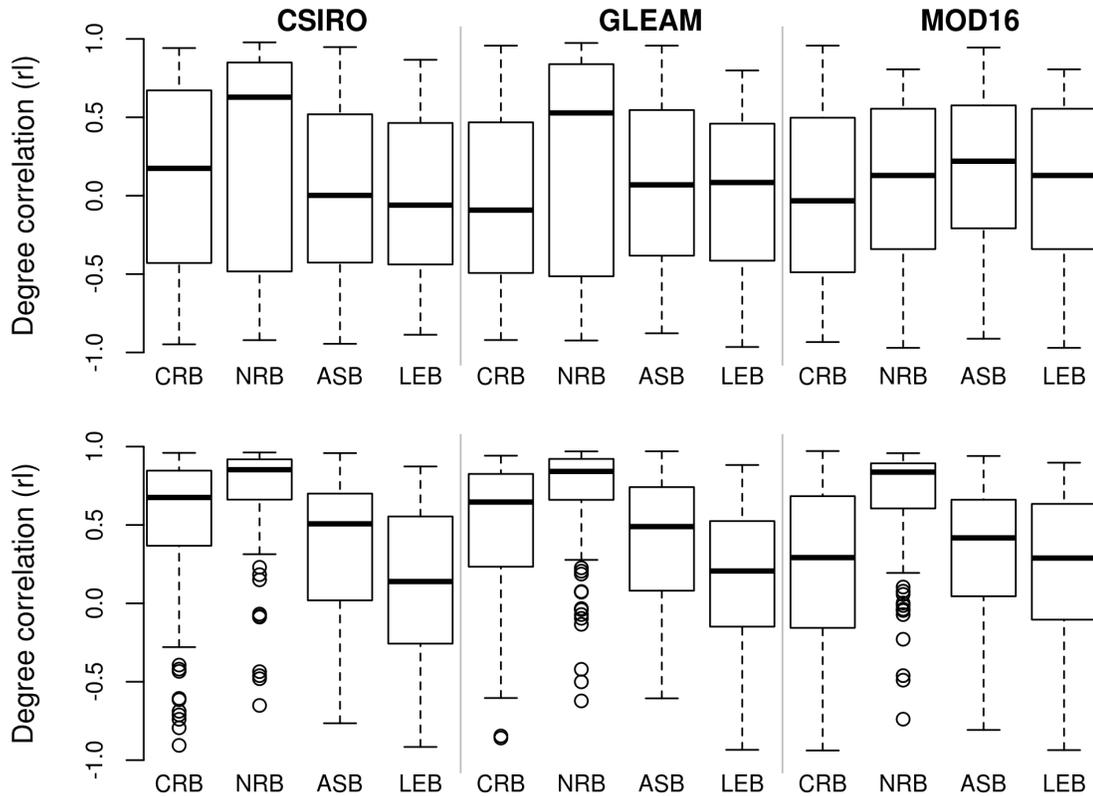


Figure 3.8: Top: average degree correlation statistics per study region and evaporation product. Bottom: GRACE data were shifted by two months to match the phase with  $P - E$  anomalies. The boxplots show the first, second (median) and third quartiles. Outliers, defined as data outside the 1.5 inter-quartile range (IQR) whiskers below or above the first and third quartiles are shown as circles.

### 3.6 Discussion

The development of methods and sensors to retrieve the various components of the water cycle has for the most part been undertaken independently of any evaluation against interrelated processes (see *McCabe et al. (2008)* and *Brocca et al. (2014)* for some examples of complementary retrieval). Large-scale retrievals of hydrological variables such as evaporation, soil moisture, and rainfall products do not come with well-defined accuracy metrics, let alone uncertainty bounds. This lack of any well-defined error structure associated with individual products complicates the task of product assessment. As such, the question of how to evaluate large-scale data sets re-

mains an outstanding one. This is especially important in the context of global-scale products. While a number of global evaporation (and precipitation) evaluation papers have been published, none seek to identify consistency with related hydrological variables, and focus instead on comparisons against traditional point-scale or tower-based techniques (*McCabe et al.*, 2016; *Michel et al.*, 2016). Given the spatial mismatch between ground observations (and the lack of continuous large-scale coverage of *in situ* data in remote regions), it is perhaps inappropriate to evaluate these large-scale products in such a manner. Determining whether individual products are at least consistent with each other (i.e. they reflect hydrological expectation) is a needed first step in product assessment. The motivation behind this study was to take a step back and determine whether a first-order hydrological assessment could be achieved. Rather than comparing the uncertainties between the evaporation products and the other hydrological components (which are poorly defined), we attempt to distinguish between the different evaporation products relative to their consistency with precipitation and terrestrial water storage. That is, are observed changes or patterns in the evaporation data sets reflected in these other hydrological variables? We explore this approach precisely because of the challenges in quantifying uncertainty based upon traditional *in situ* methods. As is discussed below, the challenge on how to do this remains, raising some important questions on both product quality and also the techniques we use to evaluate global products.

### **3.6.1 Challenges to implementing hydrological consistency**

For some regions, especially those where simpler and more defined water cycle behavior dominates, it is reasonable to expect that significant and consistent inter-product agreement between hydrological components should be observable. To explore this idea, our study focused on basins where such a simplified water budget, consisting of water storage anomalies as a function of precipitation and evaporative fluxes, might

be expected to predominate. The aim was to reduce the influence of complicating variables such as snow, vegetation changes, large precipitation and streamflow contributions, and other hydrological processes from the analysis. The assumption was that arid and semi-arid regions would best fit this profile. The role of the degree correlation was to evaluate the spatial agreement between the hydrological components, assuming that any non-closure errors due to unmodeled outflow components (e.g. long-term baseflow or minimal surface runoff) would not affect this measure. However, other sources of errors that directly affect evaporation estimates, such as the choice of algorithm, implicit model assumptions, choice of parameterizations, and an incorrect representation of the land cover, can directly impact the degree correlation measure. Given the relationship between size and retrieval accuracy as relates to GRACE data, obtaining a geographical distribution of basins that could satisfy this simplified water budget assumption was non-trivial. Restrictions related to basin size affect the study in two conflicting ways. On the one hand, a large basin will inevitably present complications related to heterogeneity (including in climate zones, as was the case for the Colorado River basin and the Aral Sea basin) and also be more likely to contain areas affected by anthropogenic activities, such as irrigation, land cover changes, and building of dams and reservoirs. On the other hand, a small catchment size is more difficult to evaluate with this consistency approach, given the coarse resolution of (most) of the global products used here, but especially the GRACE data. The spatial resolution of GRACE data is further limited by the use of filters to remove errors. Considering these restrictions, a compromise in the selection of study basins was required to allow for at least a narrow range of length scales (500-800 km) to be evaluated.

In the end, our study consisted of four major globally distributed river basins, including two endorheic systems. Although they mostly have an arid climate in terms of Köppen classification, both the Colorado River and Aral Sea basins include regions

with the presence of snow and snowmelt-dominated runoff. While snow storage itself is not a problem, since GRACE detects changes in storage irrespective of their nature (snow, groundwater, soil moisture, etc.), snowmelt may contribute to delayed changes in storage that can affect gravity results. Likewise, evaporation models generally have a difficult time adequately estimating sublimation. However, the inclusion of these basins was considered important in order to test the hydrological consistency concept in regions that deviated from the ideal assumption. Indeed, the influence that snowmelt and other potential sources of lag in the system have is poorly defined and forms part of the motivation to explore the inclusion of a lag response in the GRACE data (see section 3.5.3).

Apart from the issues of spatial scale, the use of satellite-based hydrological data presents additional challenges and sources of uncertainty to any consistency-based assessment. For instance, because GRACE data are smoothed to remove errors in small-scale terms (i.e. truncation of the spherical harmonic coefficients), the gravity signal contains contamination from outside of the studied basins (leakage) and represents a potential source of uncertainty in areas neighbouring high-amplitude signals (particularly if they are out of phase with the study basin) and the ocean. Although the LSM-based scaling factor, which is static in time, has been used to correct for bias (e.g. signal reduction) and leakage contamination, dynamic changes in water storage trends outside the basin might still contaminate the signal (*Long et al., 2015*). In addition, the temporal lag in terrestrial storage response, as documented in previous studies (*Rieser et al., 2010; Ahmed et al., 2011; Wang et al., 2014; Hassan and Jin, 2016*) and observed in our analysis, represents an important source of potential error (see sections 3.5.3 and 3.6.2). Product errors are also evident in the precipitation and evaporation data sets. Global rainfall retrievals have well-recognized limitations, including the detection of both high- and low-intensity events (*Hou et al., 2014*), the discrimination of cloud-free and cloud precipitation scenes, as well as the sensitiv-

ity to parameters in the forward model of radiative transfer over different sensors (*Stephens and Kummerow, 2007*). In terms of evaporation, uncertainties related to algorithm choice, input data variability, and process parameterizations all complicate the accurate estimation of terrestrial fluxes (*Ershadi et al., 2015*). *McCabe et al. (2016)* present a thorough description of accuracy issues related to global products. However, it is not the intention of this work to explore these product uncertainty issues in detail. Determining whether or not and understanding how much these sources of product uncertainty affect hydrological consistency studies are important areas requiring further investigation. What is clear from this analysis is that there is still some way to go in terms of being able to confidently assert that any single global product outperforms any other, at least in terms of its inter-product consistency.

### 3.6.2 Temporal lag in terrestrial storage response

In exploring the relationship between GRACE water storage changes and precipitation and evaporation data, it was evident that water storage anomalies peaked at a significantly later time than the corresponding  $P - E$  values. One possibility for this apparent lag in GRACE data is that, due to the inability of GRACE to detect small-scale changes in the gravity field (a rough estimate of GRACE accuracy averaged over the entire Earth is  $20 \text{ mm.month}^{-1}$ ; *Wahr et al. (2006)*), the corresponding mass is not detected until a sufficient amount has accumulated within the catchment via natural drainage processes (*Rieser et al., 2010; Ahmed et al., 2011*). The intensity and duration of the precipitation events, antecedent soil moisture condition, as well as the hydrogeological and geomorphological characteristics of the basin all influence the accumulation and detection time. A simple way to account for this phenomenon was to apply a constant phase lag to GRACE data across the whole study period. Doing this improved the behavior of degree correlation, not only in time (less variability in the results), but also increased the value of  $r_l$  as well. This was particularly

evident in the Niger River basin, which was expected due to the well-defined seasonal behavior of its hydrological cycle throughout the study period, and to a lesser extent in the Colorado River basin and Aral Sea basin, where changing trends in the seasonal patterns of precipitation made it more challenging to apply this simple correction. In the Lake Eyre basin, applying a lag to GRACE data did not seem to have an effect on the degree correlation. Clearly, this is a simple approach to accounting for delayed flow contributions to the catchment, and further understanding the implications and physical rationale behind the attribution of this lag is required.

### 3.6.3 Discriminating between satellite evaporation products

One motivating aspect of this work was to explore whether differences in available global evaporation products (i.e. satellite-based evaporation models forced with global input data) impacted the results of the consistency analysis i.e. could we identify better agreement between water storage anomalies and  $P - E$  in any particular evaporation product? While the analysed products covered a wide range of resolutions ( $0.05^\circ$ - $0.25^\circ$ ), the effective resolution in the analysis was ultimately determined by the truncation degree ( $l_{max}$ ) of the spherical harmonic transformation. Even after accounting for this, absolute differences were evident from a qualitative basin-scale analysis (figure 3.2). Results indicated that MOD16 underestimated evaporation when compared to CSIRO-PML and GLEAM, even though both the CSIRO-PML and MOD16 products are based on the Penman-Monteith equation. Several recent studies (*McCabe et al.*, 2016; *Michel et al.*, 2016; *Miralles et al.*, 2016) also suggest that the MOD16 product (or variants using the PM-Mu approach) underestimate evaporation when compared to other products (including GLEAM), and that most products show large discrepancies in reproducing results during periods of water stress. *Ershadi et al.* (2015) demonstrated that the parameterization of aerodynamic and surface resistances were critical controls on evaporation through both soil and vegetation.

Furthermore, both GLEAM and CSIRO-PML include dynamic constraints on evaporation (stress module and soil moisture assimilation in GLEAM; dynamic ratio of actual to potential evaporation at the soil surface in CSIRO-PML) that are critical in arid regions due to hydrological and plant physiological stresses and the subsequent importance of soil evaporation. Whether these differences in model parameterization are the sole cause of the apparent underestimation by MOD16 remains to be investigated. However, these differences in absolute values did not affect the overall results of degree correlation dramatically, i.e. we could not identify a significant advantage or disadvantage in terms of hydrological consistency.

As the focus of the study was to discriminate between evaporation products, the question of whether the choice of precipitation product affected the hydrological consistency analysis was somewhat beyond the scope of this work. However, a preliminary analysis was undertaken by replacing the GPCP precipitation data with another data set and reproducing the analysis. To do this, we processed the Precipitation Estimation from Remote Sensing Information using Artificial Neural Network (PERSIANN) product, which uses an artificial neural network to approximate spatiotemporal non-linear relationships between physical variables and remotely sensed signals (*in Hsu et al., 1997*). PERSIANN uses data from the long-wave infrared imager onboard the Geostationary Operational Environmental Satellite (GOES) as well as from the Tropical Measurement Mission (TRMM) microwave imager (TMI). The results of this new analysis did not reveal any significant difference when compared to those based on the GPCP analysis.

Evaluating global evaporation products remains an outstanding challenge. The purpose of implementing a hydrological consistency approach was to explore the evaluation of evaporative fluxes by comparing the spatial patterns between precipitation and changes in water storage. If such an approach could be shown to perform well in a relatively simple hydrological system, the potential for broader-scale application

in regions with more complex behavior would be the next logical step. However, the study showed that even in these relatively simple basins, it was not possible to demonstrate a consistent hydrological agreement between independent observations. Improvements in satellite-based evaporation products are likely to be delivered through advances in algorithm development, increases in the observable resolution, and also via the development of multi-product ensembles (with weighting based on validation analyses and uncertainty assessments). The prospects for improved precipitation monitoring are also promising given the Global Precipitation Measurement mission, which will allow for a more accurate representation of light rains, a challenge that has been a limitation in other precipitation products, including the GPCP (*Huffman et al.*, 2001). Likewise, the next generation gravity missions (GRACE follow on and GRACE II) with the incorporation of improved sensor design (*Christophe et al.*, 2015) are anticipated to provide more accurate estimates of the water storage anomalies, albeit with no significant increase in resolution.

### 3.7 Conclusions

Given the inherent challenges in validating satellite-based products via the use of ground-based observations, we examined the capacity of independent observations of the water cycle to reflect hydrological consistency. The study focused on regions where it would be most expected to observe such responses: arid and semi-arid regions with a simplified water budget, consisting primarily of precipitation and evaporation, and assuming a minimal runoff and other long-term outflow components. It was determined that, even in these simple environments, hydrological consistency was difficult to obtain. While there are times and locations at which some consistency was observed, there were a greater number for when it was not. The lack of any persistent behavior is problematic, both in the attempt at independently evaluating remote sensing data and also in any effort to discriminate between individual prod-

ucts. The lack of persistent agreement in some of the studied basins may be explained in part by the added complexities that limit the validity of the assumption of a simple water cycle, i.e. snowmelt runoff, complex geomorphology or hydrogeology, changing patterns of precipitation, as well as anthropogenic influences on the water system.

Despite these challenges, the expectation is that retrievals of global and regional products will inevitably improve with advances in resolution, process understanding, and forcing data accuracy. In concert with such product improvements, the way in which we evaluate remotely sensed variables should also evolve beyond comparisons against *in situ* data that form the basis of most current assessments. Such a strategy would include evaluation against related hydrological variables, reflecting the underlying rationale of hydrological consistency and hydrological closure studies. Only by implementing a more comprehensive evaluation framework in our assessment schemes will greater confidence in component retrievals be realized.

The tradeoff in resolution, model assumptions and robustness of global evaporation products limits the ability to include some components of evaporation. For example, changes to land cover and extensive use of water in largely developed agricultural regions might not be captured by these large scale products, partly explaining the lack of agreement with observed water storage changes. Currently, efforts to capture these anthropogenic influences, at least within the limits of the retrieval models, require the use of hyper-resolution data as well as larger computation requirements. By identifying and focusing on small regions (e.g. figure 1.2) with intense agricultural footprints, it will be feasible to apply these models and capture this component of evaporation. This will be the topic of chapter 4, where an examination of methods to obtain high-resolution data from remote sensing for use in a relatively more complex evaporation model will be explored. Then, the use of a land surface model will allow us to retrieve not only evaporation, but also other important fluxes including drainage into the subsurface, providing more insight from a water management perspective.

## Chapter 4

### High resolution modeling of water fluxes in an extreme arid environment

---

This chapter describes the need for (and challenges of) high-resolution modeling, within the context of water cycle modeling as well as crop evaporation retrieval. The availability of numerous satellites monitoring the Earth at increasingly higher resolutions enables the characterization of a range of land surface parameters needed for higher resolution modeling. Coupling this data with output from a numerical weather prediction model, a relatively simple energy balance model can be used to retrieve evaporation from both soil and vegetation. Using the same data to feed a land surface model, a simulation of the hydrologic fluxes across a range of spatial and temporal scales can be performed. Both the energy budget evaporation model and the land surface model are described in this chapter, along with the techniques and data required to force the models and correctly characterize the land surface within the study domain. While the high-resolution remote sensing based approach clearly captures the hydrological impact of agricultural activities in the study region, simulations of the land surface model using meteorological data (i.e. without additional irrigation input) are unable to capture this signal, highlighting a key limitation of the approach and setting the stage for a new model-data fusion framework in Chapter 5.

---

## 4.1 Introduction

Although there have been significant developments in hydrologic models and a much wider availability of observational data needed to drive them, there remains a need to increase the resolution of land surface models to accurately represent the water, energy and carbon cycles with local relevance, i.e. the concept of “everywhere and locally relevant” (*Bierkens et al.*, 2015). Large-scale models might not capture local variations in small catchments, fluxes from sparse irrigation (see section 3.7), or be able to accurately predict and monitor flood risks. *Wood et al.* (2011) synthesized six major challenges that need to be addressed to develop the next generation of global “hyper-resolution” models (100 *m* - 1 *km*). The first challenge is the need for a better representation of surface and subsurface interactions, while accounting for effects of local topography (e.g. flooded hill slopes, effects of slopes on reflected radiation). Second, interactions between land and atmosphere are currently based on a one-dimensional vertical transfer, but will need to account for heterogeneous surfaces at smaller resolutions. Third, linking the water cycle to the carbon and nitrogen cycles, as well as including biogeochemical processes to model the global distribution of water quality. Fourth, impacts of human-related activities including reservoirs, land cover changes and other environmental man-made modifications. In order to include all these sources of impact on the water cycle, more and more data will need to be collected globally, which is already an ongoing effort driven in large part by the current era of satellite remote sensing (RS). The fifth challenge will be to manage (store, process and assimilate) this growing input of data at higher and higher resolutions. Finally, in order to process all these data and the inclusion of these additional complexities, global hyper-resolution models will need a large number of computing resources: an ongoing effort in terms of managing massively parallel computations.

While there is still a considerable amount of work needed to overcome many of

these issues, a number of models have been used at increasingly higher resolutions (e.g. 5' degree cells) compared to what has been traditionally used ( $>0.1$  degrees) to explore global water consumption (*Flörke et al.*, 2013), global water balances (*van Beek et al.*, 2011), continental-scale flood risks (*van Beek et al.*, 2011; *Rojas et al.*, 2012) and terrestrial carbon cycles (*Haverd et al.*, 2013). Recent smaller-scale studies at high-resolutions have also provided insight into the impacts of urban landscape in ET and LST studies (*Vahmani and Hogue*, 2014), and aided in the calibration and validation of land surface models by directly incorporating model resolution commensurate with available remote sensing data (*Vahmani and Hogue*, 2014).

In the context of advancing land surface models towards the goal of hyper-resolution, *Singh et al.* (2015) ran multi-year continental-scale simulations using the Community Land Model at resolutions of 1, 25 and 100 *km* using high-resolution (100 *m* - 1 *km*) topography and soil texture data, and showed significant improvements in model output as a result of the high-resolution surface data. This and similar works draw attention to the need for parallel efforts aimed at developing higher-resolution land surface data products, making use of new data sources and processing techniques (*Houborg and McCabe*, 2016). *Singh et al.* (2015) also showed some of the limitations at hyper-resolutions (100 *m* and lower), resulting from the use of embedded model physics that were designed for larger scales, e.g. a lack of consideration of lateral flow, or the effects of slope and aspect on runoff, infiltration, drainage, and groundwater storage *Singh et al.* (2015). This “scale separation and scale-related breakdown of concepts and assumptions” (*Bierkens et al.*, 2015) also include other aspects of the model (i.e. other than surface characteristics) such as the neglect of horizontal advection for land-atmosphere heat and moisture exchange, the connections between compartments of hydrological systems (e.g. water stress calculations), among others.

Using an ensemble average of the same model with different parameterizations (i.e. multi-physics modeling) could potentially improve LSM estimates, although necessar-

ily increasing computation costs even more (*Niu et al.*, 2011; *Yang et al.*, 2011; *Cai et al.*, 2014). Another approach, which will require the efficient use of growing computational resources, is the concept of integrated hydrologic models (or distributed hydrologic models) at larger domains. For example, PARFLOW (*Kollet et al.*, 2010; *Kollet and Maxwell*, 2006) was designed and developed with parallelism and scalability in mind, to include small-scale physically-based processes at increasing scales (in the order of  $1 \times 10^3 \text{ km}^2$  with resolutions of  $1 \times 10^1 - 1 \times 10^2 \text{ m}$ ). These models offer a more realistic representation of the interactions between surface and subsurface, such as an overland flow equation used as an upper boundary condition, path-based description of surface flow across drainage basins and consideration of groundwater storage and fluxes (*Camporese et al.*, 2010), variable-density flow and transport (including reactive chemical species) (*Brunner and Simmons*, 2012), and better representation of topographic effects (*Maxwell*, 2013). However, the application of these type of models requires an active coupling of the land-subsurface components, demand a rethink on model initializations (*Ajami et al.*, 2015), and require serious computational infrastructure, of the order of tens of thousands of processors (*Kollet et al.*, 2010).

Since the groundwater systems in the Arabian Peninsula are generally hundreds to thousands of meters below the surface (*Sultan et al.*, 2014), they are effectively decoupled from the surface, so the application of coupled surface water-groundwater models is not appropriate. For this reason, we examine the use of a more standard land surface model (LSM) coupled with high-resolution (30 m) satellite-based data (section 4.5) and meteorological data derived from a state-of-the-art numerical weather prediction model (section 4.4), to describe the water and energy interactions in our study region. In this study, we use Landsat data to derive land and vegetation parameters (albedo, leaf area index, land surface temperature) at a sufficient resolution to capture the evaporation from agricultural fields (section 4.7). These same

inputs will be used to drive a simpler RS-based evaporation model (Section 4.2) with the goal of comparing the evaporation output of both models.

## 4.2 High-resolution RS-based evaporation retrieval

A number of remote sensing-based large-scale evaporation products were evaluated in chapter 3. These products were developed with a global perspective in mind, for which the goals were simplicity, robustness, and enough accuracy to predict large-scale trends of evaporation in different environmental conditions around the globe. With the balance of these objectives in mind, the number of input data requirements for these models is small. For example, the global evaporation MOD16 product (section 3.2) and other Penman-Monteith-based evaporation products do not require the radiometric surface temperature ( $T_R$ ) nor wind speed data as input (*Cleugh et al.*, 2007), as other evaporation models do (e.g. thermal-based  $E$  models). Given the relatively small regional focus (i.e. single ( $< 1 \text{ km}^2$ ) fields within a desert environment; figure 1.2) of the present study, the considerations related to model complexity (and number of inputs) can be relaxed in favor of a more realistic model, that can potentially incorporate the footprint of agricultural water use. One of these models, the two source energy balance model (TSEB) makes use of  $T_R$  and shares some of the meteorological inputs with the LSM as well (section 4.3). The use of  $T_R$  (at high enough resolution) enables TSEB to potentially retrieve the evaporation signal from these irrigated fields.

### 4.2.1 Description of the Two Source Energy Balance model (TSEB)

*Norman et al.* (1995) developed a two-source energy balance model (TSEB), which partitions evaporation into soil and vegetation components. It is based on the energy balance concept, consisting of net radiation ( $R_N$ ), ground heat flux ( $G$ ), and the

turbulent fluxes (i.e. sensible  $H$  and latent  $LE$  heat fluxes).  $LE$  is obtained as a residual of the energy balance for both soil (S) and vegetation (C):

$$R_N - G = H + LE \quad (4.1)$$

$$LE_C = R_{N,C} - H_C \quad (4.2)$$

$$LE_S = R_{N,S} - H_S - G \quad (4.3)$$

The heat transport to the ground (G) was originally modeled as a constant portion of the radiation flux to the soil, but this was later refined to account for the difference in phase between the daytime ground heat flux and the radiation that reaches the soil (*Santanello Jr. and Friedl, 2003*). The sensible heat flux is modeled based on the idea of a network of temperature gradient-transport “resistances” (figure 4.1) between the air (A), the canopy boundary layer (AC), the canopy (C), and the soil (S):

$$H_C = \rho C_p \frac{T_C - T_{AC}}{r_x} \quad (4.4)$$

$$H_S = \rho C_p \frac{T_S - T_{AC}}{r_S} \quad (4.5)$$

$$H = H_C + H_S = \rho C_p \frac{T_{AC} - T_A}{r_A} \quad (4.6)$$

Because of a lack of separate measurements for the temperature of the canopy ( $T_C$ ) and soil ( $T_S$ ), the radiometric surface temperature ( $T_R$ ) is used to indirectly infer the soil temperature assuming a linear relation where  $T_R$  is a composite of  $T_C$  and  $T_S$ :

$$T_S^4 = (T_R^4 - f_{VR}T_C^4)/(1 - f_{VR}) \quad (4.7)$$

where  $f_{VR}$  is the vegetation fraction within the field of view (zenith angle  $\phi$ ) of the

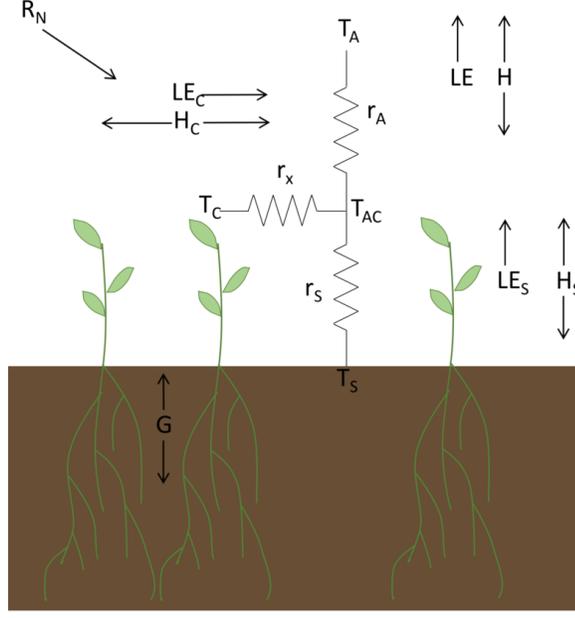


Figure 4.1: Conceptual model of TSEB including the network of resistance (adopted from *Norman et al. (1995)* and *Colaizzi et al. (2012)*)

radiometer (*Norman et al., 1995*):

$$f_{VR} = 1 - \exp\left(\frac{-0.5LAI}{\cos\phi}\right) \quad (4.8)$$

In this study,  $T_R$  is obtained from atmospherically corrected land surface temperature (LST) based on Landsat data (*Rosas et al., 2017*). Initially,  $T_C$  is given a value based on non-water stressed conditions using either a Penman-Monteith (PM) approach (*Colaizzi et al., 2012*) or the Priestley-Taylor (PT) equation (*Kustas and Norman, 1999; Norman et al., 1995*), and  $T_S$  is then obtained from equation (4.7). In this study, we use the Priestley-Taylor approach with an adjustment to the original PT coefficient  $\alpha_{PT}$  (increase from 1.3 to 2) for vegetation transpiration, as in *Kustas and Norman (1999)*. The net radiation ( $R_N$ ) was estimated as a function of shortwave radiation ( $S_d$ ,  $W.m^{-2}$ ) and air temperature ( $T_a$ ) as in *Zhuang and Wu (2015)*:

$$R_N = (1 - \alpha)S_d + \epsilon_\alpha\sigma T_\alpha^4 - \epsilon\sigma T_{rad}^4 \quad (4.9)$$

where  $\epsilon_\alpha$  is the emissivity of the atmosphere and  $\epsilon$  is the emissivity of the land surface.  $\epsilon_\alpha$  is a function of air temperature and water vapor pressure, while  $\epsilon$  is composed of the emissivity of the soil ( $\epsilon_s$ ) and canopy ( $\epsilon_c$ ) via the vegetation fraction:

$$\epsilon = f_c \epsilon_c + (1 - f_c) \epsilon_s \quad (4.10)$$

A more detailed description of TSEB (including the calculation of resistances) can be found in *Anderson et al. (2008)*, *Colaizzi et al. (2012)*, *Norman et al. (1995)* and *Kustas and Norman (1999)*. Several studies have shown an advantage of TSEB compared with other energy balance residual models (*Gao and Long, 2008*; *Gonzalez-Dugo et al., 2009*; *Tang et al., 2011*), as well as with PM and PT approaches (*Sun et al., 2014*). However, depending on vegetation type and scale, other approaches have also illustrated an advantage over TSEB (*Yang et al., 2015*; *Hoffmann et al., 2016*).

### 4.3 The Community Atmosphere Biosphere Land Exchange (CABLE) model

The Community Atmosphere Biosphere Land Exchange (CABLE) model, developed at the Commonwealth Scientific and Industrial Research Organization (CSIRO) in Australia, is a Land Surface Model (LSM) designed to calculate the temporal evolution of energy, water, momentum and carbon fluxes across the land atmosphere interface and across different land covers (i.e. bare ground, different types of vegetation and snow) (*Kowalczyk et al., 2006*). LSMs have evolved from simple “bucket-type” schemes (e.g. with a single soil type and no vegetation) to more complex representations of the land-atmosphere interaction, including vegetation components and more recently biogeochemical processes such as carbon production (i.e. biosphere-atmosphere interactions). These processes are important, since land surface processes

may affect weather and climate at different spatial and temporal scales (*Zhang et al.*, 2011). CABLE, as with many other land surface models, can be used “offline” at single sites, applied at regional and even global scales, or incorporated directly into Global Climate Models (GCMs) and Earth System Models. For example, CABLE constitutes the land surface model component in the Australian Community Climate Earth System Simulator (ACCESS) (*Kowalczyk*, 2013), while the Catchment Land Model (*Koster et al.*, 2000) forms part of the Goddard Earth Observing System, Version 5 (GEOS-5) Atmospheric General Circulation Model (AGCM) (*Molod et al.*, 2015).

The development of CABLE somewhat mirrors the development of land surface models in general. Climate and weather forecast models used land surface schemes in order to set the lower boundary conditions for the atmosphere. The land surface scheme in CSIRO’s General Circulation Model (GCM) in 1990 consisted of a soil moisture scheme (*Deardorff*, 1977) and a force-restore method (*Deardorff*, 1978) to calculate the surface temperature, although with no vegetation. A few years later, a stand-alone model for soil and canopy (single “big” leaf, i.e. a large leaf acting as a source or sink of water vapor and sensible heat) and a force-restore model for soil were added to the land surface scheme, along with different soil types, soil percolation, vegetation features such as canopy resistance and interception of rainfall, as well as other surface characteristics (e.g. albedo and roughness length). For the canopy temperature, a stomatal resistance (modeled as a function of radiation, saturation deficit, temperature and water stress) was calculated from the surface energy balance. Subsequent improvements during the following years (up to 1997) included the partition of the soil into six layers as well as the addition of a snow model with three snowpack layers. A stand-alone offline version of the land surface scheme was developed (*Raupach et al.*, 1997) and included a canopy layer above the soil surface, the parameterization of an aerodynamic conductance that accounted for turbulent

exchanges between the soil, vegetation and the atmosphere, as well as radiation, saturation deficit, temperature and water stress effects on the canopy stomata. A comparison of two canopy models: a one layer two-leaf model (sunlit and shaded leaves) and a multilayer model, showed that both yielded similar fluxes (*Wang and Leuning, 1998*).

The first release of CABLE in 2003 included a canopy turbulence model, multi-layer soil and snow models, a simple carbon pool model, and an improved version of the two-leaf canopy model selected for its higher computational efficiency (*Kowalczyk et al., 2006*), which allowed for the effects of soil water deficit on photosynthesis and respiration, among other improvements (*Wang, 2000*). Later major improvements included the modification of the root water uptake function and the addition of a hydraulic redistribution function, which helped improve soil water dynamics and latent heat flux estimation during dry conditions (*Li et al., 2012*). CABLE consists of five modular components (*Wang et al., 2011*): 1) radiation, 2) canopy micrometeorology, 3) surface flux, 4) soil and snow and 5) ecosystem respiration. A detailed description of each of five CABLE modules is available in *Kowalczyk et al. (2006)*, with subsequent improvements detailed in *Wang et al. (2011)*, *Kowalczyk (2013)* and *Haverd et al. (2016)*.

The main calculation in CABLE is the fluxes of water, momentum and heat, which are modeled through the use of aerodynamic resistances. The general form of sensible and heat fluxes is:

$$H/\rho_a c_p = -u_* T_* = (T_{sur} - T_{ref})/r_H \quad (4.11)$$

$$E/\rho_a = -u_* q_* = (q_{sur} - q_{ref})/r_E \quad (4.12)$$

where  $T$  and  $q$  are air temperature and specific humidity at the reference level (subscript  $_{ref}$ ) and surface values (subscript  $_{sur}$ ),  $u_*$  is the friction velocity (subscript  $_*$

represents turbulent scale),  $r$  represent resistances (H for heat and E for water exchange between surface and reference level, the latter including aerodynamic and plant stomatal resistance). Finally,  $\rho_a$  is air density and  $c_p$  is specific heat.

The energy balance equation (4.1) for the surface (consisting of a combination of bare ground, vegetation and snow/ice) is solved in combination with equations (4.11) and (4.12) to obtain the surface temperature  $T_{sur}$ . Both sensible and latent heat fluxes are partitioned into a flux from the soil to the canopy, and from the canopy to the atmosphere.

In the parameterization of aerodynamic resistances, the friction velocity ( $u_*$ ), calculated using Monin-Obukhov similarity theory (*A.S. and A.M.*, 1954), depends on the thermal stability parameter ( $\xi$ ). Based on the model for canopy aerodynamics of *Raupach et al.* (1997),  $\xi$  depends on the total grid heat fluxes, and thus on surface temperature as well. Given that the surface temperature must be solved for in the energy balance equation (4.1), and that there is a strong interdependence between the thermal stability parameter  $\xi$  and the surface temperature, these main calculations must be done iteratively. In CABLE, four iterations are allowed to obtain the values of  $\xi$ , surface fluxes and canopy temperature (*Kowalczyk et al.*, 2006). The main flow diagram of CABLE, simplified from a more complete diagram in *Kowalczyk et al.* (2006), is shown in figure 4.2, where the key iterative procedure for the calculation of the stability parameter is evident. After this iterative procedure, where mainly diagnostic variables are calculated (i.e. that depend only on the current time step), the calculation of prognostic variables (i.e. those that are linked to future time steps) is undertaken, including the carbon flux, soil moisture and temperature.

The soil profile in CABLE is divided into six layers of increasing thickness downwards, with model levels prescribed at 2.2, 5.8, 15.4, 40.9, 108.5, and 287.2 *cm*. The soil module computes the soil moisture and heat fluxes between the layers. Vertical soil moisture is considered in terms of liquid and ice components (*Wang et al.*, 2011)

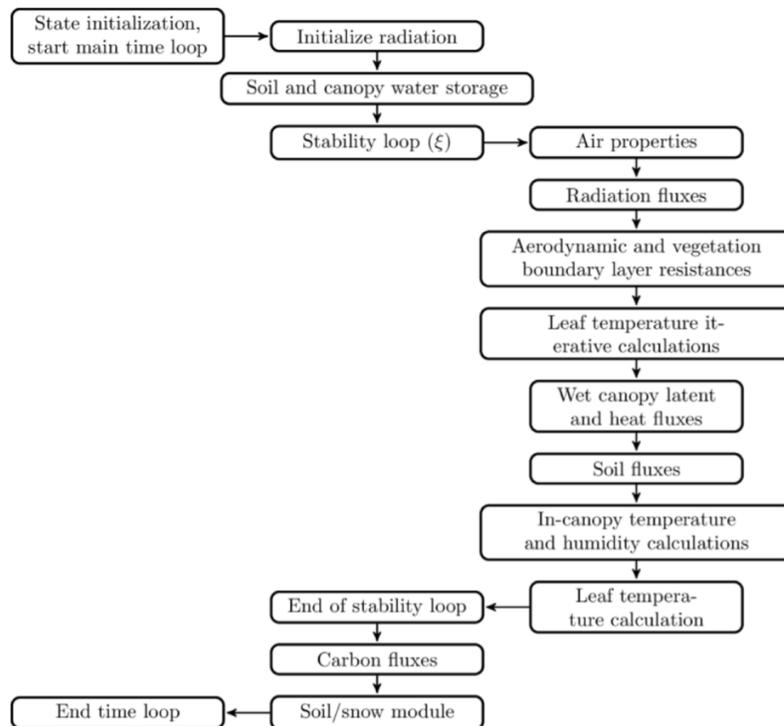


Figure 4.2: Flow diagram of CABLE model showing the importance of the stability parameter iterative calculations. A more complete diagram can be found in *Kowalczyk et al.* (2006)

and CABLE includes three snow layers above the ground. The soil moisture flux is calculated based on the one-dimensional conservation equation and Darcy's Law, and are combined to form the Richards' equation:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial}{\partial z} \left( K + D \frac{\partial \theta}{\partial z} \right) + F_w(z) \quad (4.13)$$

where  $\theta$  is the volumetric soil moisture content ( $V_o/V_o$ ),  $q$  is the kinematic soil moisture flux (defined as positive downward), the term  $F_w$  contains the water uptake by vegetation from depth  $z$ ,  $K$  is the hydraulic conductivity and  $D$  is the diffusivity. Water infiltration from the top layer is controlled by rainfall, evaporation, runoff and snowmelt, while the boundary condition at the bottom is given by:

$$q = c_{\text{drain}} \theta \quad (4.14)$$

where  $c_{drain}$  is the soil drainage coefficient. The representation of water uptake from plants (i.e. the term  $F_w$ ) was improved by incorporating an alternative root water uptake and hydraulic redistribution (*Li et al.*, 2012). These improvements help model how soil moisture is redistributed by vegetation during the night under dry conditions, when moisture is driven upwards from the deep moist layers.

In recent versions of CABLE, canopy transpiration and photosynthesis processes are coupled through the stomatal conductance:

$$G_s = f_{w,soil} \left( G_0 + \frac{a_1 A_c}{(C_s - \Gamma^*)(1 + D_s/D_0)} \right) \quad (4.15)$$

where  $A_c$  is net photosynthesis,  $D_s$  and  $C_s$  are water vapor pressure deficit and CO2 concentration at the leaf surface,  $\Gamma^*$  is a function of canopy temperature,  $a_1$  is a model parameter that depends on type of crop (4 for C4 plants and 9 for C3 plants),  $D_0$  is another model tuning parameter ( $D_0 = 1500 Pa$ ),  $G_0$  is residual conductance and  $f_{w,soil}$  is the stomatal conductance drought response factor (*Haverd et al.*, 2016). When the coupled equations for net photosynthesis and energy balance are solved, an initial value of the transpiration flux ( $q_{trans,0}$ ) is obtained. This value is then adjusted based on water availability in the soil and the surface energy balance is calculated using this adjusted transpiration. However, the net photosynthesis is not adjusted based on the actual transpiration. When the demand for root water extraction exceeds availability of water in the soil, this leads to an incorrect calculation of photosynthesis in the absence of extractable water. To correct for this inconsistency, (*Haverd et al.*, 2016) introduced a coupling of the stomatal conductance drought response and the root water extraction, which leads to a “root shut-down” by testing for over-extraction in each of the soil layers (i.e. setting  $f_{w,soil}$  to 0).

## 4.4 Meteorological inputs

In order to run both the LSM and RS-E model, CABLE needs to be forced with the following surface meteorological variables: shortwave and longwave radiation, surface air temperature, air humidity, surface pressure, wind speed and rainfall. Given the lack of a sufficient number of meteorological stations in the region (or reliable global forcing data at high resolution), and the expense that installing these in all agricultural regions would represent, we used meteorological data from a numerical weather prediction simulation. The meteorological data used in this study is based upon the Weather Research Forecasting (WRF) model: specifically, the Advanced Research WRF (ARW, *Skamarock et al. (2008)*) model version 3.7.0 developed by the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR). The simulation was performed over two, two-way nested domains with resolutions of 9 and 3 *km* and 40 vertical levels covering the Arabian Peninsula and neighboring regions (figure 4.3). An analysis of this forcing dataset has featured in several studies focusing on the climatology of the Red Sea (*Yesubabu et al., 2017; Langodan et al., 2017*), the potential for wind and wave energy of the Red Sea (*Langodan et al., 2014, 2016, 2017*) and extreme rainfall events in Saudi Arabia (*Yesubabu et al., 2016*).

The data from these WRF runs span the years 2011 to 2015 and were produced at an hourly time resolution. Simulations were performed every 36 hours following a cyclic 3DVAR assimilation approach, assimilating available observations every six hours. Assimilated data included quality controlled observational data (Prep-BUFR; <http://rda.ucar.edu/datasets/ds337.0>) from the NCEP Atmospheric Data Project (ADP) featuring: conventional observations from surface stations (synoptic stations, Metar, Ship, and Buoy), upper-air soundings (Rawinsonde and Pilot balloon), and satellite observations such as wind vectors from the Quick Scatterometer (QSCAT), Windsat and ASCAT scatterometers, and atmospheric motion vectors

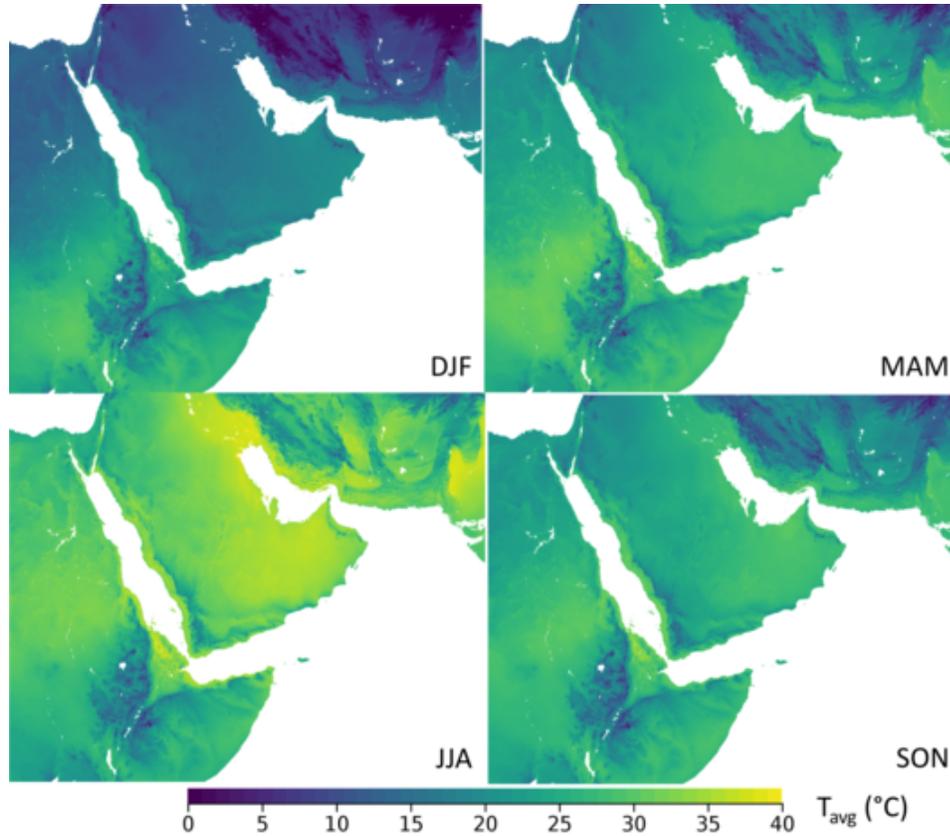


Figure 4.3: Seasonality of daily average temperature over the 2011-2015 WRF simulation period within the interior (3 km resolution) domain consisting of 1522 x 1210 grid points covering the Arabian Peninsula and neighboring regions.

from geostationary satellites. Further details are provided in *Langodan et al.* (2014) and *Yesubabu et al.* (2016).

The WRF model physics are described in *Langodan et al.* (2014) and are the same as in *Jiang et al.* (2009). The model physics features include: Yonsei University's non-local diffusion scheme (*Hong et al.*, 2006) for boundary layer processes, the Kain-Fritsch scheme for cumulus convection, the WRF Single-Moment 3-class (WSM3) for microphysical processes, the Noah land surface scheme (*Chen and Dudhia*, 2001) for surface processes, Rapid Radiation Transfer Model (RRTM) for long-wave radiation (*Mlawer et al.*, 1997), and the *Dudhia* (1989) scheme for short-wave radiation. Elevation data, soil type and other topographic data for the model were interpolated from USGS data.

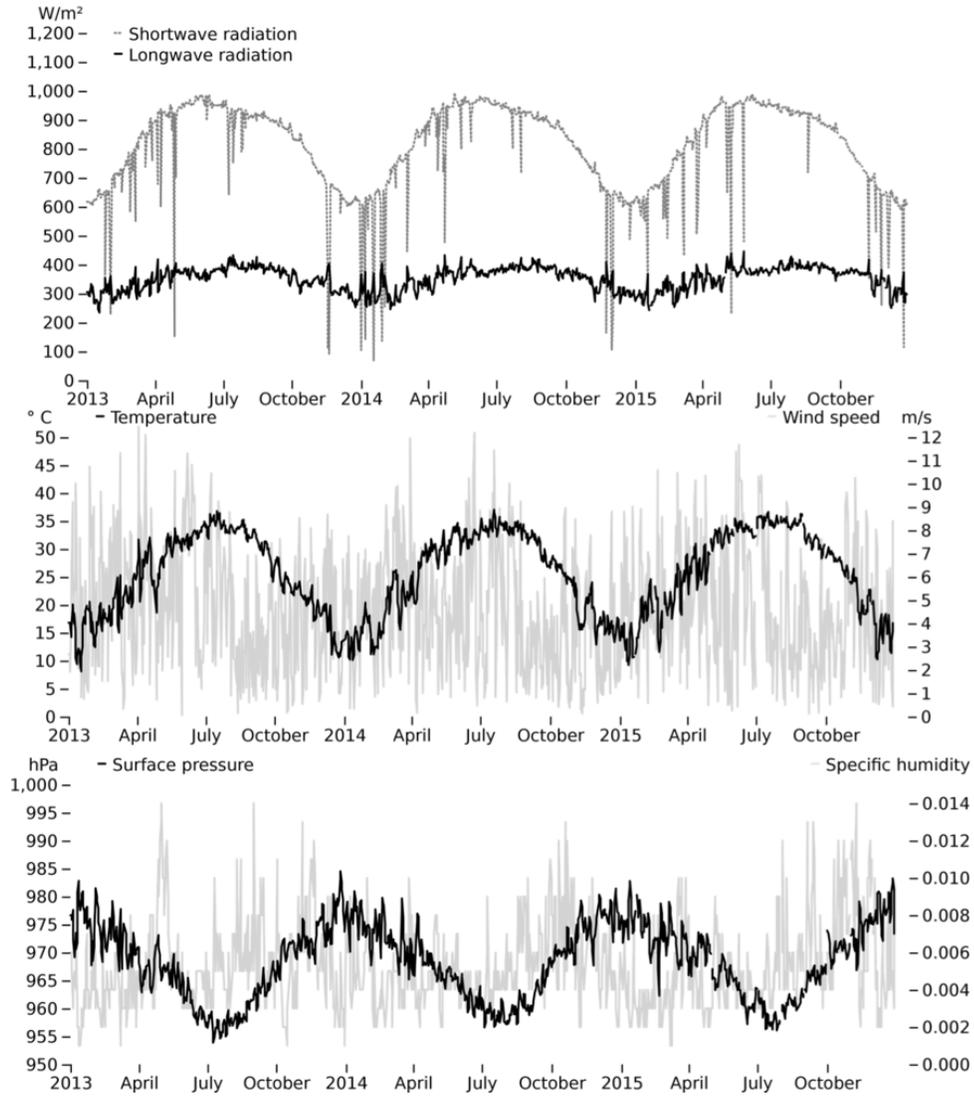


Figure 4.4: 10 AM (local time) snapshots of WRF variables relevant for CABLE extracted from the location of the Tawdeehiya Farm for the period 2013 - 2015.

Figure 4.4 shows a sample seasonal time series of the relevant variables for CABLE from the WRF simulation for the years 2013 - 2015. To enhance the visualization of the long-term trends, data from the 10AM (local time) hourly step were extracted for temperature, while the average daily value was computed as the mean of the maximum and minimum values for each day.

#### 4.4.1 CABLE meteorological input module modification

CABLE offline runs are typically done for single sites, with a single meteorological input file containing the required variables for that site only in netCDF format. While this is suitable for single sites, it is not a scalable approach. For the Arabian Peninsula domain, there are multiple agricultural sites of variable sizes, some covering a number of (3 km) pixels of the WRF data. Creating individual netCDF files for CABLE for each of these sites would be an unnecessary duplication of effort and storage. For this reason, the CABLE input module was modified to receive a static look-up table and programmatically extract only the required data for each of the fields within the Tawdeehiya domain (figure 1.2). This will enable an easy transition (and scalability) to other subdomains within the Arabian Peninsula (figure 2.7), independent of the WRF domain and/or resolution.

#### 4.4.2 Evaluating the WRF-predicted rainfall

The spatial distribution of precipitation within the Arabian Peninsula domain predicted by the WRF simulation (figure 4.5) is consistent with recent studies comparing the performance of rain gauges and global climate models over the region (*El Kenawy et al.*, 2014; *El Kenawy and McCabe*, 2016), with highest rainfall in the Peninsula occurring in the southwestern mountainous region, and minimal precipitation elsewhere. However, these studies also emphasize the poor performance of global climate models in reproducing long-term trends over the region, while gauge-based products exhibited significant inter-product variability (*El Kenawy and McCabe*, 2016).

In the study region (Tawdeehiya Farm; figure 1.2), annual precipitation rarely exceeds 150 mm (figure 4.6), a value that is a small fraction in comparison with typical irrigation values (on the order of  $1 \times 10^3 \text{ mm.yr}^{-1}$ ). Importantly, much of this rainfall falls in only a few extreme events, rather than being evenly distributed throughout the year. It is therefore expected that using solely the WRF predicted rainfall, the

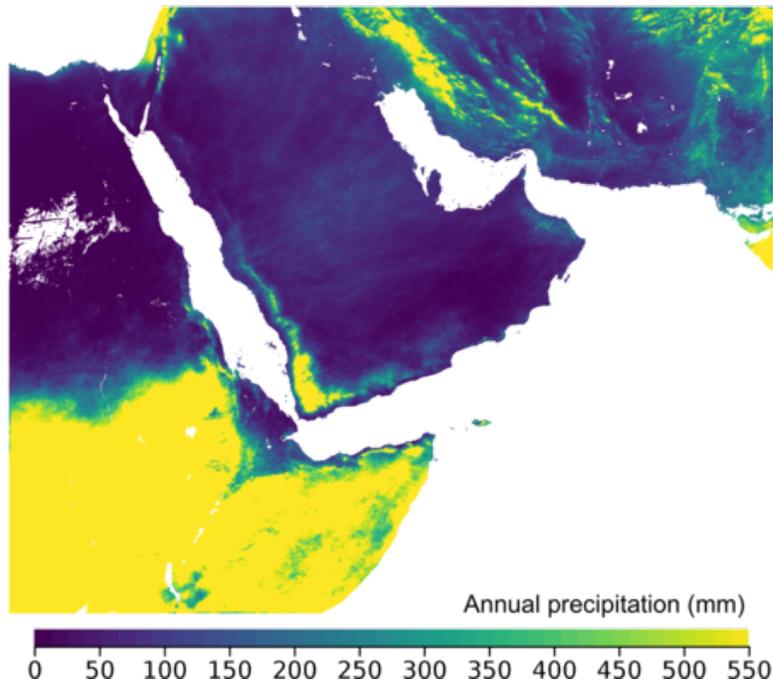


Figure 4.5: Average annual precipitation (2011 - 2015) over the WRF interior ( $3\text{ km}$  resolution) domain.

model will not be able to reproduce the observed evaporation rates (section 4.7). Given the type of irrigation in the fields (sprinkler irrigation system), it makes sense to modify the rainfall input (i.e. prescribe an irrigation rate) in order to provide a realistic amount of water to the system.

#### 4.5 Surface input data for the CABLE model

CABLE includes a default static look-up table of global vegetation classification as defined in *Loveland et al.* (2000) along with global soil texture types *Zobler* (1999). Both classifications are available at  $1^\circ$  resolution (figure 4.7), and are suitable for single site simulations or large-scale regional across the globe (*Wang et al.*, 2011). Given the needs of this study in terms of spatial resolution (sections 1.2 and 4.1), the default vegetation classification is insufficient, especially since the entire Arabian Peninsula is defined as “bare ground” (vegetation type number 14). Also included as a default data set in CABLE auxiliary files is a  $1^\circ$  resolution monthly LAI dataset

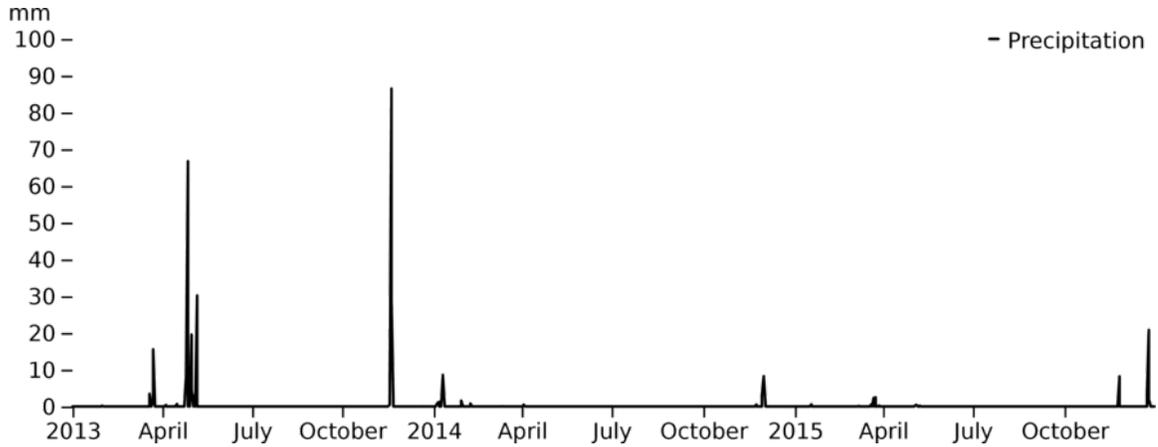


Figure 4.6: Daily precipitation at the Tawdeehiya site for the years 2013 - 2015.

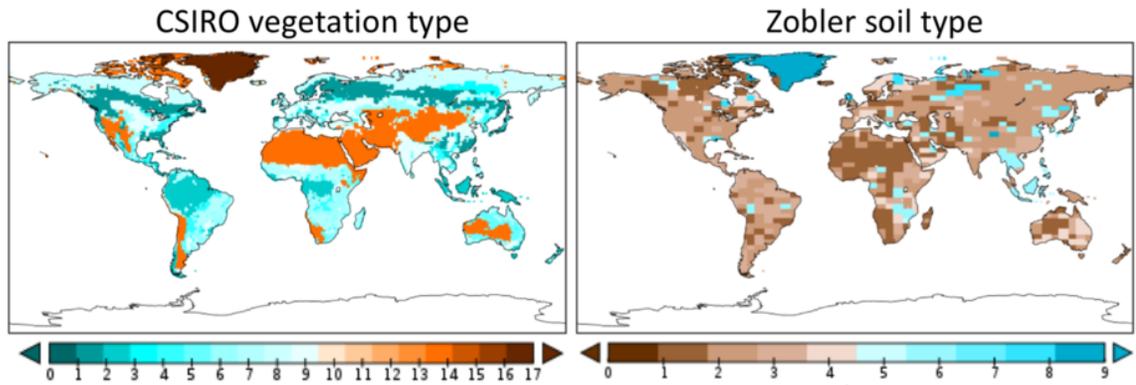


Figure 4.7: Default vegetation and soil classification types in CABLE.

based on MODIS data (averaged from 2002 to 2009) (figure 4.8). In this section, a method to obtain a higher-resolution land cover classification, as well as LAI and albedo values will be introduced. For the soil properties however, we used the default soil classification. Further examination on the accuracy of these soil properties in the study site is needed (e.g. by means of in situ measurement, or by comparison with other higher resolution soil datasets).

#### 4.5.1 High-resolution surface and vegetation characteristics based on Landsat 8 data

One of the challenges with hyper-resolution modeling is the requirement for equivalent high-resolution data characterizing the surface. This is especially true in landscapes

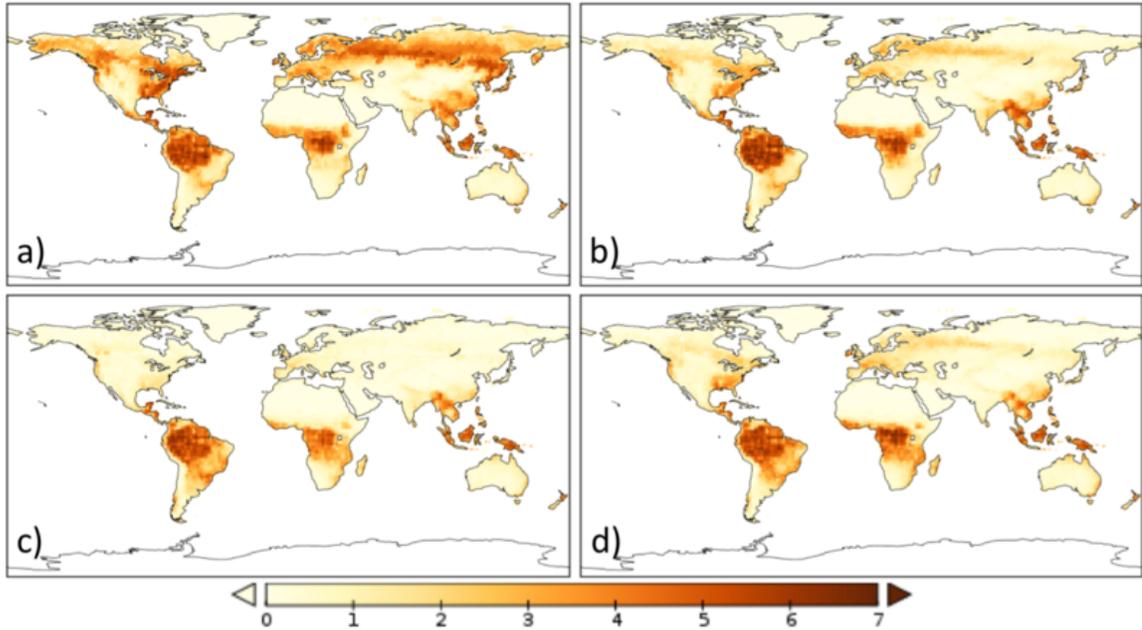


Figure 4.8: Default global monthly LAI dataset based on MODIS data (averaged from 2002 to 2009) included in CABLE auxiliary files (a: July, b: October, c: January, d: April).

with heterogeneous features, such as in pivot irrigation farms, where there is a sharp contrast between the vegetation and the soil (figure 1.2). However, direct measurement of some of the required surface parameters from satellite observations at the desired resolution is not possible. One of such variables is the land surface temperature (LST), a required input for TSEB (section 4.2.1). In such cases, statistical downscaling techniques may be used, such as a simple regression relationship with another variable that is available at the finer resolution. For some parameters however, this relationship is not straightforward but can be inferred through the use of a rule-based decision tree. This machine learning technique can also be used to improve the estimation of other derived parameters, such as LAI, which is traditionally estimated from a simple relationship to a single vegetation index. A decision tree is a type of empirical learning system (*Quinlan, 1996*) that can be used to reveal a set of relationships (or “rules”) between a target “property” and a set of explanatory or “predictor” variables (independent variables), without the explicit knowledge of the

underlying processes. A training dataset is required to establish the non-parametric, multi-variate regression models (rules) that form the terminal “leaves” of the tree, and the quality of this data will influence the model’s validity and portability (*Houborg and McCabe, 2018*).

*Gao et al.* (2012) developed an operational framework that uses data mining to predict thermal infrared (TIR) imagery at higher resolution than it would otherwise be typically available (TIR is acquired at a coarser resolution than shortwave bands). This “data mining sharpening technique” (DMS) used a regression tree approach and performed better than a “traditional thermal sharpening tool”, which resulted in artificial “box-like” results when applied at scenes with heterogeneous vegetation (*Gao et al., 2012*). Furthermore, the DMS is extendable, in the sense that additional predictor variables (e.g. albedo) could be easily included without a “pre-selection” step to determine the best combinations for a specific scene. In this study, LST, which is a required input for TSEB (section 4.2), was downscaled to 30 *m* resolution using a sharpening technique after applying atmospheric correction (*Rosas et al., 2017*). Cubist (RuleQuest; [www.rulequest.com](http://www.rulequest.com)), an open-source implementation of the regression tree model (*Quinlan and Rivest, 1989; Quinlan, 1996*) was used to derive the regression rules.

The landcover classification used in this work (figure 4.9) was derived from 30 *m* resolution Normalized Difference Vegetation Index (NDVI) based on Landsat 8 data corrected for atmospheric effects (*Houborg and McCabe, 2017*). The classification for the year 2015 included four types of crops (alfalfa, maize, rhodes grass and carrots), and a no-vegetation class. In CABLE, the land cover classification includes a bare ground class, C3 and C4 croplands, among others. In this study, we assigned either the C3 cropland class (alfalfa and carrot), C4 cropland class (maize and rhodes grass) or bare ground (no vegetation). During the year 2015, 41 active center-pivot fields were identified and will be the focus of this work (figure 4.9, right). The pivots were

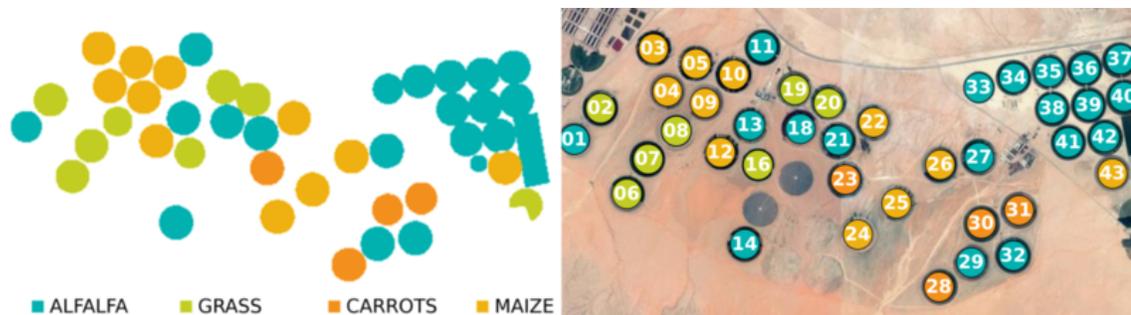


Figure 4.9: Left: Landsat 8 based land cover classification at Tawdeehiya Farm for the year 2015. Right: The 41 pivots numbered approximately in SW-NE stripes. The smaller pivot in the eastern part of the farm was omitted for this study, as well as the pivot located in the SE corner and the rectangular alfalfa field in the eastern part of the farm. Two pivots that did not grow vegetation for the year 2015 were omitted as well (their numbers, 15 and 17, were skipped).

numbered roughly in SW-NE stripes, to distinguish pivots affected by the swath edge of Landsat 8 (e.g. pivots 1 to 11-21, depending on each scene).

Recent work by *Houborg and McCabe* (2018) used Cubist to obtain Landsat 8 based leaf area index (LAI) over the Tawdeehiya agricultural site. For the training dataset, they used a combination of *in situ* and physically-based estimates derived from satellite data. The *in situ* data consisted of 87 sampling units covering a wide range of plant development stages for all crop types at the site (alfalfa, grass, maize and carrots), collected over five field campaigns between November 2014 and October 2015. The satellite-based LAI data was obtained using REGFLEC (*Houborg et al.*, 2015), a model for the inversion of LAI from atmospherically corrected surface reflectances. The implementation of this model for the Tawdeehiya site accounted for some challenges resulting from the extreme conditions of the site, such as adjacency effects and dust deposition (*Houborg and McCabe*, 2016), and used multi-spectral imagery from RapidEye satellites. The use of this hybrid training dataset significantly improved the LAI prediction accuracies, when compared to LAI using *in situ* data alone. The satellite data in the training data set consisted of a selection (or sampling) of 1700 points within the pivots (and a few points adjacent to the pivots in

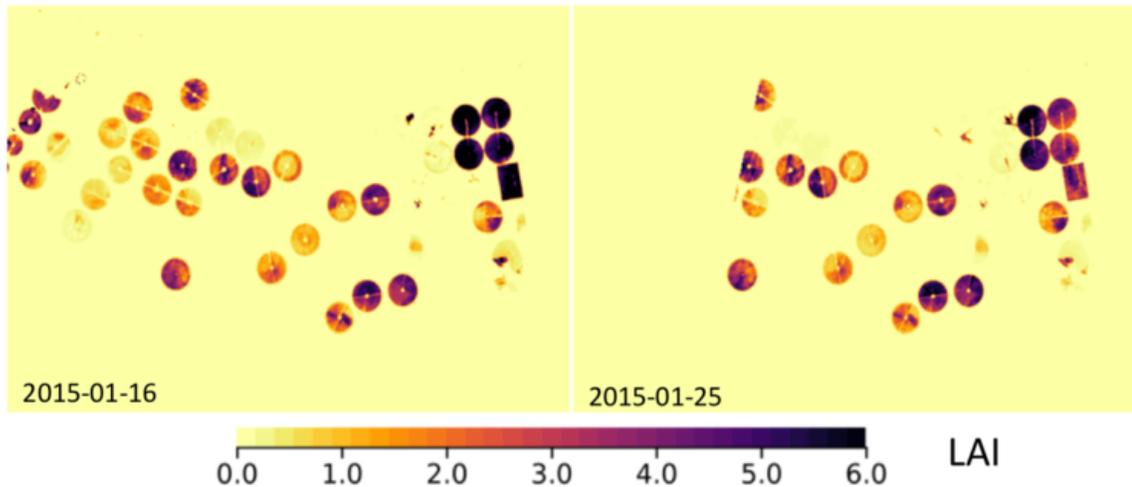


Figure 4.10: Two scenes of Landsat 8 based LAI data from (*Houborg and McCabe, 2018*). The proximity of the farm to the eastern swath edge for Landsat path/row 164/43 is noticeable in the Western side of the farm

the soil). The Landsat 8 explanatory variables consisted of different combinations of 13 vegetation indices, with results indicating that a group consisting of indices using the near infrared (NIR) and shortwave infrared (SWIR) bands improved the accuracy of the predicted LAI values.

In this work, we use NDVI, LAI and albedo data from the work of (*Houborg and McCabe, 2018*) for the years 2014 and 2015 at 30 *m* resolution. Because the farm is situated between two Landsat paths (path/rows 165/43 and 164/43), the data can be obtained with an approximate interval of 8 days, as opposed to the 16-day revisit time for Landsat data. However, the western part of the farm is often not covered by the path/row 164/43 due to the proximity of the farm to the eastern swath edge (figure 4.10). Scenes where cloud coverage was detected were also omitted (figure 4.11). The number of scenes where information is available for all pixels within each of the 41 pivots is displayed in figure 4.12, where the effect of the swath edge is clear: pivots 1-8 are covered by half the number of scenes than in pivots 20-43.

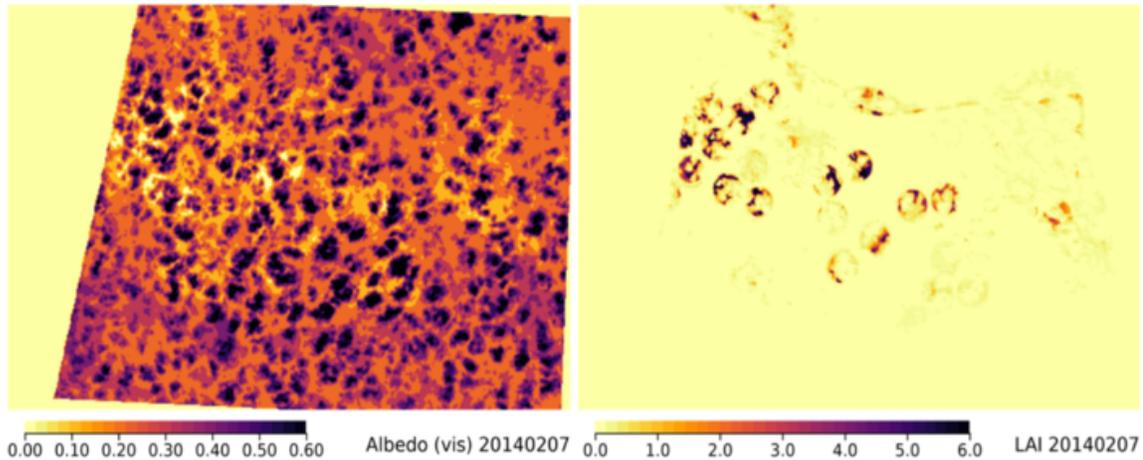


Figure 4.11: Landsat reflectance values upon which surface (e.g. albedo, left) and vegetation parameters (e.g. LAI, right) are based upon are affected when there is significant cloud cover. The number of observations is therefore reduced, affecting the number of evaporation retrievals as well.

## 4.6 Initializing the land surface model

At every point in time, the land surface model requires information on the state of the model at previous time steps, e.g. from soil moisture, canopy water storage, among others. However, at the model start time there is obviously no antecedent information, and thus initial conditions to the model must be given prior to the intended run of the model. Some approaches for initializing or “spinning up” a model include running the model using meteorological forcing from a few years, and running the model consecutively for a representative year of meteorological data (*Coon et al.*, 2016; *Yang et al.*, 2016; *Shi et al.*, 2013). However, given the sporadic nature of rainfall in this region, and the fact that the agricultural fields are irrigated with a much larger amount of water anyway, initializing the model using WRF predicted rainfall would not leave the model in a realistic state representative of the agricultural activities in the region.

Furthermore, given that there is generally no monitoring of actual irrigation rates for every field and time, an assumed value of irrigation must be used. Of course, this could also present some problems, as the uncertainty of these assumed conditions

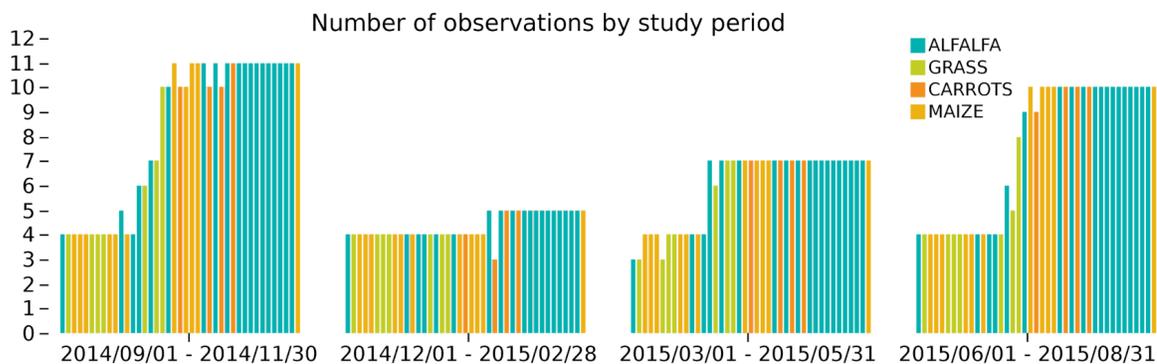


Figure 4.12: Coverage of Landsat scenes (cloud-covered scenes excluded) per pivot and crop type for four three-month periods starting from September 2014. Pivots located on the west side of the Farm receive less coverage due to the swath edge of Landsat 8

could affect the actual run of the model (*Ajami et al., 2015; Nikolopoulos et al., 2011*).

Fortunately, even if there is no actual flow data (i.e. groundwater abstraction from the wells at the center of each pivot), in the case of the Tawdeehiya farm there is at least some information: the reported number of hours of applied irrigation at each pivot per month, and an approximated constant flow rate for each pivot (in gallons per minute). With this information, an “effective” monthly irrigation rate can be calculated each month as  $EMI = Cn_{applied}/n_{tot}$ , where  $C$  is the constant flow rate,  $n_{applied}$  is the number of hours of irrigation application, and  $n_{tot}$  is the maximum number of hours in a given month. Figure 4.13 shows this calculation for one pivot field. Of course, the reported information is missing any temporal variations in the irrigation, e.g. the actual timings when it is turned on or off. The EMI rate is only an approximated constant monthly rate that can be applied to the model in order to get a somewhat more realistic initialization.

The farm data covers the years 2014, 2015 and part of 2016. Therefore, we selected 2014 for model initialization, and the period September 2014 - August 2015 as the study period. The model was initialized under two different scenarios: 1) an assumed constant irrigation application of  $12 \text{ mm.day}^{-1}$ , and 2) the calculated

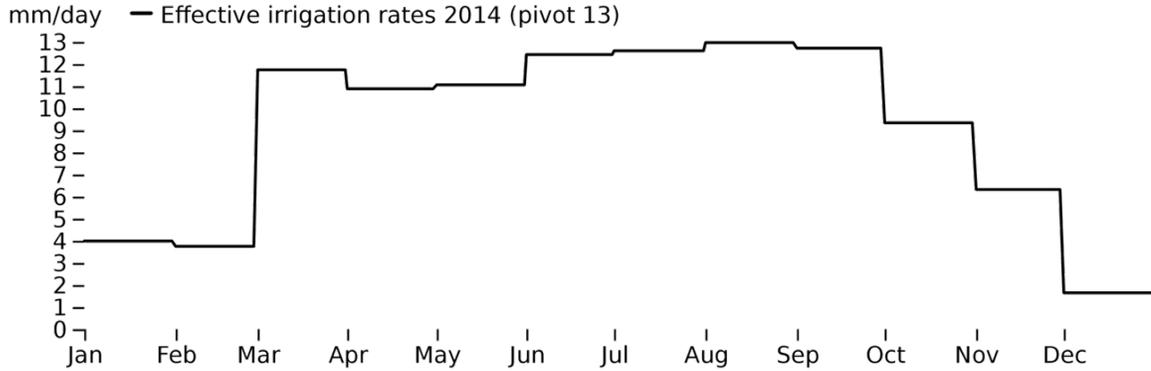


Figure 4.13: Calculated “effective monthly irrigation” based on farm data for pivot number 13 (see figure 4.9)

“effective” monthly irrigation (EMI) rates based on reported farm data. The model was initialized by consecutively running the model five times for the year 2014. For both scenarios, the initial soil moisture was set to a value of 0.3 for all six layers. An additional experiment was set for both scenarios but with an initial soil moisture of 0.1 for all six layers, in order to verify if different initial conditions would have a long-term effect on the state of the model.

Figure 4.14 shows the average soil moisture across all six layers for the same pivot shown in figure 4.13, during the last consecutive run of the model, using the assumed constant irrigation (a-b) and EMI rates (c-d). The effect of initializing the model at different soil moisture values ( $0.3 \text{ m}^3.\text{m}^{-3}$  for a and c;  $0.1 \text{ m}^3.\text{m}^{-3}$  for b and d) is negligible. However, the sharp changes in irrigation rates for the EMI scheme cause sharp changes in the soil moisture.

#### 4.7 Evaporation from land surface modeling using WRF rainfall only: comparison with remote sensing based evaporation at Tawdeehiya farm

After the “spinup” period of 2014, the model was run from March 2014 to August 2015 using only the predicted rainfall from the WRF model. As shown in figure 4.15,

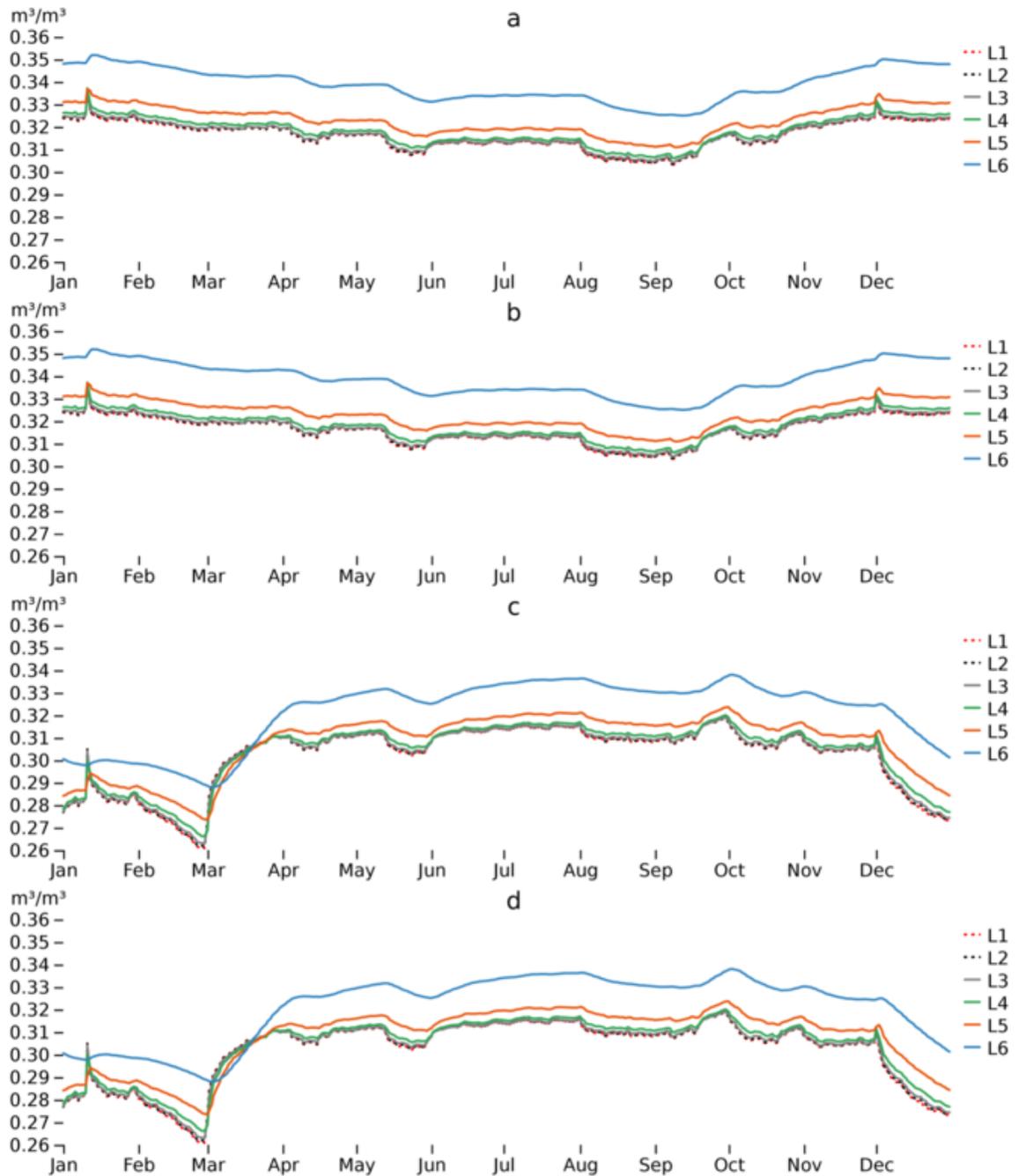


Figure 4.14: Soil moisture time series (six layers) for the constant (a-b) and “effective irrigation” (c-d) irrigation schemes for the fifth consecutive run of the model (year 2014). The model was initialized with values of 0.3 (a, c) and 0.1 (b, d)  $m^3.m^{-3}$  for all layers.

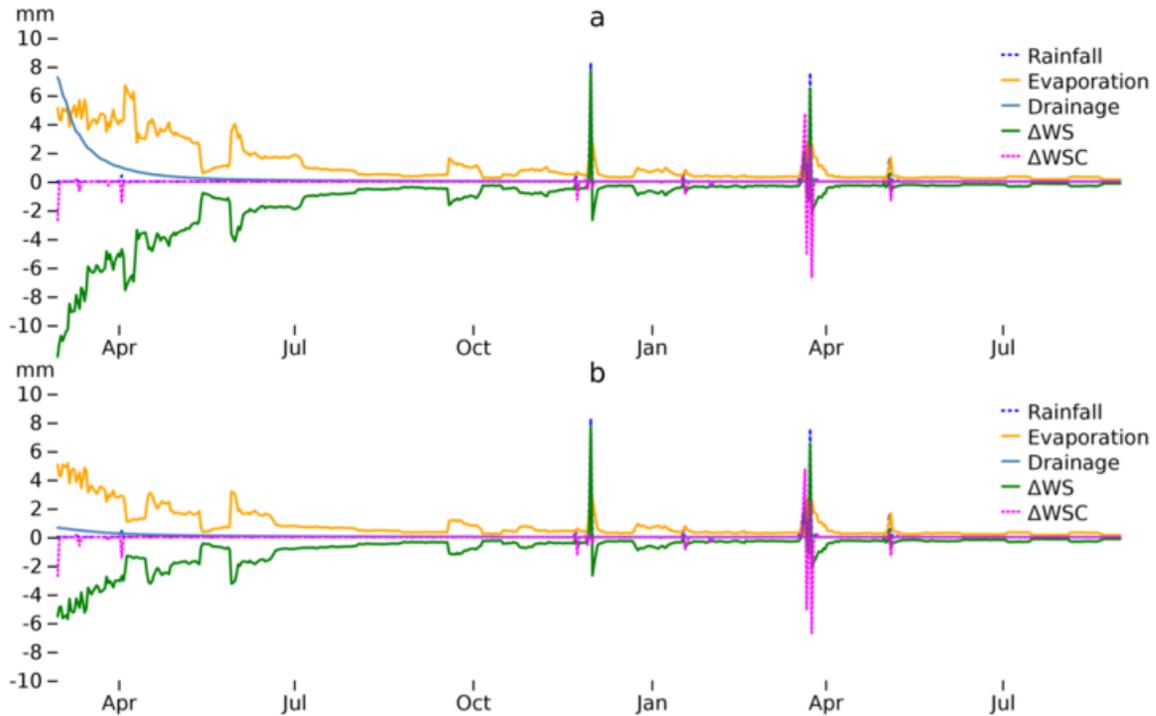


Figure 4.15: Daily water budget for the WRF-rainfall model run from March 2014 to August 2015 after the spinup period, using two different spinup scenarios: constant irrigation (a) and EMI (b).  $\Delta WS$  is the change in water storage from the soil, while  $\Delta WSC$  is the change in water storage from the canopy layer.

initially the remaining water stored in the model from the spinup period causes the model to output large evaporation rates, even with the absence of rainfall. However, after only a few months, i.e. around September 2014, both runs of the model shows similar responses, both in terms of magnitude and partitioning of the fluxes. For this reason, and also to avoid the effect that sharp changes in the rainfall input (either from different initialization scenarios, or from the change from initialization to the actual run of the model) might have on the model state, the study period was selected to start from September 2014, but with a model run starting from March 2014. The remote sensing based evaporation model (TSEB; section 4.2) was run for 40 days within the study period, constrained upon the availability of adequate Landsat scenes. Figure 4.17 shows spatial maps of the high-resolution evaporation results for eight selected dates representative of the study period (September 2014

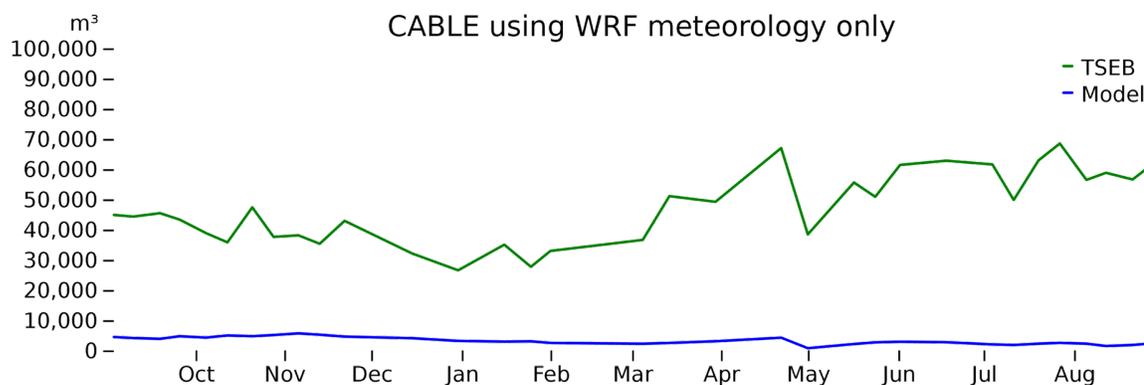


Figure 4.16: Estimated total evaporated water in  $m^3$  from TSEB and CABLE without prescribed irrigation.

- August 2015). At this resolution, it is possible to identify a range of evaporation rates across different pivot fields as well as some sub-field variation. Furthermore, an impact of the vegetation dynamics is also evident, with most pivots showing larger evaporation rates during spring and summer.

The same spatial maps of high-resolution daily evaporation obtained by CABLE with only WRF predicted rainfall are shown in figure 4.18. As expected, the evaporation rates are negligible in comparison with those predicted by TSEB, with some minimal evaporation caused by sporadic rainfall events during winter of 2014/2015 and spring 2015 (see figure 4.6). Finally, we computed the total amount of water evaporated from the system in  $m^3$  from the 17 easternmost pivots during the study period (i.e. the ones with the most coverage by Landsat and therefore TSEB as well), as predicted by both the RS model and the land surface model (figure 4.16) during days in which there were TSEB data available. Clearly, the land surface model forced with meteorological data alone is not able to account for the vast amounts of water being used by these agricultural fields.

## 4.8 Conclusions

In an effort to capture realistic hydrologic interactions over small catchments (including farm-scale irrigation), there have been concerted efforts to improve both land surface schemes and the input data required to drive these models. For example, high-resolution satellite imagery, combined with advanced machine learning techniques, enables a capacity to retrieve surface and vegetation parameters (e.g. LAI, albedo, NDVI) at sub-farm scales, with the ability to characterize individual center-pivot fields. These high-resolution datasets (e.g. 30 *m*), coupled with hourly meteorological predictions from advanced numerical weather prediction models, can be used to drive high-resolution simulations of a modern land surface model. However, these efforts have not fully overcome the challenges relating to the need to account for heterogeneities in the surface, the inclusion of biogeochemical processes, anthropogenic changes affecting water movement, as well as the associated increased computational needs, both in terms of storage and processing time.

In this chapter, the requirements for running (1) a remote sensing based evaporation model and (2) a land surface model, both at high-resolution (30 *m*) was introduced, with the goal of determining whether the impact from agricultural activities in Saudi Arabia (figure 1.2) could be detected. The land surface model can be used to study the full water and energy exchanges between the soil, canopy and atmosphere, but it requires the inclusion of rainfall data. In the case of this study region, it also requires the additional (and significantly higher) input of water from irrigation, a requirement that is generally not available. The thermal-based remote-sensing model on the other hand does not demand this input and is able to capture the impact from irrigated agriculture, as can be seen clearly in figure 4.18. Unfortunately, it can only do so for a select number of days when satellite information on LST and vegetation characteristics is available. Furthermore, given that it is a model focused on evaporation retrieval, it cannot be used to provide a complete water budget.

Clearly, the anthropogenic component (i.e. irrigation forcing) is missing from the purely meteorological (i.e. rainfall) runs of the model (figures 4.17 and 4.18). While there is some information available (although not directly what is needed for the model) for the study region (and indeed it was used for one of the two initialization approaches; see Section 4.5), the goal of this work is to develop a method that could be generalized and applied to any irrigated agricultural region where this data might not be available (which is normally the case). This idea will be explored further in Chapter 5, where the evaporation information obtained by the remote sensing model will be combined with the land surface model in order to infer irrigation, and hence groundwater abstraction rates.

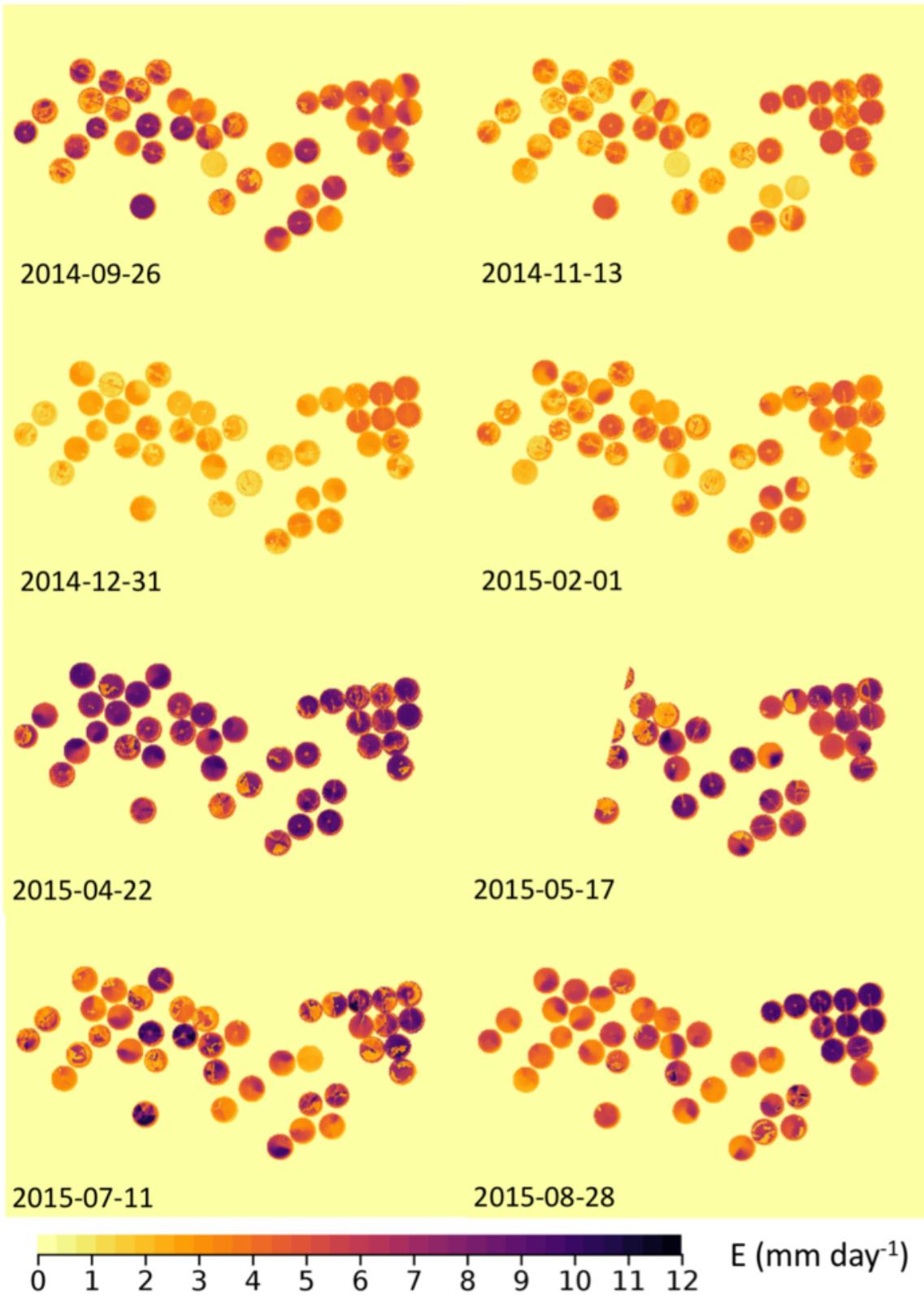


Figure 4.17: High-resolution daily evaporation maps obtained by TSEB.

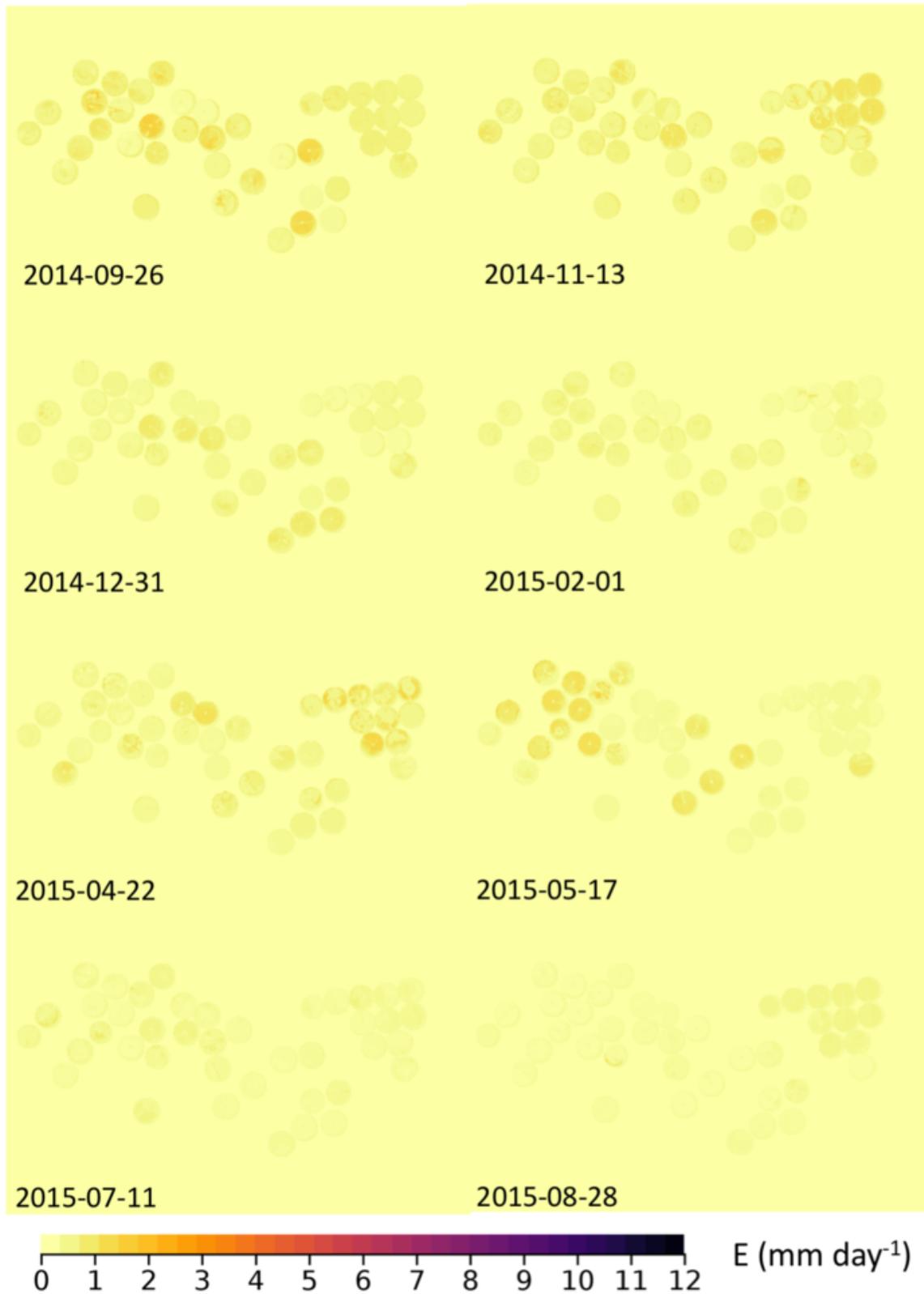


Figure 4.18: High-resolution daily evaporation maps obtained by CABLE using only WRF rainfall as “irrigation”.

## Chapter 5

### **Estimating water use in irrigated agriculture: combining high-resolution remote sensing of evaporation and land surface modeling**

---

This chapter presents a novel method aimed to retrieve groundwater abstraction rates used in irrigated fields by constraining a land surface model (LSM) with remote sensing-based evaporation observations (model details in chapter 4). Here we present the optimization technique central to this method and the results of applying it to 41 irrigated fields. While low coefficient of efficiency (COE) values were obtained, mean absolute errors were below  $3 \text{ mm.day}^{-1}$ , and the LSM generally underestimated irrigation when compared with farm reported data (bias of  $-1.6 \text{ mm.day}^{-1}$ ). It was also shown that these metrics were significantly better for fields located on the east of the farm, likely due to the higher frequency in observations. Soil evaporation was shown to be overestimated by the LSM, particularly between vegetation growing cycles during summer and fall months (MAE =  $-1.41 \text{ mm.day}^{-1}$ , bias =  $1.49 \text{ mm.day}^{-1}$ ). The total estimated groundwater abstraction consumed by the 41 fields was  $65.8 \text{ Mm}^3$ , with about 28% representing drainage to the subsurface and 72% evaporation.

---

## 5.1 Introduction

Efforts to capture a realistic water balance at sufficient resolution consistent with the presence of agriculture in an extreme arid landscape were described in chapter 4. A remote sensing based evaporation model (sections 4.2.1) forced with high-resolution vegetation and surface parameters was used to retrieve daily evaporation rates, showing spatial and temporal variations consistent with the presence of agricultural fields (figure 4.17). However, these estimates were retrieved only at semi-regular frequencies, depending on the availability of satellite imagery, the presence of clouds and the performance of statistical techniques used to retrieve certain parameters required for the model (section 4.5.1). To fill these temporal gaps, as well as to study other hydrologic interactions (e.g. drainage to the subsurface), the use of a land surface model (LSM) was also explored (section 4.3). However, without information that adequately describes the agricultural water input (i.e. irrigation), the LSM cannot offer a realistic description of the hydrologic interactions.

In this study, we inquire as to whether the remote sensing-based evaporation retrievals (RS-E) can be used to constrain the LSM in order to obtain this needed information. To do this, we propose a methodology that treats RS-E as “observations”, to be matched by the evaporation obtained by the land surface model (LSM-E) in an iterative procedure that aims to minimize the difference between these two (figure 5.1). The procedure is as follows: after initializing the LSM (section 4.6), it is run for a period of time assuming a constant irrigation rate. The difference between RS-E and LSM-E is used to update the irrigation rate and re-run the LSM for this period. This is repeated until the difference in values between consecutive runs is negligible, or if a maximum number of iterations is reached. The LSM is then run with the final value of the irrigation rate for this period of time, and the state at the end of the period is used as the initial conditions for the next optimization period. Details of the optimization algorithm are covered in section 5.2. To simplify the procedure,

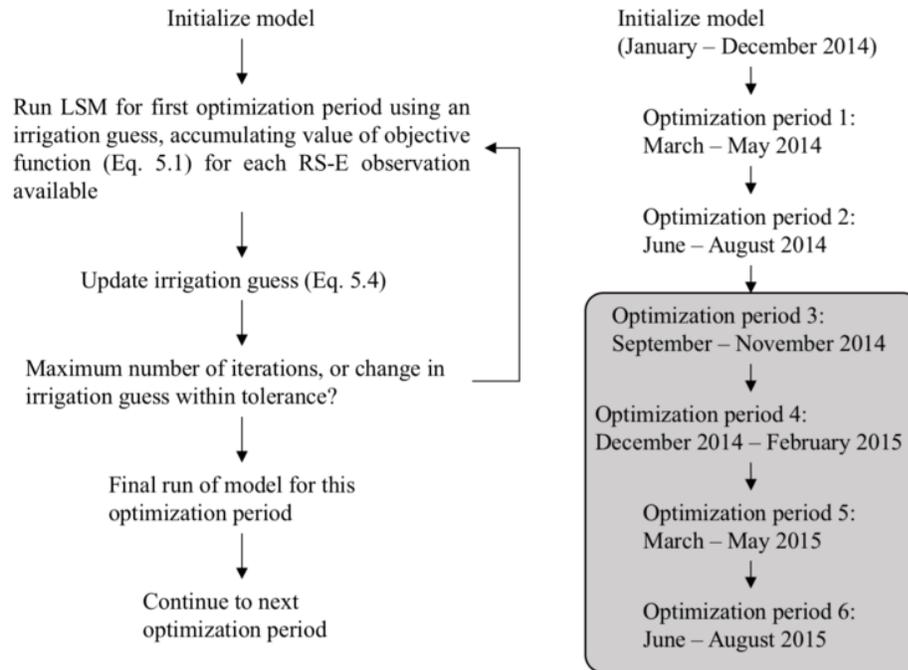


Figure 5.1: Diagram showing the general methodology for estimating irrigation rates (left) and optimization periods used in this study (right). The grayed area represents the study period.

the optimization was based solely upon the retrieval of evaporation and assuming different irrigation rates applied depending on the season. Further refinements to the methodology could include optimization of soil parameters in the model.

Because the number of observations is irregular both in time and across the study region (see figures 4.10 and 4.12), a choice has to be made to find a balance between the length of the optimization periods and the number of observations that are used to constrain the LSM. To keep the methodology as general as possible, the size of each optimization period was set to a constant of three months. The optimization periods were selected as follows: 1) March 2014 - May 2014, 2) June 2014 - August 2014, 3) September 2014 - November 2014, 4) December 2014 - February 2015, 5) March 2015 - May 2015, 6) June 2015 - August 2015, i.e. allowing for two optimization periods (1 and 2) between the initialization of the LSM and the actual study period (September 2014 - August 2015; see section 4.6).

## 5.2 Methodology

### 5.2.1 Optimization

The simultaneous perturbation stochastic approximation (SPSA) algorithm is an automated calibration method for problems where a large number of parameters need to be determined (i.e. multivariate optimization problems). The key distinguishing feature is that all model parameters are perturbed in the same iteration to generate a search, and it does not depend on direct measurements of the gradient of the objective function with respect to the parameters. Instead, an approximation of the gradient is computed based on the simultaneous perturbation of the parameters, requiring only two objective function evaluations independent from the number of parameters to optimize. This method has received attention in multiple disciplines that require the calibration of models with a large number of parameters, including for example aerodynamic shape design optimization (*Xing and Damodaran, 2005*), wind farm design (*Ahmad et al., 2014*), tidal models (*Altaf et al., 2011*), and aquifer parameter estimation (*Goyal, 1982*). In this section, the computational cost and effectiveness of SPSA will be explored, in the context of determining the irrigation amounts applied to individual fields within the Tawdeehiya farm. The optimization is based on the ability of the LSM to reproduce remote sensing-based daily evaporation rates, evaluated using the objective function (5.1):

$$J(\gamma) = \sum_i [Y(t_i) - H(X(t_i))]^T R_i^{-1} [Y(t_i) - H(X(t_i))] \quad (5.1)$$

where  $Y(t_i)$  are observations of evaporation (RS-E) in  $mm.day^{-1}$ , and  $X(t_i)$  is the LSM output evaporation (LSM-E) in  $mm.day^{-1}$ .

We start with an initial guess, a set of parameters  $\gamma^0$  that represents the irrigation values applied to each of the sub-domains in Tawdeehiya field. The set of parameters

is perturbed around the initial guess to obtain two new sets:

$$\gamma^k + c_k \Delta_k \quad (5.2)$$

$$\gamma^k - c_k \Delta_k \quad (5.3)$$

where  $\Delta_k$  is a column vector with  $n_p$  (number of parameters) independent samples from the  $\pm 1$  Bernoulli distribution (i.e. random samples of +1 or -1), and  $c_k$  is a scaling factor given by:

$$c_k = \frac{c}{(k+1)^\beta} \quad (5.4)$$

Using these perturbed sets of parameters, a stochastic approximation (or stochastic gradient)  $g_k(\hat{\gamma}^k)$  of the objective function  $J(\gamma)$  to minimize is computed:

$$\hat{g}_k(\gamma^k) = \frac{J(\gamma^k + c_k \Delta_k) - J(\gamma^k - c_k \Delta_k)}{2c_k} \Delta_k^{-1} \quad (5.5)$$

i.e. it is computed from two evaluations of the objective function  $J(\gamma)$  based on the simultaneous perturbation around the current set of parameters. The set of parameters are iteratively corrected based on the stochastic gradient as:

$$\gamma^{k+1} = \gamma^k - a_k \hat{g}_k(\gamma^k) \quad (5.6)$$

where  $a_k$  is given by

$$a_k = \frac{a}{(A+k+1)^{\hat{\alpha}}} \quad (5.7)$$

The choice of parameters  $\beta$ ,  $c$ ,  $A$ ,  $a$ , and  $\hat{\alpha}$  is dependent on the optimization problem. However, following the guidelines of *Spall* (1998) and recommendations by *Altaf et al.* (2011), setting  $\hat{\beta} = 0.101$ ,  $\hat{\alpha} = 0.602$ , and  $A$  to 10% of the maximum number of iterations allowed yields effective results. It remains to determine the SPSA parameters  $a$  and  $c$  for this specific problem.

### 5.2.2 Tuning the SPSA parameters using synthetic data

The parameter  $a_k$  in the SPSA method controls the weight that the cost (value of the objective function) will have on determining the value of the model parameters for the next iteration. However, this parameter itself decays over each iteration, so that the impact of the cost is lower over the course of the optimization. The decay of  $a_k$  is controlled by the value of  $\hat{\alpha}$ , which for this study is set to a recommended value. However, the initial value of  $a_k$ , i.e. the value for the first iteration, is a critical factor in the optimization because it determines the initial “jump” that the irrigation guess will have. This initial value is affected by the choice of  $a$ . Another critical parameter is  $c$ . Similarly to the impact that  $a$  has on  $a_k$ ,  $c$  affects the initial value of  $c_k$ , the decay of which is controlled by a recommended value of  $\hat{\beta}$ . In this section, several values of  $a$  and  $c$  were explored.

Four fields across the Tawdeehiya field (field numbers: 13, 20, 25, and 30; see figure 4.9) were selected for this experiment. Following the “spinup” period (see sections 4.6 and 5.1), we set the LSM to run for three months: March, April and May of 2014 (i.e. the first optimization period in figure 5.1) using an initial guess for the constant irrigation rate to be applied during this period.

For the purpose of tuning the SPSA parameters, the evaporation values obtained during the last run of the spinup period were selected as “observations”, but only for a select number of days coinciding with TSEB observations availability. After each LSM run, the objective function was evaluated by comparing the obtained evaporation ( $X$  in equation 5.1) and the “observed” value (i.e. from the last run of the spinup period;  $Y$  in equation 5.1). The number of maximum iterations was set to 30. We repeated this experiment for each of the four fields, and using nine pairs of values for  $a$  and  $c$ : s1 (0.01, 0.001), s2 (0.05, 0.001), s3 (0.1,0.001), s4 (0.01,0.005), s5 (0.05, 0.005), s6 (0.1, 0.005), s7 (0.01, 0.1), s8 (0.05, 0.1) and s9 (0.1, 0.1).

A comparison of the optimized irrigation values and its associated cost from these

sets of SPSA parameter values is shown in figures 5.2-5.3. Shown in grey dotted lines are the three monthly “effective irrigation” rates used during the spinup period for the months of March, April and May 2014, while the average of these three values is shown as a solid black line. For this set of experiments (where the observed values come from the spinup period), this average value would be the “true” value of the irrigation that the optimization algorithm is aiming to retrieve.

The last three sets (s7-s9, i.e. with  $c=0.1$ ) diverged to unrealistically high irrigation values ( $>20$ ,  $>50$  and  $>100 \text{ mm.day}^{-1}$ , respectively) and were therefore excluded from figure 5.2. The first (s1) and fourth (s4) sets of parameters produced similar responses (same value for a, with c five times larger in “s4”) in the irrigation guess, correctly arriving at the “optimal” value for B and C, but not for A, with only a slow increase in the irrigation guess. In the second selected field (B), the sets s2, s3, s5 and s6 reached the average value after only 4-6 iterations, then oscillated around this value. This oscillation also occurred for C, but in this case, sets s3 and s6 reached a slightly higher value. Notice however, that the differences in final cost values were minimal between the six sets shown in B and C. In D, the large spread in “true” irrigation values (notice the different Y-scale compared to A, B and C) hindered the ability of the algorithm to find an “optimal” value for this three-month period, as evidenced by the larger costs (which actually increased initially rather than decrease). Finally, in order to test the sensitivity of the method to a different initial guess (see also section 5.2.3), the experiment was repeated with an initial value almost twice as large as the original one ( $15.6 \text{ mm.day}^{-1}$ ). The results for A and B are shown in figure 5.4, with similar oscillatory behavior exhibited by s2, s3, s5 and s6 for B, and in general the same smooth transitions shown in figure 5.2 although mirrored in the Y-axis.

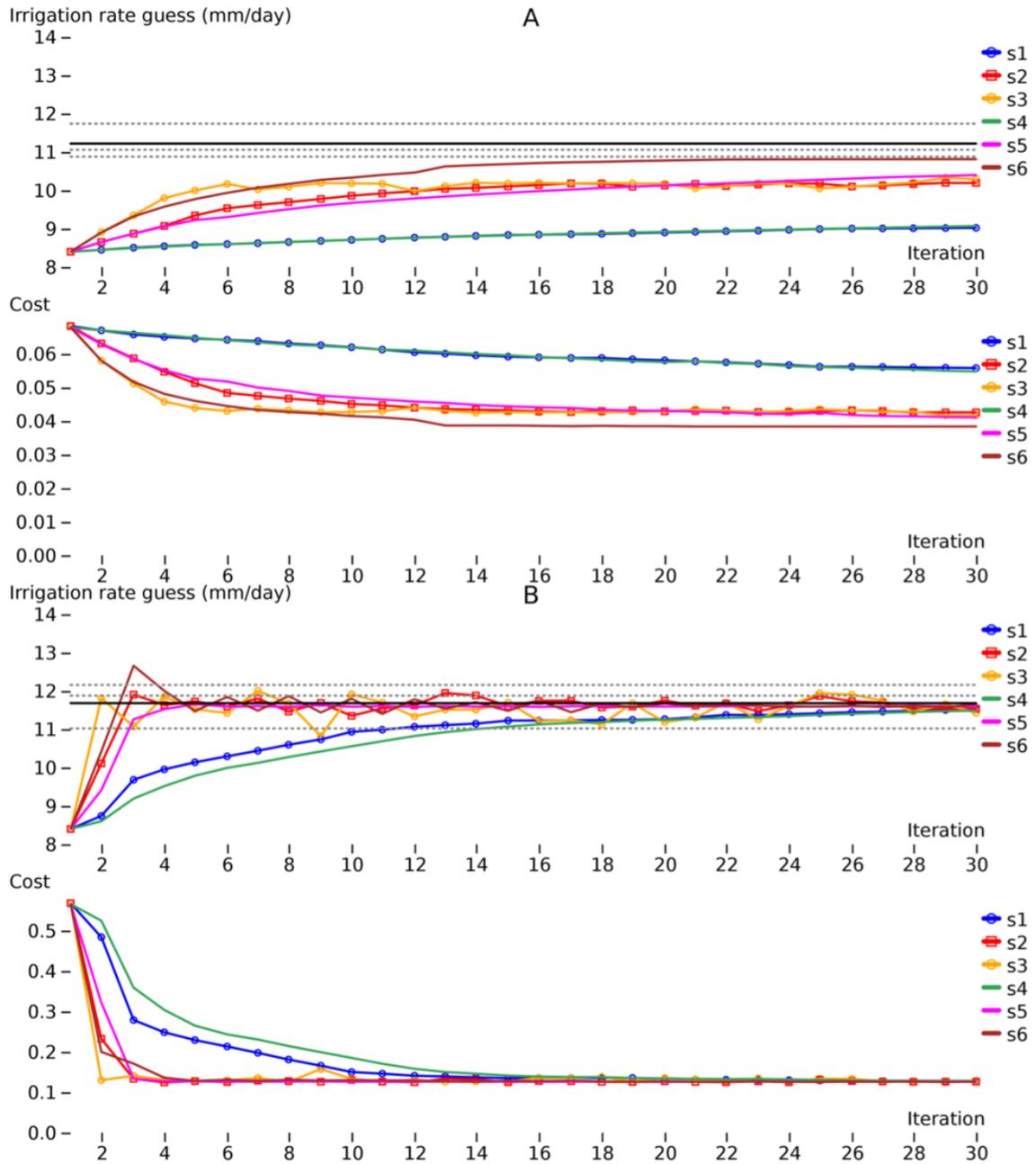


Figure 5.2: Results of six tuning experiments for two selected fields (A:13, B: 25)

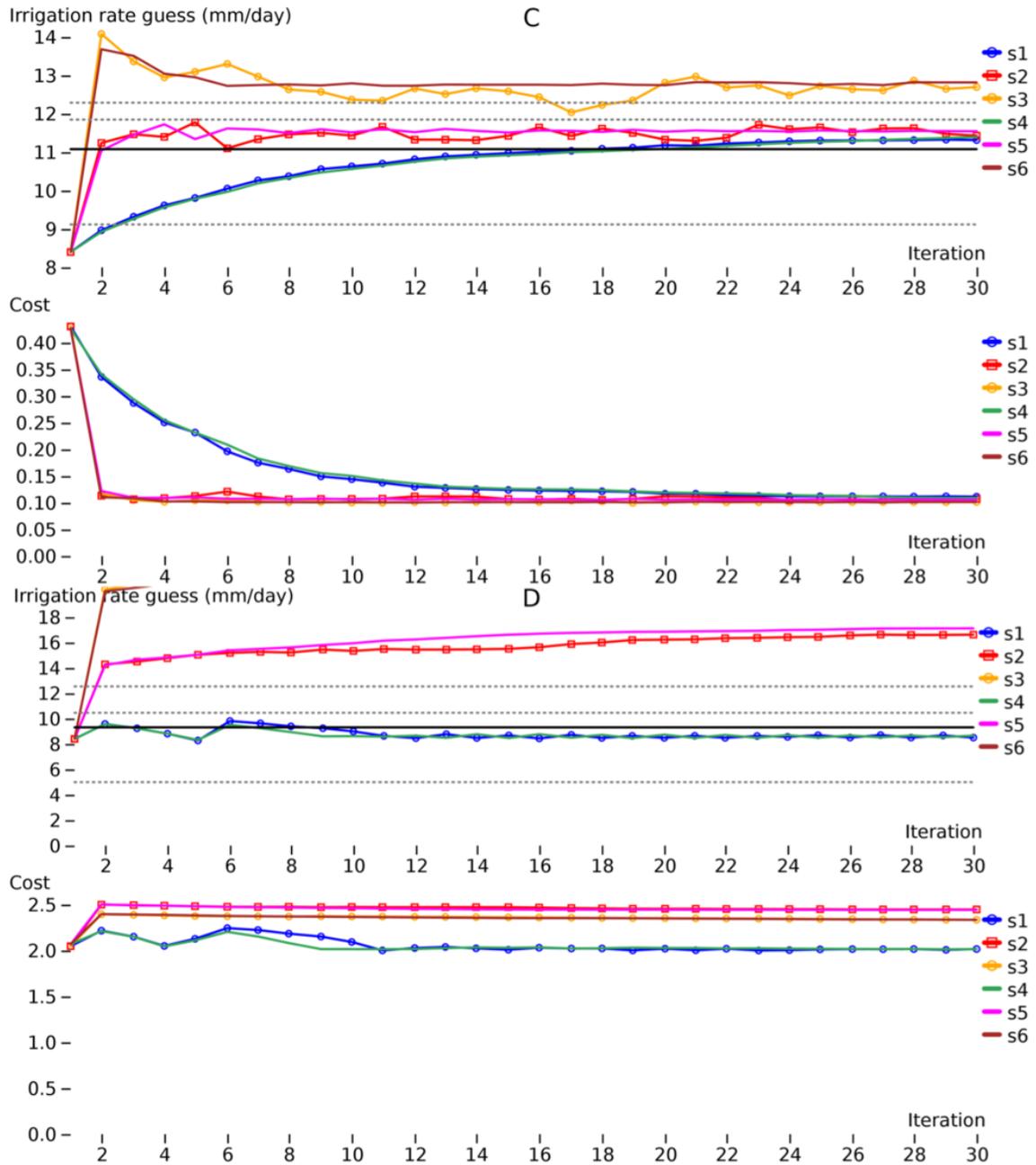


Figure 5.3: Results of six tuning experiments for two selected fields (C: 25, D: 30).

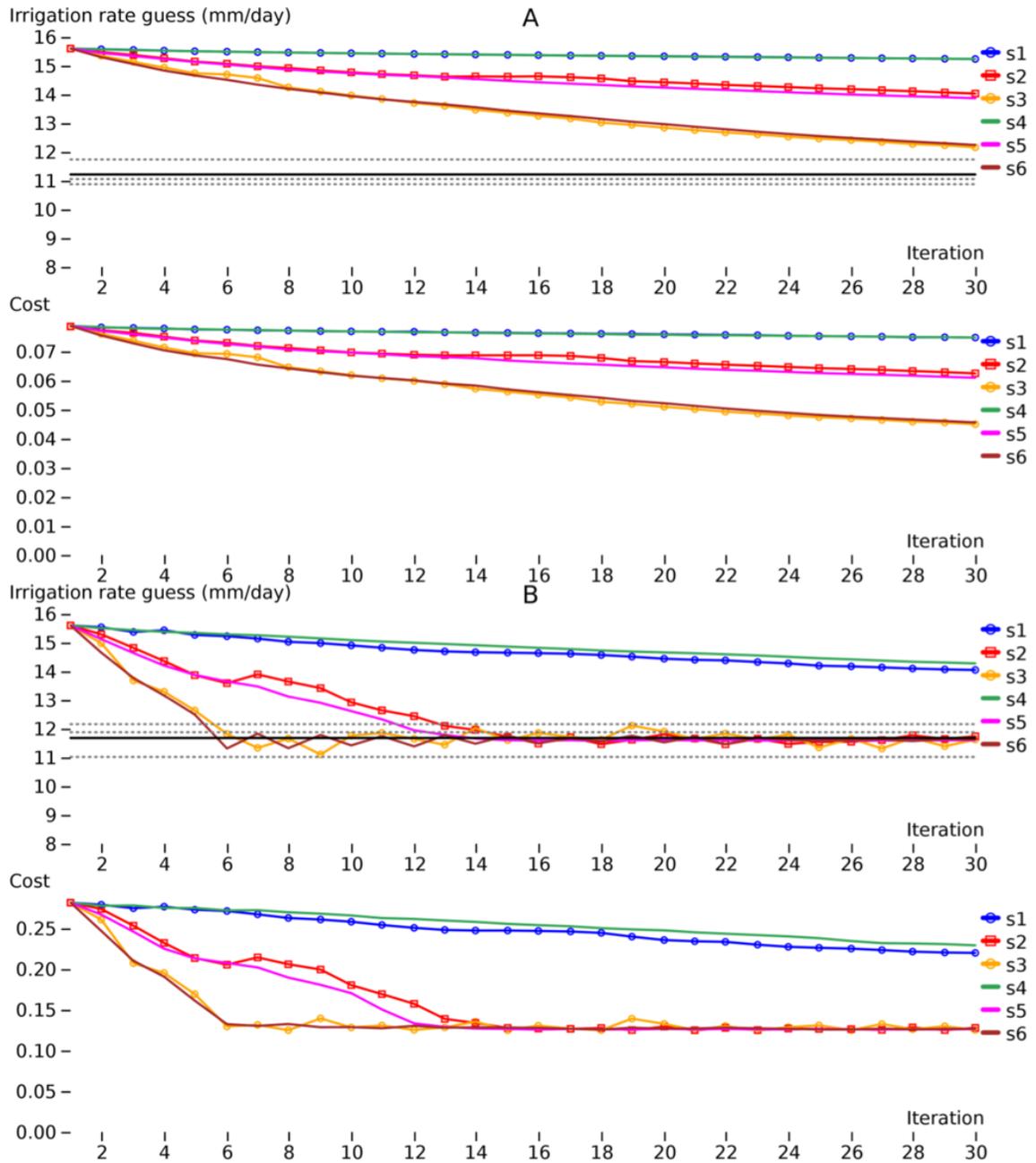


Figure 5.4: Results of six tuning experiments for two selected fields (A:13, B:25) using a higher initial guess

### 5.2.3 Testing the performance of SPSA using synthetic data

In order to test the performance of the SPSA algorithm with a larger set (all 41 fields), we repeated the experiment from section 5.2.2 (i.e. using the evaporation from the last run of the spinup as “observations” to constrain the LSM), using only the set of SPSA parameters “s5”. The results are shown for all 41 fields in figure 5.5 by first taking the average value of the “effective irrigation” data as the x coordinate, while the optimized value is the y coordinate.

The “optimal” parameter obtained by SPSA was selected based on the lowest cost (A) or taken from the final iteration value (B). Both need not be the same because of the oscillatory nature of the methodology (e.g. figure 5.2B). Although the difference is minimal in figure 5.5, the first approach (lowest cost) was ultimately selected for the methodology. In order to show the relative range of values that were taken for the average (i.e. March, April and May EMI values), the difference between the maximum and minimum values was used to scale the size of the circles in the scatterplot. When this range is large (e.g. larger than  $5 \text{ mm.day}^{-1}$ ), then the value that the optimization is attempting to retrieve is not well-defined and therefore the optimization might be “trapped” in a local minimum, as was the case in figure 5.2D. This can be seen in figure 5.5: in general, the largest circles are farther from the 1:1 line (shown as a dotted gray line).

### 5.2.4 Sensitivity to the initial guess

We repeated the optimization of all 41 fields with different values for the initial guess: 0.75, 0.9, 1.1 and 1.25 times the initial guess used in section 5.2.2 ( $8.4 \text{ mm.day}^{-1}$ ). The optimization failed for most fields with the lowest initial guess ( $6.3 \text{ mm.day}^{-1}$ ), irrespective of the difference between the three irrigation rates (i.e. size of circles in figure 5.6), with most of the results yielding unrealistically high values ( $>18 \text{ mm.day}^{-1}$ ). For the next initial guess ( $7.56 \text{ mm.day}^{-1}$ ), the results were closer to the ones found

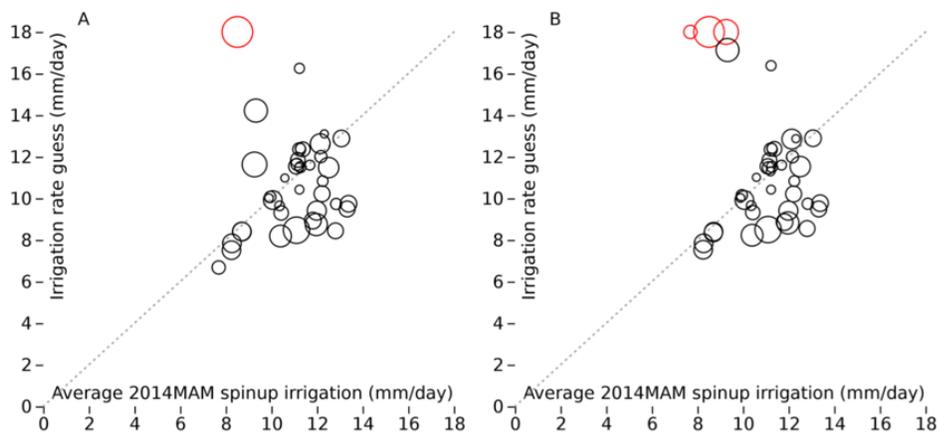


Figure 5.5: Results from applying one set of SPSA parameters to all 41 fields for the first optimization period using synthetic data as observations. Circles in red are out of scale (larger than  $18 \text{ mm.day}^{-1}$ ). See text for additional details.

in section 5.2.2 (figure 5.5), although there were still a few unrealistically high values (shown in red). For larger initial guesses (i.e. 10% and 25% more than the original one used in section 5.2.2), the results had better agreement with the “true” irrigation rate values.

### 5.2.5 Validation measures

The sole source of information available in terms of irrigation values is the “effective monthly irrigation” (EMI) rates calculated from reported monthly number of hours of (constant) irrigation applied (section 4.6). In order to provide a validation measure for the optimized irrigation values, a number of statistical indicators were computed based on the difference between the optimization results and the EMI rates. *Krause et al.* (2005) provided a comparison of efficiency measures normally used in the evaluation of hydrological models. The Nash-Sutcliffe (*Nash and Sutcliffe, 1970*) coefficient of efficiency (COE; equation 5.7) ranges from  $-\infty$  to 1 (perfect fit), with a threshold value of 0 indicating whether the model (M) performed at least as well as the mean of the observed values ( $\bar{O}$ ):

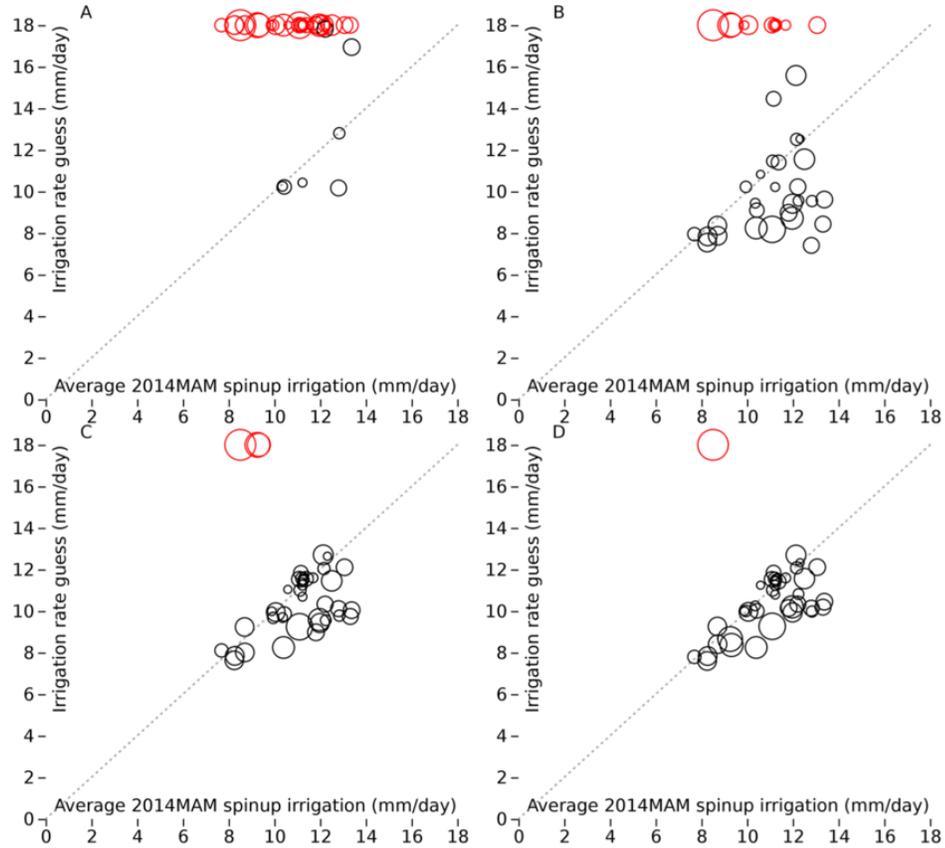


Figure 5.6: Testing different initial guess values (A: 6.3; B: 7.56; C: 9.24 and D: 10.5  $mm.day^{-1}$ ). Optimization during the first three-month period (2014MAM) using synthetic data as observations.

$$COE = 1 - \frac{\sum_{i=1}^n (O_i - M_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (5.8)$$

However, because this measure is not very sensitive to systematic bias, it is recommended to include bias ( $\bar{M} - \bar{O}$ ) as a separate measure. Furthermore, because COE is calculated from squared differences, larger values have a larger impact on the measure. In order to account for this, the mean absolute error (MAE; equation 5.8) will also be used as a supplementary indicator of model performance:

$$MAE = \frac{1}{N} \sum_{i=1}^n |O_i - M_i| \quad (5.9)$$

## 5.3 Results

### 5.3.1 Comparison with reported farm data (EMI rates)

The SPSA algorithm was applied to the six optimization periods sequentially, i.e. using the final state of the LSM after finding the optimal irrigation rate for the previous period (figure 5.1). Here we evaluate the results for the last four optimization periods (2014SON, 2015DJF, 2015MAM and 2015JJA). In order to establish some sort of validation in terms of the obtained irrigation rates, we again make use of the farm data as the only source of information (see section 4.6). Figure 5.7 shows a comparison of the obtained irrigation rates (41 fields and four study periods) against the average of the observed data for each corresponding three-month periods. In general, the optimized values followed similar spatial and temporal patterns as the average EMI values, with lower values during the winter months (2015DJF) and higher during spring (2015MAM) and summer (2015JJA). Several fields showed values that overestimated the observed (average EMI) values during the last optimization period.

A scatterplot comparison of all values (4 periods x 41 fields) is shown in figure 5.8A, with one value that exceeded  $18 \text{ mm.day}^{-1}$  omitted (shown as white in figure 5.7). Evaluating the overall performance of the optimized values quantitatively, the results indicate a poor correlation (COE = -0.56; MAE =  $3.04 \text{ mm.day}^{-1}$ ; bias =  $-1.59 \text{ mm.day}^{-1}$ ). However, considering the fact that the observed irrigation data is an average of three-month periods, some of which might be based on a large range of irrigation rates, we excluded data with a range (maximum - minimum) of observed values exceeding  $5 \text{ mm.day}^{-1}$  in figure 5.8B (94 “valid” points out of a total of 163). This led to a marginal increase in COE (-0.25) and decrease in mean absolute error (MAE =  $2.98 \text{ mm.day}^{-1}$ ), i.e. still indicating a poorer agreement than with the mean of the observed values (negative COE).

Furthermore, considering the difference in availability of “observed” evaporation

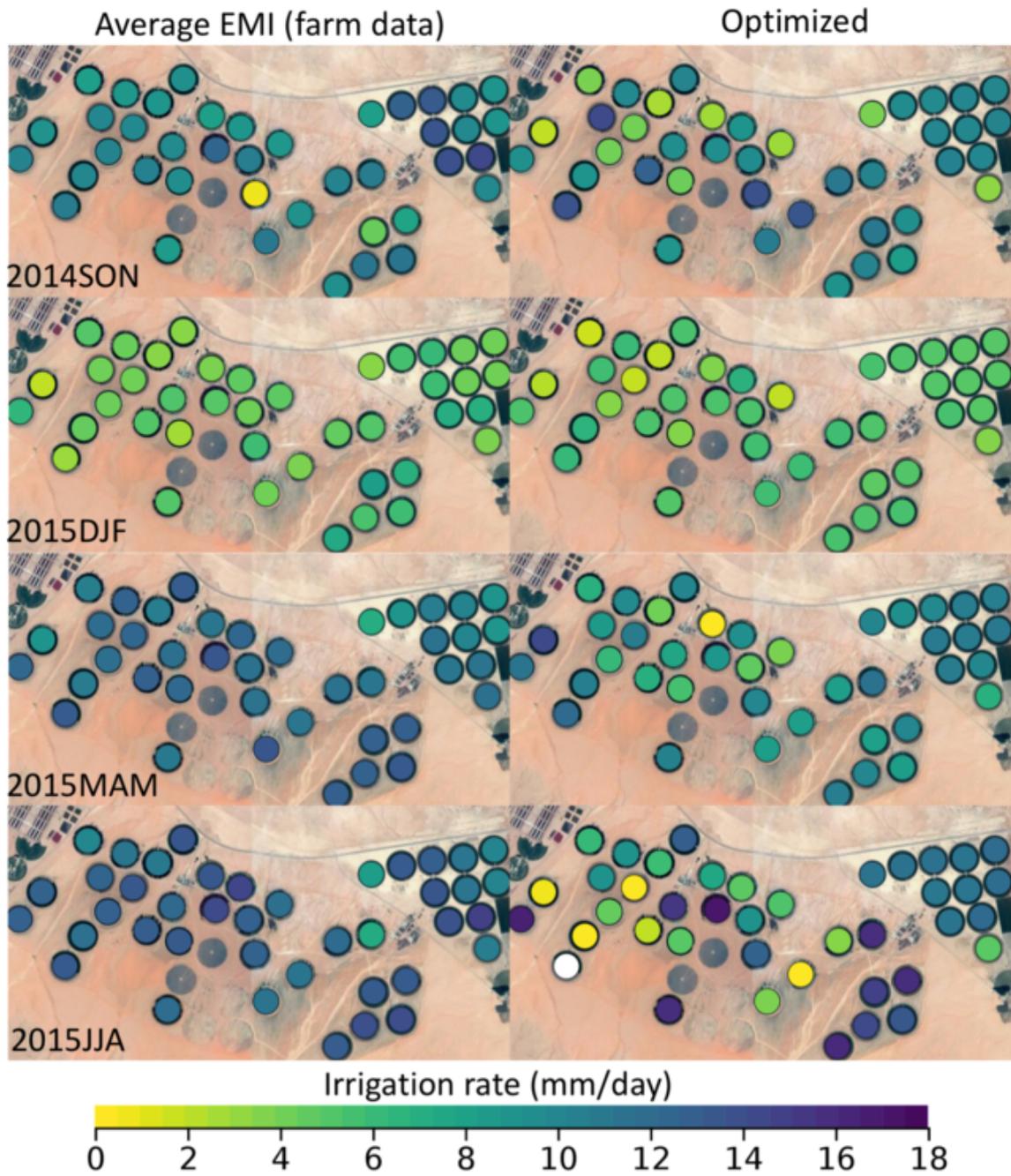


Figure 5.7: Comparison of optimized irrigation values (right) against farm data (average of each corresponding three-month period for each field).

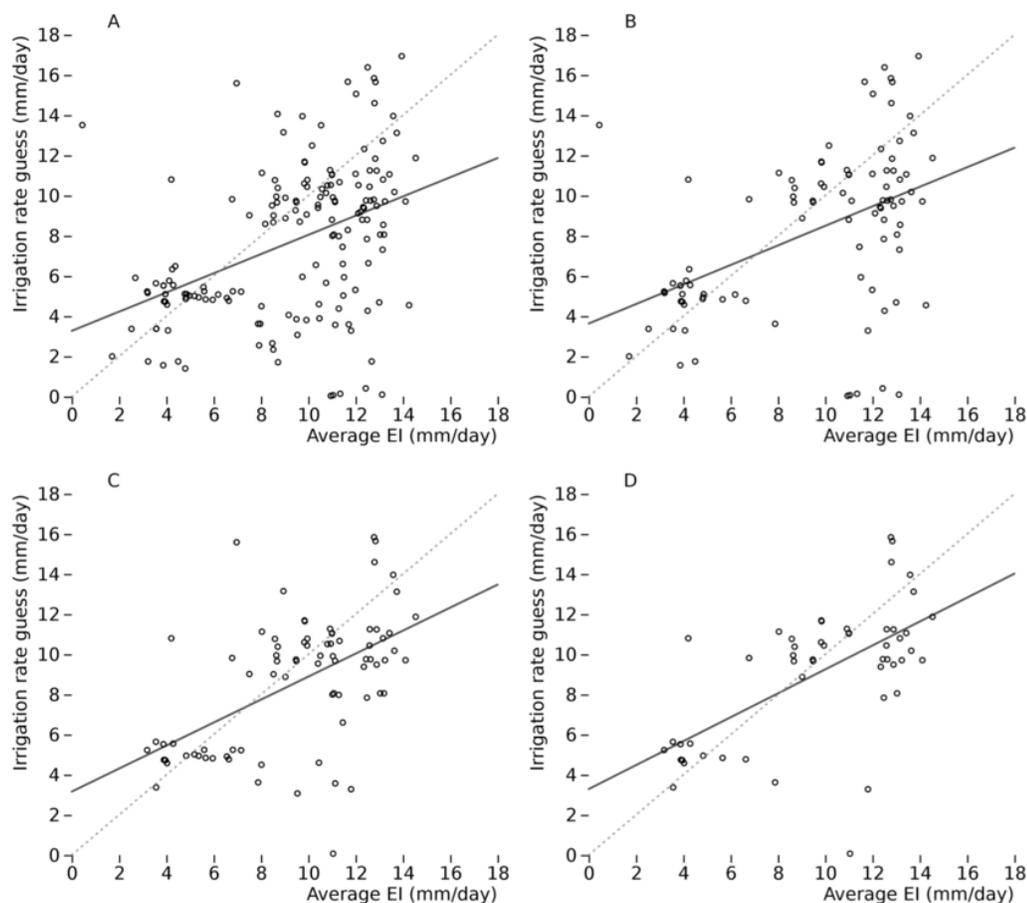


Figure 5.8: Comparison of optimized irrigation values against “observed” values (average of each corresponding three-month optimization period for each field). A: Data for all 41 fields during the four study periods. B: Data for all 41 fields during the four study periods, but omitting observed values with a large monthly variability ( $>5 \text{ mm}\cdot\text{day}^{-1}$ ). C: Data for the East fields (24-43) during the four study periods. D: Same as C, but omitting observed values with high variability.

data across the study site (figure 4.12), we repeated the analysis with two groups: 1) fields 1-23 (West) and 2) fields 24-43 (East; figure 5.8 C-D). There is a marked difference in spatial agreement between these two groups (table 5.1), with a poor agreement with the first group and a positive COE value for the second group after removing large-range values (COE = 0.26, highlighted in bold in table 5.1).

However, analyzing the individual optimization periods revealed poor agreement (negative COE values), even after removing values with large differences in the monthly farm data (figure 5.9). The statistical measures were consistently better

Table 5.1: Statistical measures obtained for the four study periods, grouped by A: fields with irrigation values  $<18 \text{ mm.day}^{-1}$ , and B: same as A but excluding values with range  $>5 \text{ mm.day}^{-1}$ . The analysis was split for half the number of fields in the West and East sides of the Tawdeehiya farm.

		<i>All</i>	<i>West(1-23)</i>				<i>East(24-43)</i>					
		<i>Study year</i>	<i>Study year</i>	<i>2014 SON</i>	<i>2015 DJF</i>	<i>2015 MAM</i>	<i>2015 JJA</i>	<i>Study year</i>	<i>2014 SON</i>	<i>2015 DJF</i>	<i>2015 MAM</i>	<i>2015 JJA</i>
A	<i>COE</i>	-0.56	-1.00	-4.00	-1.72	-22.8	-43.6	-0.02	-0.87	-0.12	-1.80	-4.76
	<i>MAE</i>	3.04	3.72	3.65	1.32	4.11	5.91	2.33	2.47	1.18	2.30	3.37
	<i>Bias</i>	-1.59	-2.28	-1.26	0.16	-3.60	-4.52	-0.87	-0.89	-0.17	-1.4	-1.02
	<i>n</i>	163	83	21	21	21	20	80	20	20	20	20
B	<i>COE</i>	-0.25	-0.72	NA	-1.09	-38.4	-68.9	<b>0.26</b>	-0.33	-0.81	-0.90	-4.46
	<i>MAE</i>	2.98	3.93	NA	1.29	3.71	5.90	2.22	2.62	1.08	2.34	2.67
	<i>Bias</i>	-1.24	-2.07	NA	0.22	-3.71	-4.05	-0.58	-1.07	0.60	-0.80	-0.92
	<i>n</i>	94	42	2	13	12	15	52	12	12	11	17

for the East group (where there are higher number of observations) than the West group (table 5.1). Overall, the optimized LSM results in mean absolute errors of  $3 \text{ mm.day}^{-1}$ , and is underestimating irrigation with an overall bias of  $-1.6 \text{ mm.day}^{-1}$ . However, during the winter months (2015 DJF), the LSM overestimates irrigation.

### 5.3.2 Analysis of hydrologic fluxes

Figure 5.10 shows the LSM results for daily evaporation during eight representative days (out of 34 for which there exist RS-E data). Overall there is a reasonable match in terms of temporal and spatial variability in evaporation between the optimized LSM results and the RS estimates (figure 4.17). However, the optimized LSM seems to overestimate evaporation during most of the year. Consequently, the LSM requires lower irrigation rates to achieve the observed evaporation rates, underestimating irrigation when compared to farm data. The overestimation of evaporation is reflected as well in the statistical measures: the comparison of field-average evaporation (figure 5.11) result in a coefficient of efficiency value of -1.41, a mean absolute error of  $2.1 \text{ mm.day}^{-1}$  and bias of  $1.49 \text{ mm.day}^{-1}$ . Similar values are obtained when computing these statistics for the East fields (24-43) alone. A range of values of COE

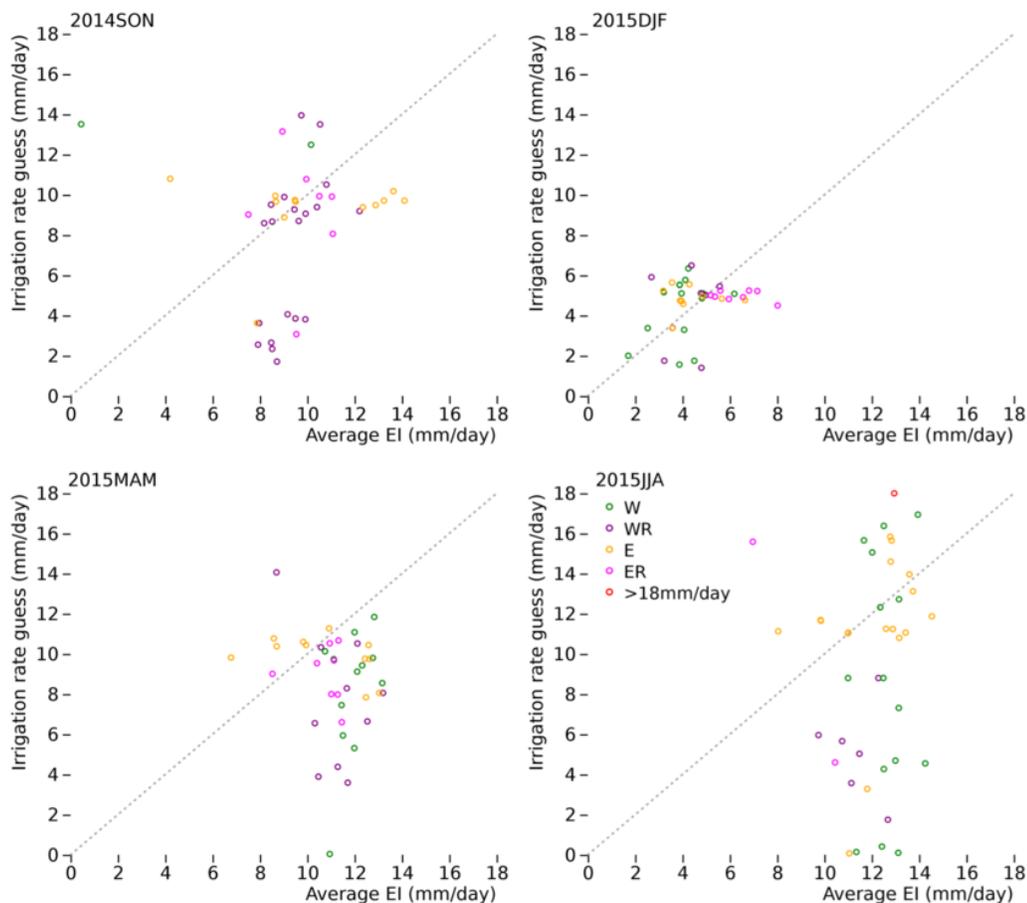


Figure 5.9: Comparison of optimized values and farm data (averaged over each three-month period) for each of the four study periods. E: East fields (24-43); W: West fields (1-23); R: values with more than  $5 \text{ mm.day}^{-1}$  difference between the three values considered for the average effective irrigation.

between  $-0.65$  and down to  $-14.6$  are obtained when comparing the individual daily evaporation values (34 days, e.g. figure 5.11 right), with mean absolute errors between  $0.57$  and  $4.15 \text{ mm.day}^{-1}$  and values for the bias from  $-0.57$  to  $4.15 \text{ mm.day}^{-1}$ ).

Figure 5.12 shows a group of six alfalfa fields in the East (fields 33, 34, 35, 38, 41 and 42) with similar LAI dynamics, irrigation patterns and observed evaporation responses during the study period. For these fields, the irrigation was underestimated most of the year (differences of up to  $4 \text{ mm.day}^{-1}$ ), particularly during the 2014SON period. Despite this, evaporation was still overestimated. Analyzing the partitioning of evaporation for this same period (figure 5.13), we can see that the soil evaporation

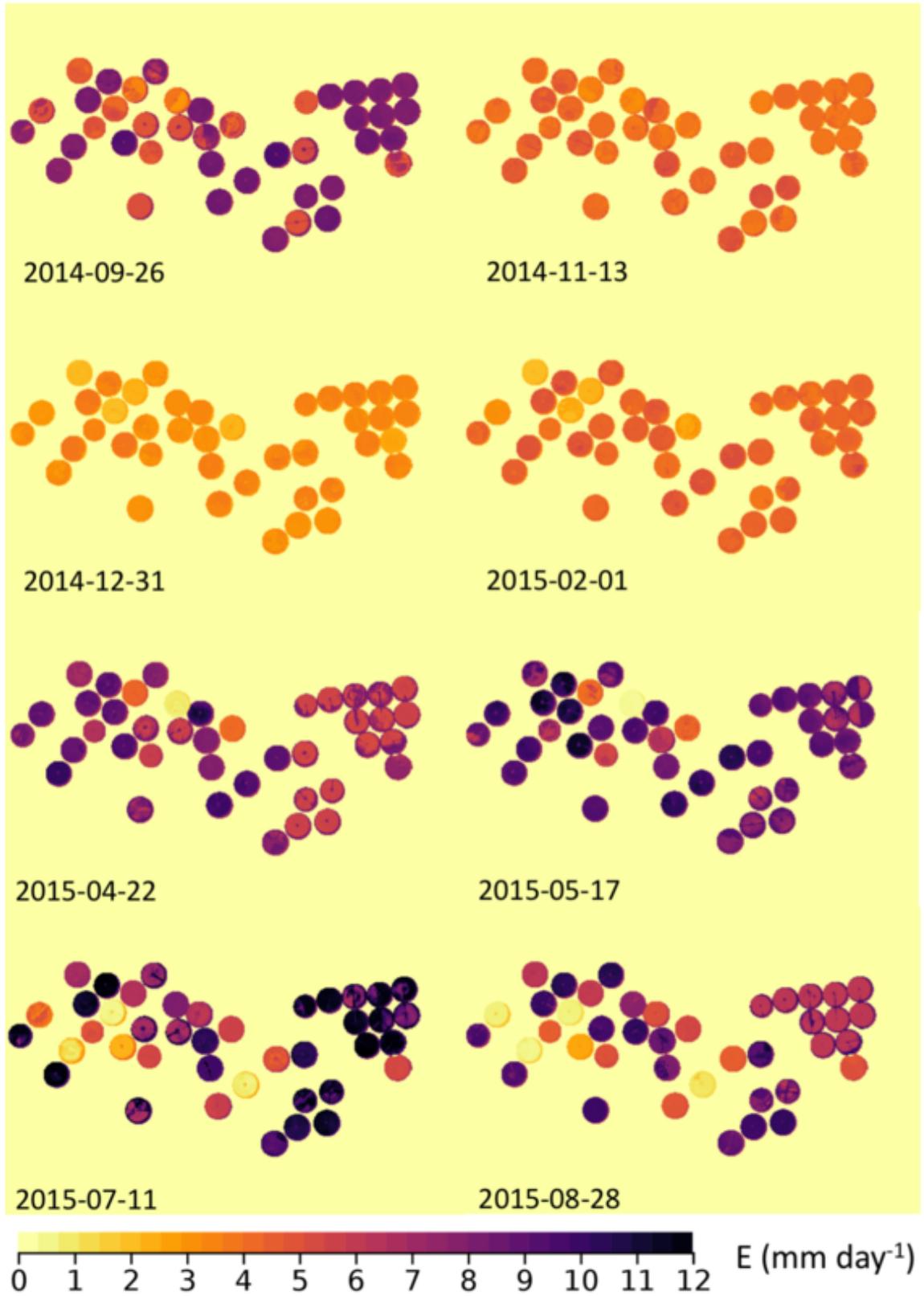


Figure 5.10: High-resolution daily evaporation maps obtained using the optimized model.

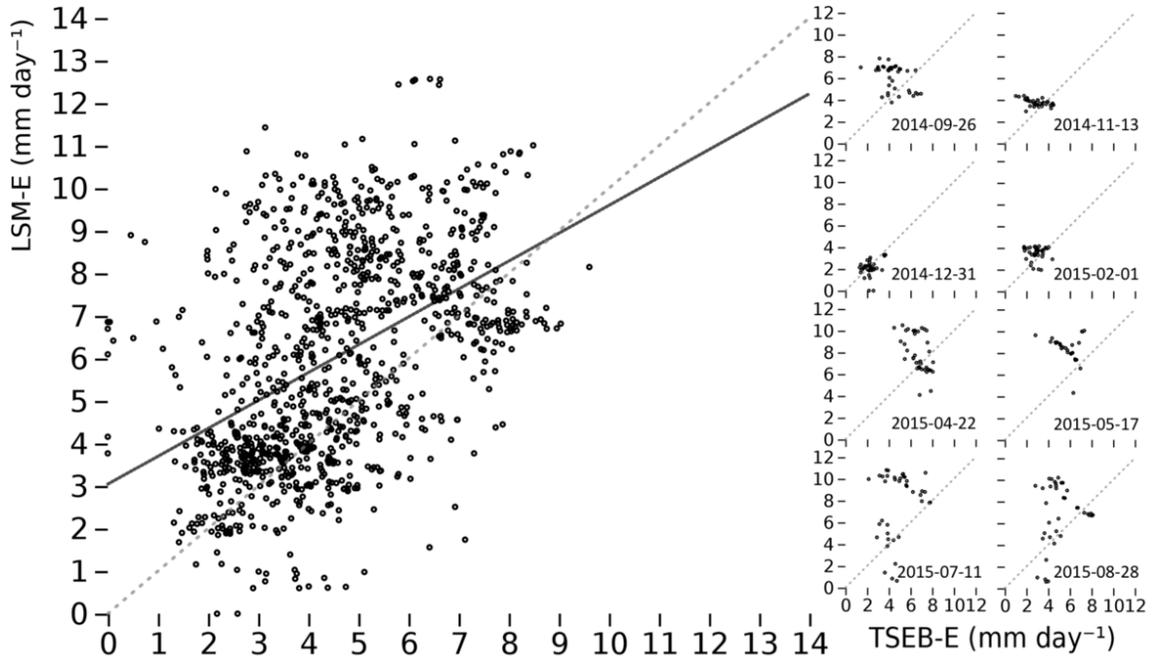


Figure 5.11: Left: scatter plot of average evaporation for each field (34 days) comparing RS-E data (i.e. from TSEB, see section 4.2.1) to the LSM results. Right: Comparisons for individual days shown in figure 5.10

(Soil-E) during September and early October was  $3\text{-}4 \text{ mm.day}^{-1}$  above the observed values. This also occurred during the summer between vegetation growing cycles, when LAI was lowest (and consequently transpiration and canopy interception were also low). However, from January to early April, soil evaporation (again the only source of  $E$  during these months) was closer to the observed values.

Another set of alfalfa fields with similar LAI and  $E$  dynamics included fields 36, 37, 39 and 40, located on the northeastern corner of the farm (figure 5.14). However, the long growing cycle that occurred from October to January in the previous group (figure 5.12), was shifted in time and occurred from November to February-March for this set of fields, coinciding better with the second study period (December 2014 - February 2015). Likely because of this reason, the irrigation and evaporation estimates during this period were closer to the observed values (compared to the previous group). On average, the first group underestimated the irrigation values by  $-1.16 \text{ mm.day}^{-1}$ , while the second group overestimated irrigation by about  $0.74$

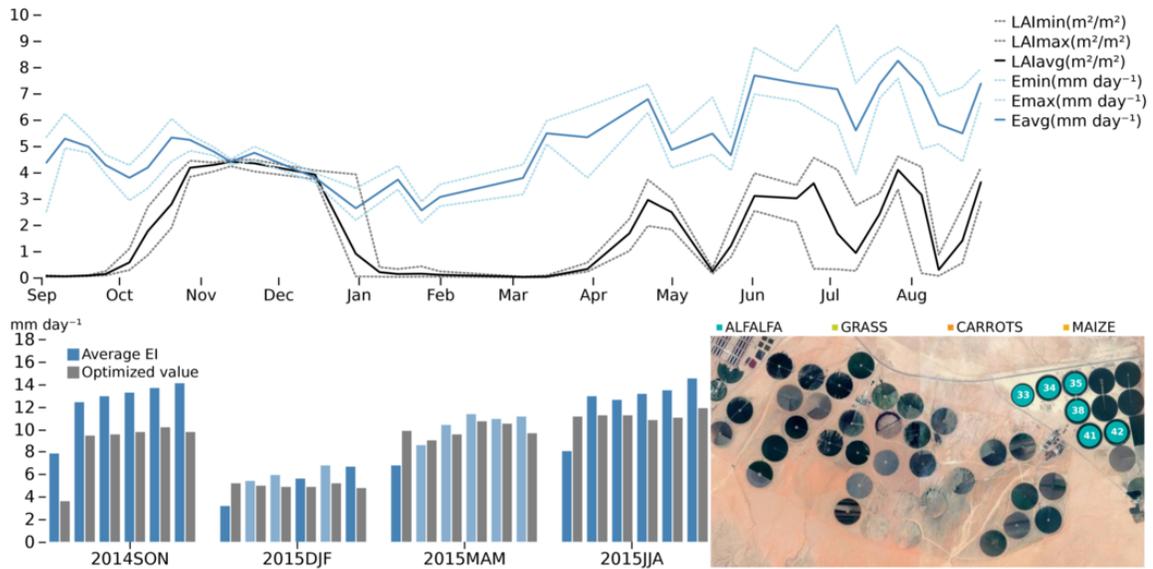


Figure 5.12: Top: LAI dynamics and observed evaporation (TSEB) response for fields 33, 34, 35, 38, 41 and 42. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (six fields). Bottom (right): location of the relevant fields.

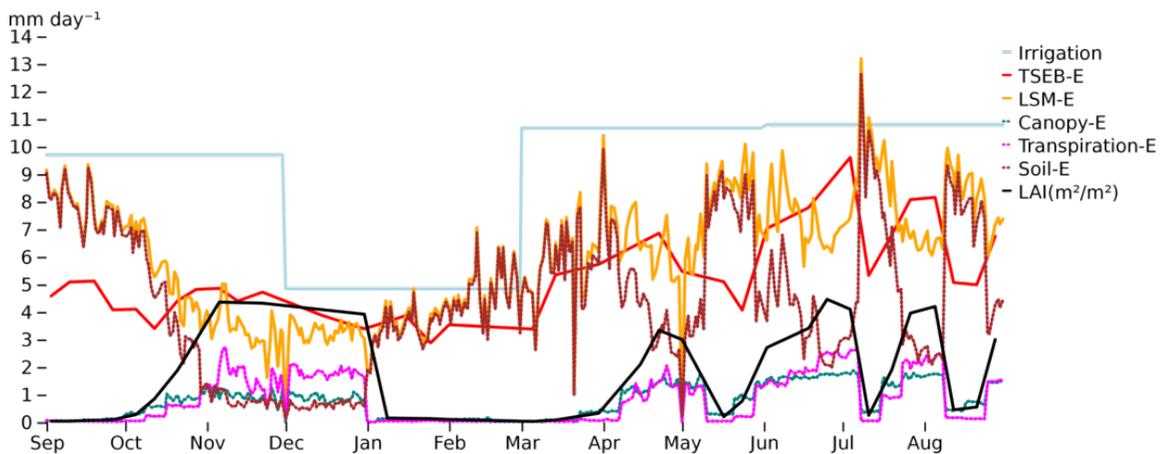


Figure 5.13: Evaporation partitioning from model and relationship to LAI (field number 38).

$mm.day^{-1}$ . However, the same overestimation of evaporation occurred during low LAI periods within the summer months, directly affecting the computed cost and therefore the effectiveness of the optimization algorithm during these periods (2014MAM and 2014JJA; figure 5.15).

Three more groups of fields with similar crop growing cycles, type of crop and

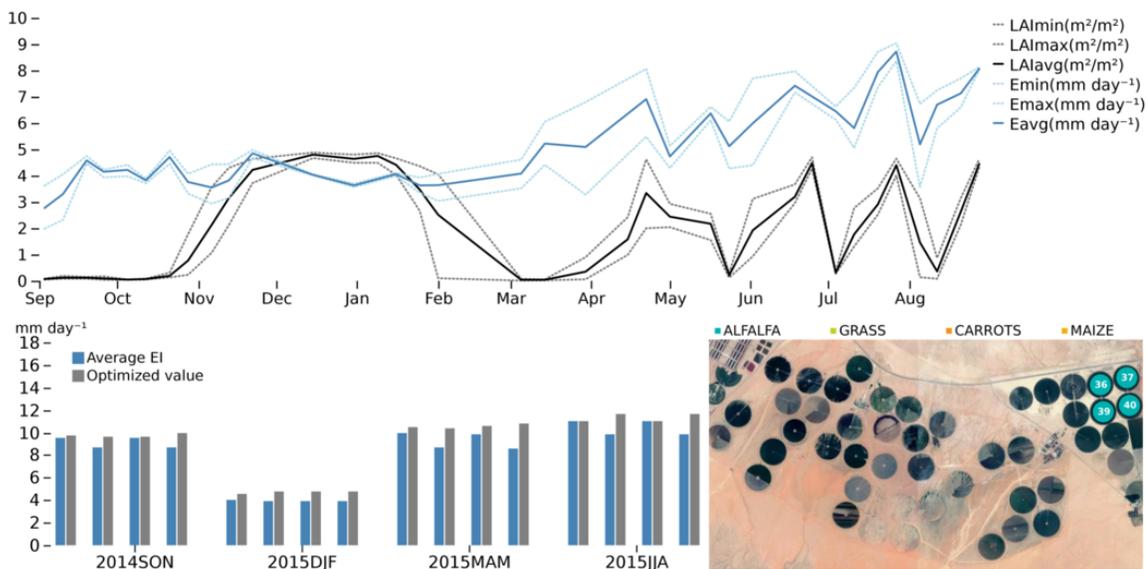


Figure 5.14: Top: LAI dynamics and observed evaporation (TSEB) response for fields 36, 37, 39, and 40. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (four fields). Bottom (right): location of the relevant fields.

observed evaporation were identified. First, a group of five maize fields (four located in the middle of the farm and one located in the eastern edge of the farm; figure 5.16) with similar (farm data-based) irrigation patterns (differences of at most  $2 \text{ mm.day}^{-1}$ ) showed large differences in optimized irrigation values, with average errors of  $-1.5$ ,  $0.4$ ,  $-4.9$  and  $-7.8 \text{ mm.day}^{-1}$  in the four study periods, respectively. The highest evaporation values coincided with a vegetation growing cycle from May - July 2015. However, another important peak in evaporation occurred previous to this cycle between April and May 2015, likely due to soil evaporation caused by irrigation previous to and during the early stages of vegetation growth. The model again overestimated the soil evaporation during low LAI periods, affecting the optimization of irrigation (underestimation) particularly during the last two optimization periods.

Next, a group of five maize fields was identified (figure 5.18). Irrigation differences with farm data averaged  $-2.2$ ,  $-0.8$ ,  $-3.6$  and  $-5.5 \text{ mm.day}^{-1}$ , respectively for the four study periods. However, despite the similar temporal variations in LAI,

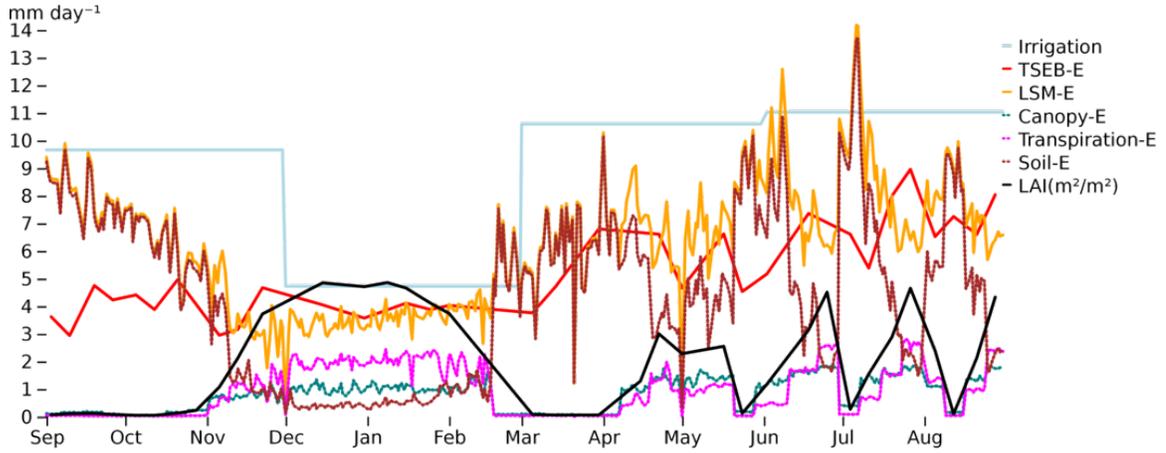


Figure 5.15: Evaporation partitioning from model and relationship to LAI (field number 39).

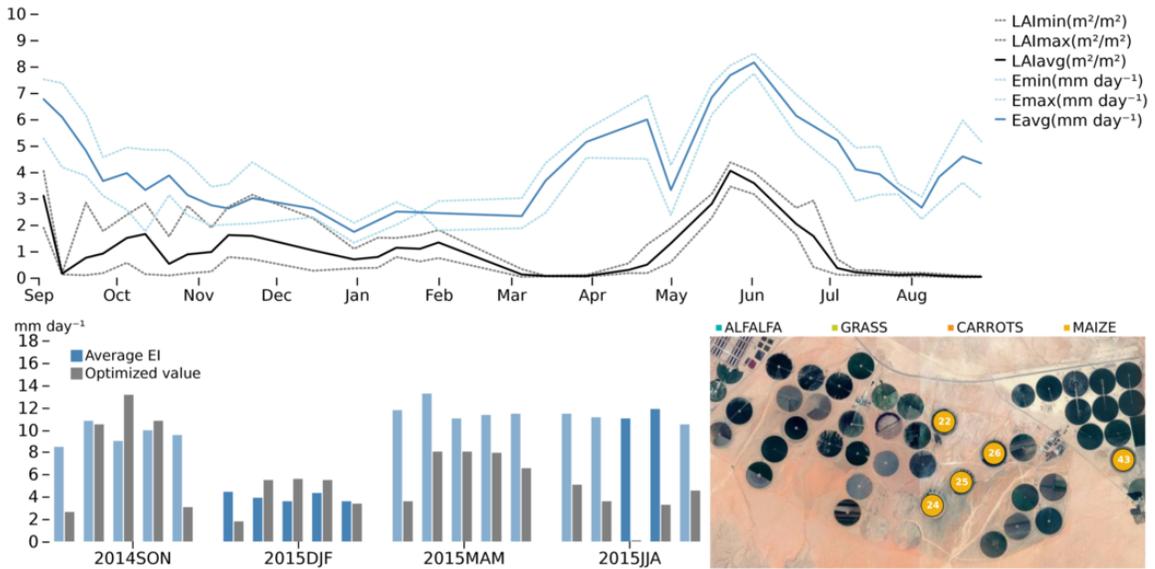


Figure 5.16: Top: LAI dynamics and observed evaporation (TSEB) response for fields 22, 24, 25, 26, and 43. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (five fields). Bottom (right): location of the relevant fields.

observed evaporation, and even location, there was a large variation in optimized irrigation values during the 2014SON period. As with the previous group, a large peak of evaporation was observed around late April 2015, before another one that coincided with an observed LAI cycle in late May 2015, the former likely being from soil evaporation prior to the start of the growth cycle. However, this soil evaporation

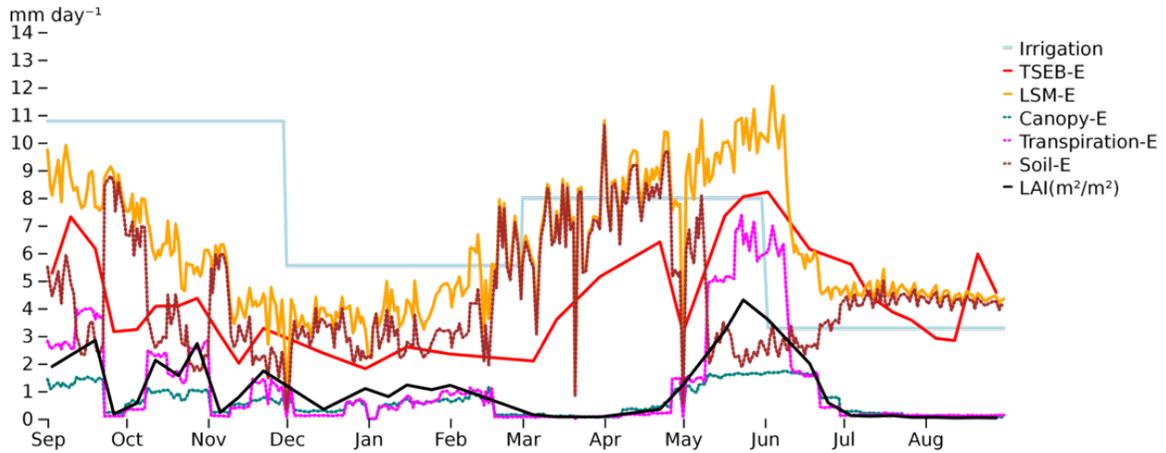


Figure 5.17: Evaporation partitioning from model and relationship to LAI (field number 26).

was overestimated, causing the irrigation to be underestimated.

Finally, a group of grass fields (located on the West) that shared similar LAI temporal variations and observed evaporation was analyzed (figures 5.20 and 5.21). Differences between the optimization irrigation results and the farm data are quite high, with errors averaging  $-2.8$ ,  $1.2$ ,  $-0.4$  and  $-5.6 \text{ mm.day}^{-1}$ , regardless of whether the monthly farm irrigation data had high ranges or not. Individual errors were much higher, in some cases overestimating by up to  $8.9 \text{ mm.day}^{-1}$  (field number 6; period 2015JJA) and in other cases diverging the irrigation estimation down to the lower constraint (i.e. close to 0) (fields number 2 and 7; period 2015JJA). This is shown in figure 5.21 for field number 7, where a sharp decrease in irrigation was predicted by the LSM from period 2015MAM to 2015JJA. In this case the system essentially withdrew water from storage in the soil (the change in water storage is shown in green), and consequently underestimated evaporation. However, the number of times this occurred (divergence of the optimization towards 0) was relatively low: seven out of a total of 164 optimizations (four study periods times 41 fields), i.e. about 4%.

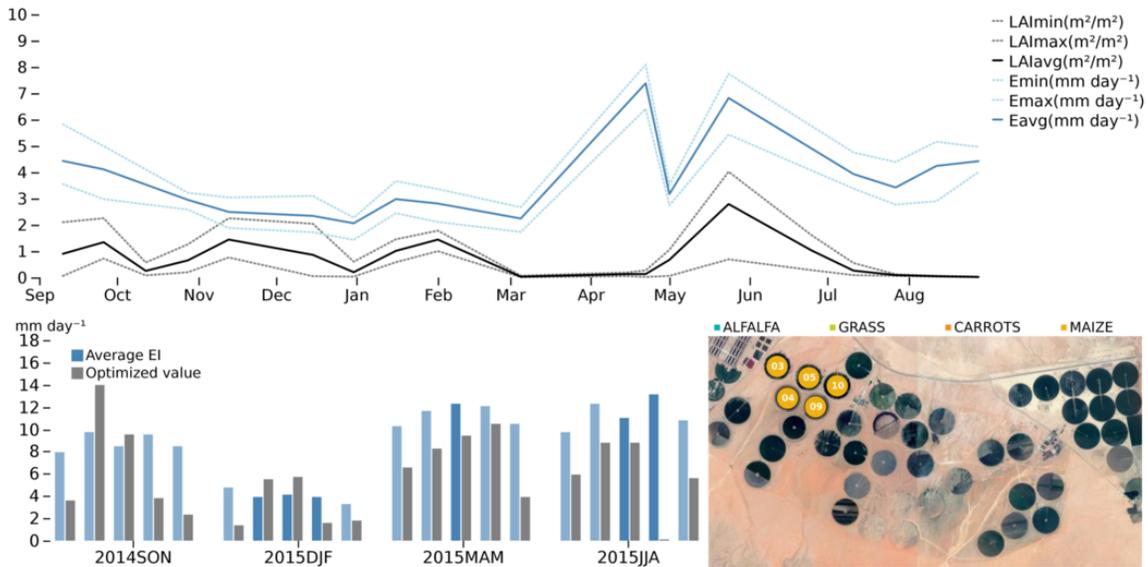


Figure 5.18: Top: LAI dynamics and observed evaporation (TSEB) response for fields 3, 4, 5, 9, and 10. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (five fields). Bottom (right): location of the relevant fields.

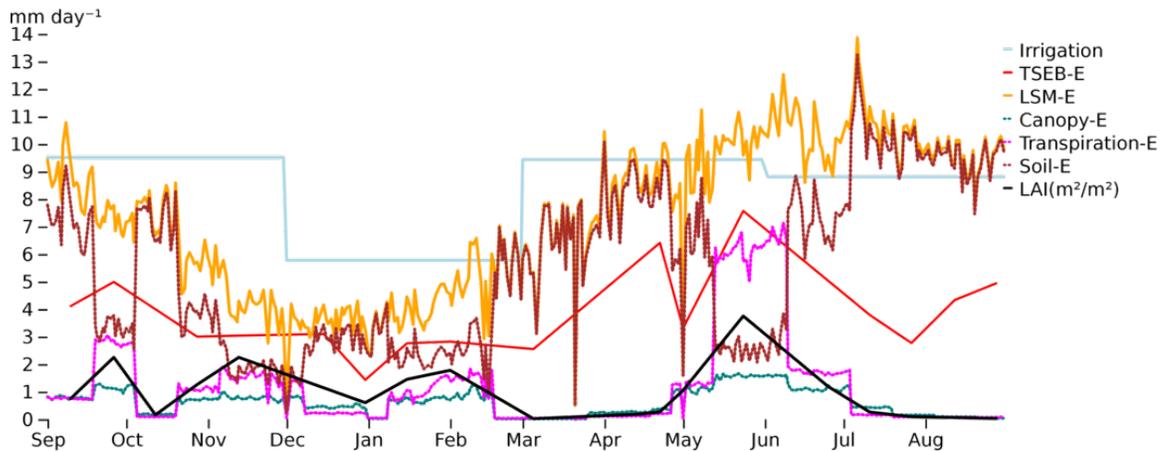


Figure 5.19: Evaporation partitioning from model and relationship to LAI (field number 5).

### 5.3.3 Estimated total water consumption

The optimized irrigation values for each field were scaled by the area (number of pixels  $\times 30 \text{ m}^2$ ), number of days in each period and by a factor of  $(1 \text{ m}/1000 \text{ mm})$  to get the amount of water used in  $\text{m}^3$  per three-month period. The total estimated water consumed by the Tawdeehiya farm is  $17.8 \text{ Mm}^3$ ,  $9.5 \text{ Mm}^3$ ,  $18.4 \text{ Mm}^3$ ,  $20.1 \text{ Mm}^3$ ,

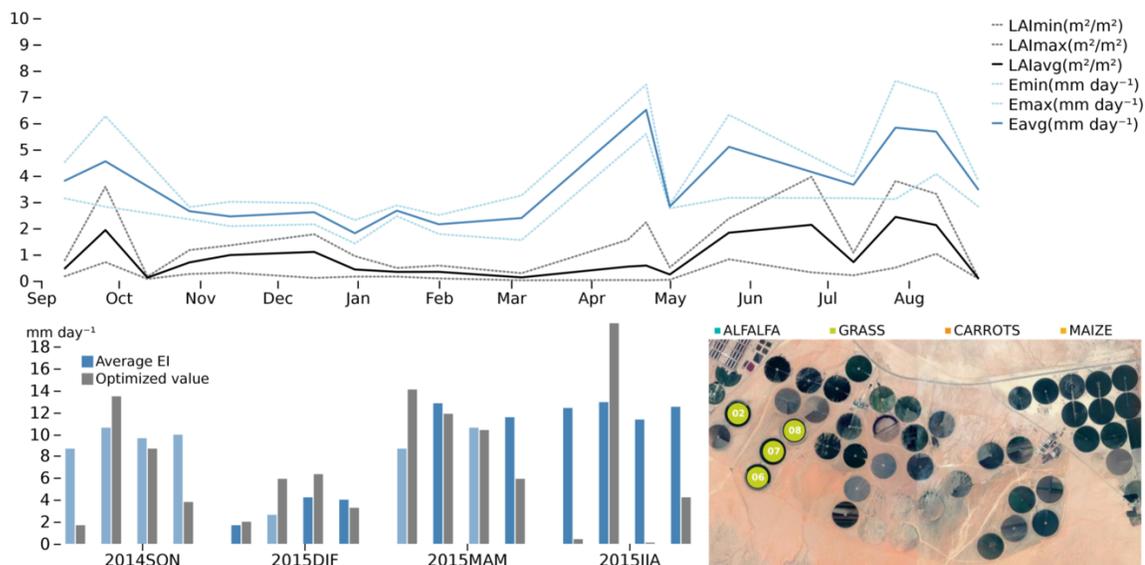


Figure 5.20: Top: LAI dynamics and observed evaporation (TSEB) response for fields 2, 6, 7, and 8. Bottom: Irrigation from average EMI values (values with large range are shown in light blue) and optimized results for the four study periods (four fields). Bottom (right): location of the relevant fields.

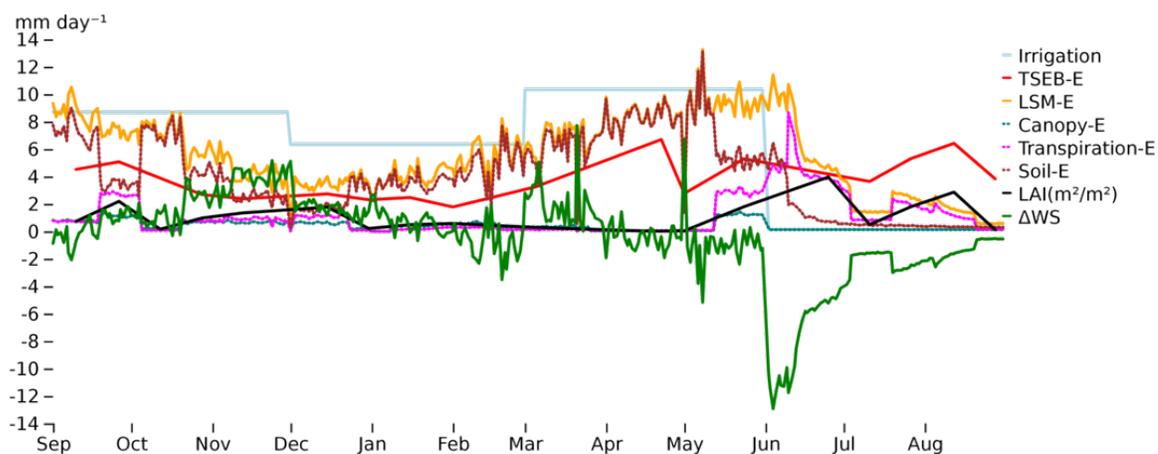


Figure 5.21: Evaporation partitioning from model and relationship to LAI (field number 7) and change in water storage (WS).

respectively for the periods 2014SON, 2015DJF, 2015MAM and 2015JJA; a total of  $65.8 \text{ Mm}^3 \cdot \text{yr}^{-1}$ , or about  $1.6 \text{ Mm}^3 \cdot \text{yr}^{-1}$  per field. By type of crop, the annual water consumption was  $35.3 \text{ Mm}^3$  for alfalfa fields (19, or  $1.85 \text{ Mm}^3$  per field),  $15.8 \text{ Mm}^3$  for maize fields (12, or  $1.31 \text{ Mm}^3$  per field),  $8.6 \text{ Mm}^3$  for grass fields (7, or  $1.22 \text{ Mm}^3$  per field) and  $8.15 \text{ Mm}^3$  for carrot fields (4, or  $2.03 \text{ Mm}^3$  per field). The annual

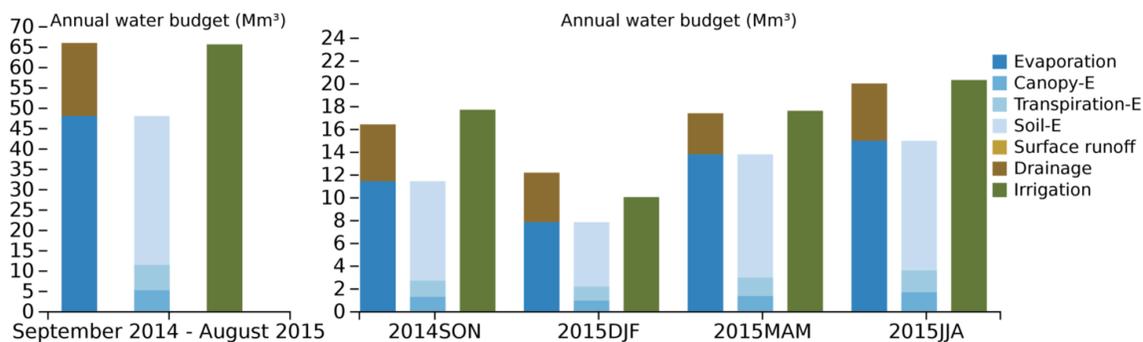


Figure 5.22: Total annual groundwater abstraction (millions of cubic meters) estimated using the optimization methodology for the entire Tawdeehiya farm, and partitioning of water fluxes.

water consumption calculated using the reported farm data for these 41 fields during the same study period is  $69.42 \text{ Mm}^3$ .

Out of the estimated  $65.8 \text{ Mm}^3$  of water irrigated throughout the year, about  $48 \text{ Mm}^3$  (or 72%) were estimated by the model to be evaporated, while drainage to the subsurface represented the remaining 28%, or about  $18 \text{ Mm}^3$  (figure 5.22). From the water being evaporated, the model attributed most of it to soil evaporation (76%), an unusually high fraction (see section 5.4.2), while vegetation transpiration and canopy interception contributed equally to the remaining 24%.

In section 4.7, the total evaporated water from 17 fields on the east side of the farm (27-43) was calculated for each of the available 34 days of TSEB observations using the land surface model with WRF rainfall as the only source of water input. Not surprisingly, the output evaporation of this model was significantly lower than the observed values (figures 4.18 and 4.16). Figure 5.23 shows the same comparison and includes the updated results using the optimized irrigation values. Clearly, this update matches the average daily water evaporation ( $58,832 \text{ m}^3$ ) more closely to the TSEB observations ( $47,180 \text{ m}^3$ ) than the model run with meteorological rainfall ( $3,225 \text{ m}^3$ ). However, as it was seen on individual field water budget analyses, the model seems to overestimate evaporation during most of the summer and fall months.

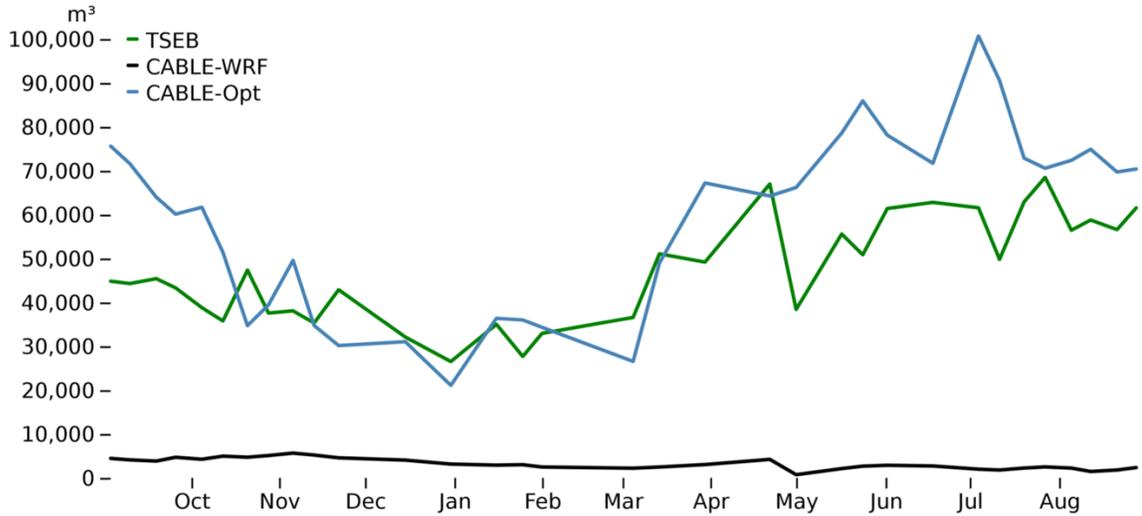


Figure 5.23: Comparison of total daily water evaporation for 17 fields (27-43) using 1) only meteorological rainfall (section 4.7) and 2) optimized irrigation values.

## 5.4 Discussion

### 5.4.1 Tradeoff between selection of optimization periods and number of observations

Ideally, the number of satellite observations should be high and uniformly distributed in time. This would allow the selection of optimization periods that would best match distinct vegetation growth cycles and consequently aim to obtain realistic irrigation dynamics. However, in reality the availability of observations in this study was constrained by the number of Landsat scenes available, cloud cover (figure 4.11), and the effectiveness of statistical techniques used to retrieve some of the required inputs for the RS-E model (section 4.5.1). Furthermore, because of the unique locations of the study site relative to Landsat's path, the number of observations available for each of the 41 fields within the farm was not uniform (figure 4.12). These factors complicated the selection of appropriate periods for optimization.

As a first attempt to get a rough estimate of irrigation patterns and to apply the same general methodology to all 41 fields, quarterly periods (three-month) were

ultimately selected. Upon inspection of figure 4.12, which shows the number of observations by study period and ordered by field number (figure 4.9), it is evident that about half the fields had more than double the number of observations than the other. The analysis performed on the validation of optimized irrigation values separately for half the fields on the West and East (table 5.1) revealed that the number of observations indeed had some influence on the effectiveness of the optimization methodology (e.g. reflected in a higher COE value).

### 5.4.2 Accuracy of the observations and retrievals

The success of the methodology implemented in this study relies on the accurate retrieval of evaporation based on satellite data. Current ongoing work aims to validate the evaporation retrievals from the TSEB model using tower flux data on the Tawdeehiya site. However, there are a number of challenges that likely affect the estimation of evaporation in this region. First, the values of shortwave radiation predicted by WRF (figure 4.4) appear to be quite high and needs to be assessed, requiring the use of ground-based data. Second, the partitioning of evaporation into bare soil and vegetation transpiration in TSEB is challenging in arid environments. Advection effects, which are strong in a constantly irrigated field, directly impact the partitioning of evaporation. There have been recent efforts to account for these effects, although requiring the use of additional soil moisture data (*Song et al.*, 2016). Further work is also required to evaluate the sensitivity of TSEB to LST values, which directly impact the estimation of sensible heat flux. Finally, the impact of the uncertainty of other assumed constant parameters (e.g. the impact of emissivities  $\epsilon_s$  and  $\epsilon_c$  on the partitioning of  $T_S$  and  $T_C$ ) needs to be evaluated as well.

Furthermore, there are challenges in validating actual groundwater abstraction rates. First, wind can carry a significant amount of water being sprayed, with losses of up to 20% (*Steiner et al.*, 1983; *Sadeghi et al.*, 2017). Next, the data used to validate

the irrigation rates does not necessarily represent the true total abstraction rates. The reported farm data consists of an assumed constant irrigation rate throughout the year, multiplied by the number of hours per month at which the irrigation was applied. This “effective monthly irrigation” rate was then averaged over each optimization period, again assuming a constant value. During periods with a high range of EMI values, it is therefore not expected that the optimized irrigation rates would closely match these rates.

### 5.4.3 Optimization methodology

The choice of parameters  $a$  and  $c$  for the SPSA algorithm had a great impact on the effectiveness of the optimization (section 5.2.2). An attempt was made to find a set of parameters that would work for the majority of the fields. This effort consisted in a “trial and error” exercise where several sets of parameters were tested. This tuning exercise was performed with conditions as close as possible to the actual implementation: the same temporal resolution of optimization periods, and the same frequency of “observations” (synthetic data, see section 5.2.2) both spatially (i.e. distribution of observations among the fields) and temporally (i.e. selecting daily evaporation snapshots only on days when the actual observations are available). Under these conditions, the variability of EMI values also had an impact on the optimization for a number of fields (e.g. figures 5.2D and 5.5).

Although no single set consistently worked for all fields within the study site, performing the tuning experiment on all fields helped to evaluate the performance of the different sets in the context of the whole study site. In generalizing the methodology to other agricultural sites, it is important to note that the selection of tuning parameters might be site-specific, and likely dependent on the length of optimization periods as well as the frequency and temporal distribution of observations.

The sensitivity of the methodology to the initial guess was also evaluated. Al-

though there were visible differences when varying the initial guess, the overall agreement between the optimized irrigation rates and the average EMI values was similar in most cases (figure 5.6 B-D). However, upon lowering the initial guess further, the optimization diverged towards unfeasibly high values for the majority of the pivots (figure 5.6 A). Therefore, when applying the methodology to other regions, identifying when such cases would occur would be relatively easy if applied to a large number of fields: either most of the values will diverge towards an unrealistically high or low value, or stay close to the initial guess. In the actual implementation of the methodology using satellite observations of evaporation (section 5.3), the number of times that this happened was relatively low: only one optimization led to a value higher than  $18 \text{ mm.day}^{-1}$ , while only seven extremely low values of irrigation were obtained.

Evaluating the efficiency of other optimization algorithms and the impact on the solution is a line of work that could be undertaken. In this work, an initial attempt to use Matlab's default optimization function (*fminunc*), based on the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, a type of quasi-Newton method (*Fletcher, 1994*) was done. Upon efforts similar to those undertaken in section 5.2 (e.g. by modifying the optional tolerances and magnitude of the perturbation to the parameters), the irrigation rates diverged to unrealistically high values. Further investigation is needed to address this issue and attempt to evaluate the use of other algorithms as well. However, the priority of current work is to evaluate and implement the use of higher frequency data into the methodology.

#### 5.4.4 Potential improvements to the methodology

Given this first attempt to infer agricultural groundwater abstraction rates from remote sensing observations, there are a number of aspects of the methodology that could be improved. First, it was shown that the frequency of observations within each optimization period greatly influenced the estimation, as evidenced by the difference

in errors obtained for the fields located on the East and West sides of the study site. Therefore, it would be of great value to incorporate other potential sources of data that could increase the frequency of observations, including data from other space agency missions such as the Global Monitoring for Environment and Security (GMES) Sentinel missions (*Donlon et al.*, 2012), as well as the use of constellations of small, lightweight satellites (e.g. CubeSats) that can provide ultra-high resolution (3 m) and increased temporal frequency (*McCabe et al.*, 2018). However, further work on the validation of these systems is needed.

One aspect specific to the land surface model used in this study (CABLE), was the apparent bias in soil evaporation. The model seemed to output significantly higher rates of soil evaporation during summer and fall months. However, this only happened between vegetation growing cycles, i.e. when LAI was close to zero, but it had a great impact on the performance of the optimization. Overestimation of soil evaporation has been reported in several studies, although in different environments than the work presented here (e.g. *Decker et al.* (2017); *Wang et al.* (2011)). *Haverd et al.* (2016) introduced improvements to CABLE aimed to reduce the errors associated to soil evaporation. In this study, we used an updated version of CABLE which includes these modifications to the soil module. However, given the extremely different environment and vegetation cover, this issue requires further investigation.

A possible alternative to mitigate this problem would be to update the soil moisture state using remote sensing observations. This would also have the additional advantage to help initialize the model with a more realistic state (section 4.6). Although several soil moisture satellite-based products are available (*Petropoulos et al.*, 2015), a product with spatial and temporal resolutions similar to the evaporation observations would be preferred, or with sufficient resolution to at least distinguish the signal between the different fields and the bare soil. Currently, these do not exist operationally. A distinction of at least surface soil moisture and root-zone moisture

would be preferable as well. Further investigation of the best way to incorporate these observations into the different soil layers in CABLE would also be needed.

## 5.5 Summary and conclusions

The work presented here represented a novel effort to capture groundwater abstraction rates used in irrigated fields through the use of remote sensing data and a land surface model. The method consists of attempting to match the evaporation output from the land surface model to the observed from satellite data, using the irrigation rates as the variable to optimize. A first implementation of this methodology was applied on a desert agriculture environment (41 center-pivot fields) on a quarterly basis for a full year (September 2014 - August 2015).

The number of observations within each field had the greatest impact on the method results, reflected in the marked difference in model agreement between fields located on the West and East of the farm. The method generally underestimated irrigation  $-1.6 \text{ mm.day}^{-1}$ , except during winter months. However, because the data used to validate the results does not necessarily represent the true irrigation patterns (section 5.4.2), the interpretation of the statistical measures (table 5.1) must be undertaken with care. The evaporation predicted by the optimized model was generally overestimated. Upon inspection of individual field time series, it was shown that this was greatest during summer and fall months between vegetation growing cycles (LAI close to 0).

Despite these numerous challenges, this study presents a first rough estimate of total annual water consumption:  $65.8 \text{ Mm}^3$ , or about  $1.6 \text{ Mm}^3$  per field, highest during the summer ( $20.1 \text{ Mm}^3$ ) but closely followed by spring and fall ( $18.4 \text{ Mm}^3$  and  $17.8 \text{ Mm}^3$ , respectively). Average annual consumption per type of crop was:  $1.85 \text{ Mm}^3$ ,  $1.31 \text{ Mm}^3$ ,  $1.22 \text{ Mm}^3$  and  $2.03 \text{ Mm}^3$  respectively for alfalfa, maize, grass and carrot fields. However, most of the grass fields were located on the West side of

the farm, while the estimate for carrots was based on four fields only, therefore these two values might be less reliable than the first two. The model estimated that most of the water was evaporated (72%) while the remaining drained to the subsurface, representing an important management insight that may improve the sustainability of irrigated agriculture in these environments. Because of the observed bias in the soil evaporation predicted by the model, actual drainage might be higher. No surface runoff was predicted by the model.

A number of future research directions might help to improve upon these first estimates. First, the frequency and uniformity of evaporation observations could be improved by incorporating data from other satellites, e.g. from commercial and other space agency missions. Furthermore, in order to address the issue of soil evaporation bias and to improve model initialization, soil moisture observations could be incorporated into the methodology. Current efforts to mitigate the issue of high bare soil evaporation include implementing constraints on irrigation based on observed LAI (e.g. irrigation shut-off periods when no vegetation is present for a certain amount of time). However, the cause of overestimation in CABLE in this specific environment will also need to be investigated further. Finally, the use of other land surface models could be implemented in order to improve the accuracy of the irrigation estimates by using an ensemble mean.

## Chapter 6

### Concluding Remarks

---

This last chapter presents a brief discussion of the main contributions of this work, including the challenges involved and conclusions reached. It also presents the recommendations for further application of the method to provide a meaningful estimate of the total agricultural water use in the region and its evaluation.

---

The work presented in this dissertation describes an effort towards monitoring agricultural water use using a combination of satellite observations, available *in situ* data and land surface modeling approaches. Remote sensing data derived from satellite platforms presents a unique opportunity to capture information relating to vegetation growth and changes in the land surface condition (i.e. water and energy exchanges). It has the advantage of spatial coverage and continuous monitoring over traditional ground-based observations. However, there is still a need for *in situ* measurements to validate these observations. Currently, there is no active and wide-spread ground-based monitoring of water use in Saudi Arabia. Even with advances in modeling and remote observation, such ground based information is essential in delivering improved water management strategies and in order to ensure continuous availability of the limited water resources in the region.

To explore the potential use of remote sensing and land surface models to monitor agricultural water use and evaluate its relation to large-scale hydrologic fluxes, the study was focused upon two complementary perspectives: the first part of the work (chapters 2 and 3) was dedicated towards detecting large-scale and long-term changes in the hydrological state of the region, while the second part (chapters 4 and 5) focused on developing and testing a methodology to measure water consumption from agriculture in order to quantify the anthropogenic impact on the water resources in the region. Here, we briefly review these dissertation elements and highlight the key outcomes.

## **6.1 Large-scale estimates of water depletion and evaporation from satellite observations**

Over the last 14 years, the Gravity Recovery and Climate Experiment satellites provided information required to detect changes in the Earth's gravity. This unprecedented information resource allowed hydrologists to study large-scale changes in the

terrestrial water storage. The techniques required to relate these two phenomena were described in chapter 2. While there is a global product available to provide end-users information on water storage changes without specific knowledge on these techniques (*Landerer and Swenson, 2012*), certain regions present additional challenges that limit the use of this product. Two nested large-scale regions were analyzed in this study: the Arabian Peninsula and the Saq Aquifer System (SAS), where these retrieval challenges were illustrated. Using the standard methodology to process GRACE spherical harmonic coefficients (*Swenson and Wahr, 2002, 2006*), a long-term depletion trend of  $4 \text{ mm.yr}^{-1}$  and more than  $10 \text{ mm.yr}^{-1}$  were found for these two regions respectively (chapter 2). This compares to the global gridded product that showed no significant trend. The detected depletion trend was linked to increased agricultural activity in the region, which was illustrated using data from a moderate resolution land cover product based on MODIS (*Friedl et al., 2010*) and high resolution Landsat data. In order to establish a more definitive and quantifiable causal relation, an analysis of related components of the water cycle, particularly evaporation, was required. In doing so, the ability of these independent data to represent hydrological cycle components was also evaluated.

As an example, recent and ongoing efforts have aimed to develop large-scale estimates of evaporation with global or continental coverage. While these products are usually evaluated by comparison with ground measurements, these are not uniformly available everywhere, and the assessment is limited by the scale misrepresentation of the ground observations. Inter-product evaluations, often performed by including other types of estimations (e.g. output from land surface models and climate reanalysis; *Mueller et al. (2011)*), provide another form of product evaluation. However, neither of these type of assessments provide insight into the connectivity or relationship that evaporation might have with other hydrologic component data.

An alternative evaluation strategy was undertaken by using the concept of hydro-

logical consistency, which aims to show how “comparable” independent hydrologic data are. By focusing on regions in which outflow components are limited to mostly evaporation, the hydrologic consistency was evaluated by comparing the spatial distribution of large-scale water storage changes (i.e. those observed by GRACE), with the difference between precipitation and evaporation. The results showed a lack of persistent agreement between these independent sources of data, hindering efforts to discriminate between individual products. Due to the tradeoff in resolution, model assumptions and robustness of the global products, some factors affecting the water system are not represented accurately. Changes to land cover, extensive groundwater abstraction, and surface water management (e.g. reservoirs and aqueducts) are some anthropogenic sources of these additional influences to the water system. Efforts attempting to capture and incorporate these impacts require the use of hyper-resolution data, rather than the coarse scale retrievals commonly available from such global products.

## **6.2 Measuring agricultural water consumption using hyper-resolution remote sensing data and land surface modeling**

In concert with model developments that aim to incorporate additional interactions in the water system (e.g. WaterGAP 2; *Alcamo et al. (2003)*), there have been parallel efforts towards collecting hyper-resolution data to drive these models. In order to take advantage of this data, novel methods are required. The focus of chapter 5 consisted of developing a methodology to incorporate evaporation estimates derived from high-resolution remote sensing data into a modern land surface model, so that realistic hydrologic interactions from agricultural development could be detected. The overriding goal was to calculate total water consumption as a response of agricultural activity. The study was focused on a relatively small farm located in an arid environment, southeast of Riyadh in Saudi Arabia (section 1.2). Although demonstrated on

a relatively specific land surface and cover type, the framework was designed to be transferable to other dryland agricultural areas. Chapter 4 described the sources of data needed to drive both a remote sensing-based evaporation model (RS-E) and a land surface model (CABLE), as well as the details of these models. Compared to the evaporation model, CABLE, in its current formulation, needs additional information (e.g. irrigation input) in order to capture the evaporation output from the agricultural fields. The RS-E model provided more realistic values in terms of expected evaporation based on the presence of vegetation. However, the land surface model, if given an accurate representation of the water inputs, has several advantages over the RS-E model. It can provide estimates at more frequent intervals (daily and even sub-daily), and calculate additional water fluxes including drainage to the subsurface.

An optimization methodology and its application to 41 individual fields in the study site was described in chapter 5. Taking advantage of the low to no rainfall in the region, the precipitation component in the land surface model was iteratively updated to incorporate agricultural irrigation (capturing the groundwater abstractions). The update was constrained based on the difference between the evaporation output from the land surface model and the satellite-based model. The optimization was applied on a quarterly basis for a full year (September 2014 - August 2015), resulting in a first estimate of the total annual water consumption of about  $65.8 \text{ Mm}^3$  (compared to a total of  $69.42 \text{ Mm}^3$  using farm data), 28% of which is estimated to be drainage to the subsurface. A preliminary evaluation of the approach was undertaken using estimated irrigation rates derived from reported flow data, and was shown to generally underestimate irrigation ( $\text{MAE} = 3 \text{ mm.day}^{-1}$  with bias of  $-1.6 \text{ mm.day}^{-1}$ ), likely due to the LSM overestimation of soil evaporation, as well as the loss of sprayed irrigated water carried by the wind. There was a marked difference in agreement between fields that benefited from a higher number of observed data. However, the data used in the evaluation does not necessarily represent the actual irrigation rates.

Further investigation is required to improve the accuracy of the irrigation estimates as well as to provide uncertainty bounds prior to its general application in other agricultural areas in the region. However, this approach provided some new insights into agricultural system behavior and offers a new technique for monitoring such systems in the absence of *in situ* data.

### **6.3 Directions and recommendations for further research**

While the study offered a first approximation of the total annual water consumption at the study site, it should not be treated as a final estimate. There is certainly room for improvement on the methodology, while some limitations specific to the selected land surface model need to be addressed. First, the use of a single main source of data for characterizing the land surface parameters (i.e. Landsat) was done to simplify the methodology in its developmental stage. This limited the frequency and temporal uniformity of observations, influencing the selection of the optimization periods. Because of the relation between the location of the study site and Landsat's path, the study benefited from a unique perspective in the sense that about half the fields (west) had about half the number of observations than the rest. Other sources of data, including recently launched missions from space agencies such as Sentinel-2 satellites (*Donlon et al.*, 2012), and commercial satellites (e.g. Planet Labs; [www.planet.com](http://www.planet.com)), can be used to provide a more uniform and frequent coverage of information. This will provide the capacity to perform the optimization with consideration of the vegetation growth cycles. However, this would involve efforts to validate these different sources of data and determine the best way to incorporate them into the analysis.

Furthermore, the methodology can be applied using a number of different land surface models, so that the impact of any individual bias in the models be reduced, and at the same time provide uncertainty bounds to the estimates. The study showed

that although recent improvements in the representation of soil evaporation (e.g. leaf litter impacts on soil evaporation (*Haverd et al.*, 2016) and improved sub-grid scale soil moisture parameterizations (*Decker et al.*, 2017)) have been implemented to CABLE, there remains a significant overestimation of this component, likely related to the extreme conditions (e.g. temperature, solar radiation and possibly soil properties) of the region. In the meantime, the incorporation of soil moisture observations might assist in mitigating this issue, while also providing useful information for the initialization of the model. Further work on extracting high-resolution soil moisture from passive and active sensors is an area of ongoing research with collaborators. Another recommendation that could potentially reduce the impact of soil overestimation is to perform the optimization on some of the soil module parameters as well.

Finally, the evaporation estimates retrieved from the thermal based model also require validation. While there are current efforts to do this using flux tower measurements on site, a potential line of research that can be explored is the impacts of uncertainties in the evaporation retrievals on the optimized irrigation estimates. This can be done by first perturbing the evaporation observations directly and applying the optimization methodology, and then by perturbing some of the inputs into the satellite evaporation model (e.g. LST).

At the larger scale, estimates on the long-term trend in total water storage depletion from GRACE can be used as a benchmark to evaluate the total water abstraction in the region. While the GRACE mission ended in October 2017, a follow-on mission is set to launch early in 2018, thereby continuing the ability of retrieving independent measurements of large-scale water depletion. Therefore, there is motivation to explore further the application of the methodology developed here to retrieve water use from all agricultural fields within the region (e.g. Al-Jowf, Buraydah, Wadi Ad-Dawasir), a task that will demand a comprehensive effort in mapping the land surface parameters, phenology and dynamic vegetation response. However, further work will

also be dedicated towards automating these processes, with the goal of providing key information to water management agencies on a regular basis. This will be the only viable means of providing a detailed link between the local water abstractions and the large-scale changes of water storage in the region.

This work represented an effort to relate independent retrievals of hydrologic fluxes observed at different scales, with the ultimate goal to monitor the amount of water used in irrigated agriculture. A first rough estimate of annual water consumption from an agricultural site in Saudi Arabia was obtained by using a methodology developed in this study. Further work to retrieve groundwater losses over larger remote regions and comparison with observed total water storage depletion is required. Ongoing work also aims to improve the methodology in terms of accuracy (section 5.4.4) as well as operational aspects (e.g. automation and data visualization) so that it can be used as a water management tool.

## REFERENCES

- Abo-Ghobar, H. M. (1992), Losses from low-pressure center-pivot irrigation systems in a desert climate as affected by nozzle height, *Agricultural Water Management*, 21(1), 23 – 32, doi:[https://doi.org/10.1016/0378-3774\(92\)90079-C](https://doi.org/10.1016/0378-3774(92)90079-C).
- Adler, R. F., G. J. Huffman, A. Chang, R. Ferraro, P.-P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, and E. Nelkin (2003), The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979-present), *Journal of Hydrometeorology*, 4(6), 1147–1167, doi:[10.1175/1525-7541\(2003\)004<1147:TVGPCP>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2).
- Ahmad, M. A., S.-i. Azuma, and T. Sugie (2014), A model-free approach for maximizing power production of wind farm using multi-resolution simultaneous perturbation stochastic approximation, *Energies*, 7(9), 5624–5646, doi:[10.3390/en7095624](https://doi.org/10.3390/en7095624).
- Ahmed, M., M. Sultan, J. Wahr, E. Yan, A. Milewski, W. Sauck, R. Becker, and B. Welton (2011), Integration of GRACE (Gravity Recovery and Climate Experiment) data with traditional data sets for a better understanding of the time-dependent water partitioning in African watersheds, *Geology*, 39(5), 479, doi:[10.1130/G31812.1](https://doi.org/10.1130/G31812.1).
- Ahmed, M., M. Sultan, J. Wahr, and E. Yan (2014), The use of GRACE data to monitor natural and anthropogenic induced variations in water availability across Africa, *Earth-Science Reviews*, 136(Supplement C), 289 – 300, doi:<https://doi.org/10.1016/j.earscirev.2014.05.009>.
- Ajami, H., M. F. McCabe, and J. P. Evans (2015), Impacts of model initialization on an integrated surface water–groundwater model, *Hydrological Processes*, 29(17), 3790–3801, doi:[10.1002/hyp.10478](https://doi.org/10.1002/hyp.10478).
- Al-Rumikhani, Y. A. (2002), Effect of crop sequence, soil sample location and depth on soil water holding capacity under center pivot irrigation, *Agricultural Water Management*, 55(2), 93 – 104, doi:[https://doi.org/10.1016/S0378-3774\(01\)00190-1](https://doi.org/10.1016/S0378-3774(01)00190-1).

- Al-Salamah, I. S., Y. M. Ghazaw, and A. R. Ghumman (2011), Groundwater modeling of Saq aquifer Buraydah Al Qassim for better water management strategies, *Environmental Monitoring and Assessment*, *173*(1), 851–860, doi:10.1007/s10661-010-1428-1.
- Al-Yemeni, M. N., and J. Grace (1995), Radiation balance of an alfalfa crop in Saudi Arabia, *Journal of Arid Environments*, *29*(4), 447 – 454, doi:https://doi.org/10.1016/S0140-1963(95)80017-4.
- Alcamo, J., P. Döll, T. Henrichs, F. Kaspar, B. Lehner, T. Rösch, and S. Siebert (2003), Development and testing of the Watergap 2 global model of water use and availability, *Hydrological Sciences Journal*, *48*(3), 317–337, doi:10.1623/hysj.48.3.317.45290.
- AlSarmi, S., and R. Washington (2011), Recent observed climate change over the Arabian peninsula, *Journal of Geophysical Research: Atmospheres*, *116*(D11), D11,109, doi:10.1029/2010JD015459, d11109.
- Altaf, M. U., A. W. Heemink, M. Verlaan, and I. Hoteit (2011), Simultaneous perturbation stochastic approximation for tidal models, *Ocean Dynamics*, *61*(8), 1093–1105, doi:10.1007/s10236-011-0387-6.
- Anderson, M., J. Norman, W. Kustas, R. Houborg, P. Starks, and N. Agam (2008), A thermal-based remote sensing technique for routine mapping of land-surface carbon, water and energy fluxes from field to regional scales, *Remote Sensing of Environment*, *112*(12), 4227 – 4241, doi:https://doi.org/10.1016/j.rse.2008.07.009.
- Anderson, R. G., M.-H. Lo, S. Swenson, J. S. Famiglietti, Q. Tang, T. H. Skaggs, Y.-H. Lin, and R.-J. Wu (2015), Using satellite-based estimates of evapotranspiration and groundwater changes to determine anthropogenic water fluxes in land surface models, *Geoscientific Model Development*, *8*(10), 3021–3031, doi:10.5194/gmd-8-3021-2015.
- Arkani-Hamed, J. (1998), The lunar mascons revisited, *Journal of Geophysical Research: Planets*, *103*(E2), 3709–3739, doi:10.1029/97JE02815.
- A.S., M., and O. A.M. (1954), Osnovnye zakonomernosti turbulentnogo peremeshivaniya v prizemnom sloe atmosfery (basic laws of turbulent mixing in the atmosphere near the ground), *Trudy geofiz. inst. AN SSSR*, *24*(151), 163–187.

- Beck, H. E., A. I. J. M. van Dijk, D. G. Miralles, R. A. M. de Jeu, L. A. (Sampurno) Bruijnzeel, T. R. McVicar, and J. Schellekens (2013), Global patterns in base flow index and recession based on streamflow observations from 3394 catchments, *Water Resources Research*, *49*(12), 7843–7863, doi:10.1002/2013WR013918.
- Bierkens, M. F. P., V. A. Bell, P. Burek, N. Chaney, L. E. Condon, C. H. David, A. de Roo, P. Döll, N. Drost, J. S. Famiglietti, M. Flörke, D. J. Gochis, P. Houser, R. Hut, J. Keune, S. Kollet, R. M. Maxwell, J. T. Reager, L. Samaniego, E. Sudicky, E. H. Sutanudjaja, N. van de Giesen, H. Winsemius, and E. F. Wood (2015), Hyper-resolution global hydrological modelling: what is next?, *Hydrological Processes*, *29*(2), 310–320, doi:10.1002/hyp.10391, hYP-14-0850.
- Boening, C., J. K. Willis, F. W. Landerer, R. S. Nerem, and J. Fasullo (2012), The 2011 la Niña: So strong, the oceans fell, *Geophysical Research Letters*, *39*(19), L19,602, doi:10.1029/2012GL053055, 119602.
- Brocca, L., L. Ciabatta, C. Massari, T. Moramarco, S. Hahn, S. Hasenauer, R. Kidd, W. Dorigo, W. Wagner, and V. Levizzani (2014), Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data, *Journal of Geophysical Research: Atmospheres*, *119*(9), 5128–5141, doi:10.1002/2014JD021489, 2014JD021489.
- Bruinsma, S., J.-M. Lemoine, R. Biancale, and N. Vals (2010), CNES/GRGS 10-day gravity field models (release 2) and their evaluation, *Advances in Space Research*, *45*(4), 587 – 601, doi:https://doi.org/10.1016/j.asr.2009.10.012.
- Brunner, P., and C. T. Simmons (2012), Hydrogeosphere: A fully integrated, physically based hydrological model, *Ground Water*, *50*(2), 170–176, doi:10.1111/j.1745-6584.2011.00882.x.
- Cai, X., Z.-L. Yang, C. H. David, G.-Y. Niu, and M. Rodell (2014), Hydrological evaluation of the NOAH-MP land surface model for the Mississippi river basin, *Journal of Geophysical Research: Atmospheres*, *119*(1), 23–38, doi:10.1002/2013JD020792, 2013JD020792.
- Camporese, M., C. Paniconi, M. Putti, and S. Orlandini (2010), Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data, *Water Resources Research*, *46*(2), W02,512, doi:10.1029/2008WR007536, w02512.

- Chen, F., and J. Dudhia (2001), Coupling an advanced land surface–hydrology model with the Penn state-NCAR MM5 modeling system. Part I: Model implementation and sensitivity, *Monthly Weather Review*, *129*(4), 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.
- Chen, F., W. T. Crow, P. J. Starks, and D. N. Moriasi (2011), Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture, *Advances in Water Resources*, *34*(4), 526 – 536, doi:<https://doi.org/10.1016/j.advwatres.2011.01.011>.
- Chen, J. L., C. R. Wilson, and B. D. Tapley (2010), The 2009 exceptional Amazon flood and interannual terrestrial water storage change observed by GRACE, *Water Resources Research*, *46*(12), doi:10.1029/2010WR009383, w12526.
- Choi, H. I. (2013), Application of a land surface model using remote sensing data for high resolution simulations of terrestrial processes, *Remote Sensing*, *5*(12), 6838–6856, doi:10.3390/rs5126838.
- Chowdhury, S., and M. Al-Zahrani (2015), Characterizing water resources and trends of sector wise water consumptions in Saudi Arabia, *Journal of King Saud University - Engineering Sciences*, *27*(1), 68 – 82, doi:<https://doi.org/10.1016/j.jksues.2013.02.002>.
- Christophe, B., D. Boulanger, B. Foulon, P.-A. Huynh, V. Lebat, F. Liorzou, and E. Perrot (2015), A new generation of ultra-sensitive electrostatic accelerometers for GRACE follow-on and towards the next generation gravity missions, *Acta Astronautica*, *117*(Supplement C), 1 – 7, doi:<https://doi.org/10.1016/j.actaastro.2015.06.021>.
- Cleugh, H. A., R. Leuning, Q. Mu, and S. W. Running (2007), Regional evaporation estimates from flux tower and MODIS satellite data, *Remote Sensing of Environment*, *106*(3), 285 – 304, doi:<https://doi.org/10.1016/j.rse.2006.07.007>.
- Colaizzi, P. D., W. P. Kustas, M. C. Anderson, N. Agam, J. A. Tolck, S. R. Evett, T. A. Howell, P. H. Gowda, and S. A. O’Shaughnessy (2012), Two-source energy balance model estimates of evapotranspiration using component and composite surface temperatures, *Advances in Water Resources*, *50*(Supplement C), 134 – 151, doi:<https://doi.org/10.1016/j.advwatres.2012.06.004>, the Bushland Evapotranspiration and Agricultural Remote Sensing Experiment 2008.

- Coon, E. T., J. D. Moulton, and S. L. Painter (2016), Managing complexity in simulations of land surface and near-surface processes, *Environmental Modelling & Software*, 78(Supplement C), 134 – 149, doi:<https://doi.org/10.1016/j.envsoft.2015.12.017>.
- Crow, W. T. (2007), A novel method for quantifying value in spaceborne soil moisture retrievals, *Journal of Hydrometeorology*, 8(1), 56–67, doi:10.1175/JHM553.1.
- Deardorff, J. W. (1977), A parameterization of ground-surface moisture content for use in atmospheric prediction models, *Journal of Applied Meteorology*, 16(11), 1182–1185, doi:10.1175/1520-0450(1977)016<1182:APOGSM>2.0.CO;2.
- Deardorff, J. W. (1978), Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation, *Journal of Geophysical Research: Oceans*, 83(C4), 1889–1903, doi:10.1029/JC083iC04p01889.
- Decker, M., D. Or, A. Pitman, and A. Ukkola (2017), New turbulent resistance parameterization for soil evaporation based on a pore-scale model: Impact on surface fluxes in CABLE, *Journal of Advances in Modeling Earth Systems*, 9(1), 220–238, doi:10.1002/2016MS000832.
- Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hlm, L. Isaksen, P. Kllberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597, doi:10.1002/qj.828.
- DeNicola, E., O. S. Aburizaiza, A. Siddique, H. Khwaja, and D. O. Carpenter (2015), Climate change and water scarcity: The case of Saudi Arabia, *Annals of Global Health*, 81(3), 342 – 353, doi:<https://doi.org/10.1016/j.aogh.2015.08.005>, climate Change, Global Health and Human Rights.
- Donlon, C., B. Berruti, A. Buongiorno, M.-H. Ferreira, P. Féménias, J. Frerick, P. Goryl, U. Klein, H. Laur, C. Mavrocordatos, J. Nieke, H. Rebhan, B. Seitz, J. Stroede, and R. Sciarra (2012), The global monitoring for environment and security (GMES) Sentinel-3 mission, *Remote Sensing of Environment*, 120(Supplement

- C), 37 – 57, doi:<https://doi.org/10.1016/j.rse.2011.07.024>, the Sentinel Missions - New Opportunities for Science.
- Duan, X. J., J. Y. Guo, C. K. Shum, and W. van der Wal (2009), On the postprocessing removal of correlated errors in GRACE temporal gravity field solutions, *Journal of Geodesy*, 83(11), 1095, doi:10.1007/s00190-009-0327-0.
- Dudhia, J. (1989), Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model, *Journal of the Atmospheric Sciences*, 46(20), 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.
- El Kenawy, A. M., and M. F. McCabe (2016), A multi-decadal assessment of the performance of gauge- and model-based rainfall products over Saudi Arabia: climatology, anomalies and trends, *International Journal of Climatology*, 36(2), 656–674, doi:10.1002/joc.4374.
- El Kenawy, A. M., M. F. McCabe, G. L. Stenchikov, and J. Raj (2014), Multi-decadal classification of synoptic weather types, observed trends and links to rainfall characteristics over Saudi Arabia, *Frontiers in Environmental Science*, 2, 37, doi:10.3389/fenvs.2014.00037.
- Ershadi, A., M. McCabe, J. Evans, N. Chaney, and E. Wood (2014), Multi-site evaluation of terrestrial evaporation models using fluxnet data, *Agricultural and Forest Meteorology*, 187(Supplement C), 46 – 61, doi:<https://doi.org/10.1016/j.agrformet.2013.11.008>.
- Ershadi, A., M. McCabe, J. Evans, and E. Wood (2015), Impact of model structure and parameterization on Penman–Monteith type evaporation models, *Journal of Hydrology*, 525(Supplement C), 521 – 535, doi:<https://doi.org/10.1016/j.jhydrol.2015.04.008>.
- Evans, J. P. (2009), 21st Century climate change in the Middle East, *Climatic Change*, 92(3), 417–432, doi:10.1007/s10584-008-9438-5.
- Famiglietti, J. S., and M. Rodell (2013), Water in the balance, *Science*, 340(6138), 1300–1301, doi:10.1126/science.1236460.
- Famiglietti, J. S., M. Lo, S. L. Ho, J. Bethune, K. J. Anderson, T. H. Syed, S. C. Swenson, C. R. de Linage, and M. Rodell (2011), Satellites measure recent rates of

- groundwater depletion in California's central valley, *Geophysical Research Letters*, *38*(3), L03,403, doi:10.1029/2010GL046442, 103403.
- Ferguson, C. R., and E. F. Wood (2011), Observed land-atmosphere coupling from satellite remote sensing and reanalysis, *Journal of Hydrometeorology*, *12*(6), 1221–1254, doi:10.1175/2011JHM1380.1.
- Fiaz, S., M. A. Noor, and F. O. Aldosri (2016), Achieving food security in the Kingdom of Saudi Arabia through innovation: Potential role of agricultural extension, *Journal of the Saudi Society of Agricultural Sciences*, doi:https://doi.org/10.1016/j.jssas.2016.09.001.
- Fletcher, R. (1994), *An Overview of Unconstrained Optimization*, pp. 109–143, Springer Netherlands, Dordrecht, doi:10.1007/978-94-009-0369-2\_5.
- Flörke, M., E. Kynast, I. Brlund, S. Eisner, F. Wimmer, and J. Alcamo (2013), Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: A global simulation study, *Global Environmental Change*, *23*(1), 144 – 156, doi:https://doi.org/10.1016/j.gloenvcha.2012.10.018.
- Frappart, F., G. Ramillien, P. Maisongrande, and M. P. Bonnet (2010), Denoising satellite gravity signals by independent component analysis, *IEEE Geoscience and Remote Sensing Letters*, *7*(3), 421–425, doi:10.1109/LGRS.2009.2037837.
- Frappart, F., G. Ramillien, M. Leblanc, S. O. Tweed, M.-P. Bonnet, and P. Maisongrande (2011), An independent component analysis filtering approach for estimating continental hydrology in the GRACE gravity data, *Remote Sensing of Environment*, *115*(1), 187 – 204, doi:https://doi.org/10.1016/j.rse.2010.08.017.
- Friedl, M. A., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang (2010), Modis collection 5 global land cover: Algorithm refinements and characterization of new datasets, *Remote Sensing of Environment*, *114*(1), 168 – 182, doi:https://doi.org/10.1016/j.rse.2009.08.016.
- Fu, B. P. (1981), On the calculation of the evaporation from land surface, *Sci. Atmos. Sin. (in Chinese)*, *5*, 23–31.
- Ganguly, S., A. Samanta, M. A. Schull, N. V. Shabanov, C. Milesi, R. R. Nemani, Y. Knyazikhin, and R. B. Myneni (2008), Generating vegetation leaf area index Earth system data record from multiple sensors. Part 2: Implementation, analysis

- and validation, *Remote Sensing of Environment*, 112(12), 4318 – 4332, doi:<https://doi.org/10.1016/j.rse.2008.07.013>.
- Gao, F., W. Kustas, and M. Anderson (2012), A data mining approach for sharpening thermal satellite imagery over land, *Remote Sensing*, 4(12), 32873319, doi:10.3390/rs4113287.
- Gao, H., Q. Tang, C. R. Ferguson, E. F. Wood, and D. P. Lettenmaier (2010), Estimating the water budget of major US river basins via remote sensing, *International Journal of Remote Sensing*, 31(14), 3955–3978, doi:10.1080/01431161.2010.483488.
- Gao, Y., and D. Long (2008), Intercomparison of remote sensing-based models for estimation of evapotranspiration and accuracy assessment based on SWAT, *Hydrological Processes*, 22(25), 4850–4869, doi:10.1002/hyp.7104.
- Gash, J. H. C. (1979), An analytical model of rainfall interception by forests, *Quarterly Journal of the Royal Meteorological Society*, 105(443), 43–55, doi:10.1002/qj.49710544304.
- Gonzalez-Dugo, M., C. Neale, L. Mateos, W. Kustas, J. Prueger, M. Anderson, and F. Li (2009), A comparison of operational remote sensing-based models for estimating crop evapotranspiration, *Agricultural and Forest Meteorology*, 149(11), 1843 – 1853, doi:<https://doi.org/10.1016/j.agrformet.2009.06.012>, special Section on Water and Carbon Dynamics in Selected Ecosystems in China.
- Goyal, S. K. (1982), *Aquifer Parameter Identification Using Stochastic Approximation*, pp. 1055–1062, Springer Berlin Heidelberg, Berlin, Heidelberg, doi:10.1007/978-3-662-02348-8\_89.
- Grindle, A. K., A. Siddiqi, and L. D. Anadon (2015), Food security amidst water scarcity: Insights on sustainable food production from Saudi Arabia, *Sustainable Production and Consumption*, 2(Supplement C), 67 – 78, doi:<https://doi.org/10.1016/j.spc.2015.06.002>, sustainability issues in the foodenergywater nexus.
- Gupta, S. K., P. W. Stackhouse, S. J. Cox, J. C. Mikovitz, and T. P. Zhang (2006), 22-year surface radiation budget data set, GEWEX news.
- Hassan, A., and S. Jin (2016), Water storage changes and balances in Africa observed by GRACE and hydrologic models, *Geodesy and Geodynamics*, 7(1), 39 – 49, doi:<https://doi.org/10.1016/j.geog.2016.03.002>, special Issue: GNSS Applications in Geodesy and Geodynamics.

- Haverd, V., M. R. Raupach, P. R. Briggs, J. G. Canadell., S. J. Davis, R. M. Law, C. P. Meyer, G. P. Peters, C. Pickett-Heaps, and B. Sherman (2013), The Australian terrestrial carbon budget, *Biogeosciences*, *10*(2), 851–869, doi:10.5194/bg-10-851-2013.
- Haverd, V., M. Cuntz, L. P. Nieradzik, and I. N. Harman (2016), Improved representations of coupled soil–canopy processes in the CABLE land surface model (subversion revision 3432), *Geoscientific Model Development*, *9*(9), 3111–3122, doi:10.5194/gmd-9-3111-2016.
- Hoffmann, H., H. Nieto, R. Jensen, R. Guzinski, P. Zarco-Tejada, and T. Friborg (2016), Estimating evaporation with thermal UAV data and two-source energy balance models, *Hydrology and Earth System Sciences*, *20*(2), 697–713, doi:10.5194/hess-20-697-2016.
- Hong, S.-Y., Y. Noh, and J. Dudhia (2006), A new vertical diffusion package with an explicit treatment of entrainment processes, *Monthly Weather Review*, *134*(9), 2318–2341, doi:10.1175/MWR3199.1.
- Hou, A. Y., R. K. Kakar, S. Neeck, A. A. Azarbarzin, C. D. Kummerow, M. Kojima, R. Oki, K. Nakamura, and T. Iguchi (2014), The global precipitation measurement mission, *Bulletin of the American Meteorological Society*, *95*(5), 701–722, doi:10.1175/BAMS-D-13-00164.1.
- Houborg, R., and M. F. McCabe (2016), High-resolution NDVI from Planet’s constellation of earth observing nano-satellites: A new data source for precision agriculture, *Remote Sensing*, *8*(9), doi:10.3390/rs8090768.
- Houborg, R., and M. F. McCabe (2017), Impacts of dust aerosol and adjacency effects on the accuracy of Landsat 8 and rapideye surface reflectances, *Remote Sensing of Environment*, *194*(Supplement C), 127 – 145, doi:https://doi.org/10.1016/j.rse.2017.03.013.
- Houborg, R., and M. F. McCabe (2018), A hybrid training approach for leaf area index estimation via Cubist and random forests machine-learning, *ISPRS Journal of Photogrammetry and Remote Sensing*, *135*, 173 – 188, doi:https://doi.org/10.1016/j.isprsjprs.2017.10.004.
- Houborg, R., M. Rodell, B. Li, R. Reichle, and B. F. Zaitchik (2012), Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment

- (GRACE) terrestrial water storage observations, *Water Resources Research*, *48*(7), W07,525, doi:10.1029/2011WR011291, w07525.
- Houborg, R., M. McCabe, A. Cescatti, F. Gao, M. Schull, and A. Gitelson (2015), Joint leaf chlorophyll content and leaf area index retrieval from Landsat data using a regularized model inversion system (regflec), *Remote Sensing of Environment*, *159*(Supplement C), 203 – 221, doi:https://doi.org/10.1016/j.rse.2014.12.008.
- Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider (1997), The global precipitation climatology project (GPCP) combined precipitation dataset, *Bulletin of the American Meteorological Society*, *78*(1), 5–20, doi:10.1175/1520-0477(1997)078<0005: TGPCPG>2.0.CO;2.
- Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind (2001), Global precipitation at one-degree daily resolution from multisatellite observations, *Journal of Hydrometeorology*, *2*(1), 36–50, doi:10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2.
- Hussain, G., and A. A. Al-Jaloud (1995), Effect of irrigation and nitrogen on water use efficiency of wheat in Saudi Arabia, *Agricultural Water Management*, *27*(2), 143 – 153, doi:https://doi.org/10.1016/0378-3774(95)91233-W.
- Izquier, A., P. Carrin, J. Tarjuelo, and M. Moreno (2015), Optimal reservoir capacity for centre pivot irrigation water supply: Maize cultivation in Spain, *Biosystems Engineering*, *135*(Supplement C), 61 – 72, doi:https://doi.org/10.1016/j.biosystemseng.2015.04.015.
- Jana, R. B., A. Ershadi, and M. F. McCabe (2016), Examining the relationship between intermediate-scale soil moisture and terrestrial evaporation within a semi-arid grassland, *Hydrology and Earth System Sciences*, *20*(10), 3987–4004, doi:10.5194/hess-20-3987-2016.
- Jiang, H., J. T. Farrar, R. C. Beardsley, R. Chen, and C. Chen (2009), Zonal surface wind jets across the red sea due to mountain gap forcing along both sides of the red sea, *Geophysical Research Letters*, *36*(19), L19,605, doi:10.1029/2009GL040008, l19605.
- Jiménez, C., C. Prigent, B. Mueller, S. I. Seneviratne, M. F. McCabe, E. F. Wood, W. B. Rossow, G. Balsamo, A. K. Betts, P. A. Dirmeyer, J. B. Fisher, M. Jung,

- M. Kanamitsu, R. H. Reichle, M. Reichstein, M. Rodell, J. Sheffield, K. Tu, and K. Wang (2011), Global intercomparison of 12 land surface heat flux estimates, *Journal of Geophysical Research: Atmospheres*, 116(D2), D02,102, doi:10.1029/2010JD014545, d02102.
- Jin, S., and G. Feng (2013), Large-scale variations of global groundwater from satellite gravimetry and hydrological models, 20022012, *Global and Planetary Change*, 106(Supplement C), 20 – 30, doi:https://doi.org/10.1016/j.gloplacha.2013.02.008.
- Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie (2004), CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, *Journal of Hydrometeorology*, 5(3), 487–503, doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.
- Kalma, J. D., T. R. McVicar, and M. F. McCabe (2008), Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data, *Surveys in Geophysics*, 29(4), 421–469, doi:10.1007/s10712-008-9037-z.
- Kirtman, B., S. Power, J. Adedoyin, G. Boer, R. Bojariu, I. Camilloni, F. Doblas-Reyes, A. Fiore, M. Kimoto, G. Meehl, M. Prather, A. Sarr, C. Schr, R. Sutton, G. van Oldenborgh, G. Vecchi, and H. Wang (2013), *Near-term Climate Change: Projections and Predictability*, book section 11, p. 9531028, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi: 10.1017/CBO9781107415324.023.
- Klees, R., E. A. Zapreeva, H. C. Winsemius, and H. H. G. Savenije (2007), The bias in GRACE estimates of continental water storage variations, *Hydrology and Earth System Sciences*, 11(4), 1227–1241, doi:10.5194/hess-11-1227-2007.
- Klees, R., E. A. Revtova, B. C. Gunter, P. Ditmar, E. Oudman, H. C. Winsemius, and H. H. G. Savenije (2008), The design of an optimal filter for monthly GRACE gravity models, *Geophysical Journal International*, 175(2), 417–432, doi:10.1111/j.1365-246X.2008.03922.x.
- Kollet, S. J., and R. M. Maxwell (2006), Integrated surfacegroundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model, *Advances in Water Resources*, 29(7), 945 – 958, doi:https://doi.org/10.1016/j.advwatres.2005.08.006.

- Kollet, S. J., R. M. Maxwell, C. S. Woodward, S. Smith, J. Vanderborght, H. Vereecken, and C. Simmer (2010), Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources, *Water Resources Research*, *46*(4), W04,201, doi:10.1029/2009WR008730, w04201.
- Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure, *Journal of Geophysical Research: Atmospheres*, *105*(D20), 24,809–24,822, doi:10.1029/2000JD900327.
- Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel (2006), World map of the köppen-geiger climate classification updated, *Meteorologische Zeitschrift*, *15*(3), 259–263, doi:10.1127/0941-2948/2006/0130.
- Kowalczyk, E. (2013), The land surface model component of ACCESS: description and impact on the simulated surface climatology, *Australian Meteorological and Oceanographic Journal*, *63*, 65–82.
- Kowalczyk, E. A., Y. P. Wang, R. M. Law, H. L. Davies, J. L. McGregor, and G. Abramowitz (2006), The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model, *Tech. Rep. 13*, CSIRO Marine and Atmospheric Research paper, Aspendale, Vic.
- Krause, P., D. P. Boyle, and F. Bäse (2005), Comparison of different efficiency criteria for hydrological model assessment, *Advances in Geosciences*, *5*, 89–97, doi:10.5194/adgeo-5-89-2005.
- Kustas, W. P., and J. M. Norman (1999), Evaluation of soil and vegetation heat flux predictions using a simple two-source model with radiometric temperatures for partial canopy cover, *Agricultural and Forest Meteorology*, *94*(1), 13 – 29, doi:https://doi.org/10.1016/S0168-1923(99)00005-2.
- Landerer, F. W., and S. C. Swenson (2012), Accuracy of scaled GRACE terrestrial water storage estimates, *Water Resources Research*, *48*(4), W04,531, doi:10.1029/2011WR011453, w04531.
- Langodan, S., L. Cavaleri, Y. Viswanadhapalli, and I. Hoteit (2014), The red sea: A natural laboratory for wind and wave modeling, *Journal of Physical Oceanography*, *44*(12), 3139–3159, doi:10.1175/JPO-D-13-0242.1.

- Langodan, S., Y. Viswanadhapalli, H. P. Dasari, O. Knio, and I. Hoteit (2016), A high-resolution assessment of wind and wave energy potentials in the red sea, *Applied Energy*, 181(Supplement C), 244 – 255, doi:<https://doi.org/10.1016/j.apenergy.2016.08.076>.
- Langodan, S., L. Cavaleri, Y. Vishwanadhapalli, A. Pomaro, L. Bertotti, and I. Hoteit (2017), The climatology of the red sea - part 1: the wind, *International Journal of Climatology*, 37(13), 4509–4517, doi:10.1002/joc.5103.
- Lemoine, J.-M., S. Bruinsma, S. Loyer, R. Biancale, J.-C. Marty, F. Perosanz, and G. Balmino (2007), Temporal gravity field models inferred from GRACE data, *Advances in Space Research*, 39(10), 1620 – 1629, doi:<https://doi.org/10.1016/j.asr.2007.03.062>.
- Leuning, R., Y. Q. Zhang, A. Rajaud, H. Cleugh, and K. Tu (2008), A simple surface conductance model to estimate regional evaporation using MODIS leaf area index and the Penman–Monteith equation, *Water Resources Research*, 44(10), W10,419, doi:10.1029/2007WR006562, w10419.
- Li, L., Y.-P. Wang, Q. Yu, B. Pak, D. Eamus, J. Yan, E. van Gorsel, and I. T. Baker (2012), Improving the responses of the Australian Community Land Surface Model (CABLE) to seasonal drought, *Journal of Geophysical Research: Biogeosciences*, 117(G4), G04,002, doi:10.1029/2012JG002038, g04002.
- lin Hsu, K., X. Gao, S. Sorooshian, and H. V. Gupta (1997), Precipitation estimation from remotely sensed information using artificial neural networks, *Journal of Applied Meteorology*, 36(9), 1176–1190, doi:10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.
- Liu, X., P. Ditmar, C. Siemes, D. C. Slobbe, E. Revtova, R. Klees, R. Riva, and Q. Zhao (2010), DEOS mass transport model (DMT-1) based on GRACE satellite data: methodology and validation, *Geophysical Journal International*, 181(2), 769–788, doi:10.1111/j.1365-246X.2010.04533.x.
- Liu, Y., W. Dorigo, R. Parinussa, R. de Jeu, W. Wagner, M. McCabe, J. Evans, and A. van Dijk (2012), Trend-preserving blending of passive and active microwave soil moisture retrievals, *Remote Sensing of Environment*, 123(Supplement C), 280 – 297, doi:<https://doi.org/10.1016/j.rse.2012.03.014>.

- Liu, Y. Y., A. I. J. M. van Dijk, M. F. McCabe, J. P. Evans, and R. A. M. de Jeu (2013), Global vegetation biomass change (1988–2008) and attribution to environmental and human drivers, *Global Ecology and Biogeography*, *22*(6), 692–705, doi:10.1111/geb.12024.
- Lo, M.-H., J. S. Famiglietti, P. J.-F. Yeh, and T. H. Syed (2010), Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data, *Water Resources Research*, *46*(5), W05,517, doi:10.1029/2009WR007855, w05517.
- Long, D., L. Longuevergne, and B. R. Scanlon (2014), Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites, *Water Resources Research*, *50*(2), 1131–1151, doi:10.1002/2013WR014581.
- Long, D., L. Longuevergne, and B. R. Scanlon (2015), Global analysis of approaches for deriving total water storage changes from GRACE satellites, *Water Resources Research*, *51*(4), 2574–2594, doi:10.1002/2014WR016853.
- Lopez, O., G. Stenchikov, and T. M. Missimer (2014), Water management during climate change using aquifer storage and recovery of stormwater in a dunefield in western Saudi Arabia, *Environmental Research Letters*, *9*(7), 075,008.
- Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Yang, and J. W. Merchant (2000), Development of a global land cover characteristics database and igbp discover from 1 km AVHRR data, *International Journal of Remote Sensing*, *21*(6-7), 1303–1330, doi:10.1080/014311600210191.
- Madugundu, R., K. A. Al-Gaadi, E. Tola, A. G. Kayad, and C. S. Jha (2017), Estimation of gross primary production of irrigated maize using Landsat-8 imagery and eddy covariance data, *Saudi Journal of Biological Sciences*, *24*(2), 410 – 420, doi:https://doi.org/10.1016/j.sjbs.2016.10.003.
- Martinec, Z. (1989), Program to calculate the spectral harmonic expansion coefficients of the two scalar fields product, *Computer Physics Communications*, *54*(1), 177 – 182, doi:https://doi.org/10.1016/0010-4655(89)90043-X.
- Maxwell, R. M. (2013), A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling, *Advances in Water Resources*, *53*(Supplement C), 109 – 117, doi:https://doi.org/10.1016/j.advwatres.2012.10.001.

- McCabe, M., S. Franks, and J. Kalma (2005a), Calibration of a land surface model using multiple data sets, *Journal of Hydrology*, *302*(1), 209 – 222, doi:<https://doi.org/10.1016/j.jhydrol.2004.07.002>.
- McCabe, M., E. Wood, R. Wjck, M. Pan, J. Sheffield, H. Gao, and H. Su (2008), Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies, *Remote Sensing of Environment*, *112*(2), 430 – 444, doi:<https://doi.org/10.1016/j.rse.2007.03.027>, soil Moisture Experiments 2004 (SMEX04) Special Issue.
- McCabe, M. F., B. Aragon, R. Houborg, and J. Mascaro (2018), Cubesats in hydrology: Ultrahigh-resolution insights into vegetation dynamics and terrestrial evaporation, *Water Resources Research*, in press, doi:10.1002/2017WR022240.
- McCabe, M. F., J. D. Kalma, and S. W. Franks (2005b), Spatial and temporal patterns of land surface fluxes from remotely sensed surface temperatures within an uncertainty modelling framework, *Hydrology and Earth System Sciences*, *9*(5), 467–480, doi:10.5194/hess-9-467-2005.
- McCabe, M. F., A. Ershadi, C. Jimenez, D. G. Miralles, D. Michel, and E. F. Wood (2016), The GEWEX landflux project: evaluation of model evaporation using tower-based and globally gridded forcing data, *Geoscientific Model Development*, *9*(1), 283–305, doi:10.5194/gmd-9-283-2016.
- McCabe, M. F., M. Rodell, D. E. Alsdorf, D. G. Miralles, R. Uijlenhoet, W. Wagner, A. Lucier, R. Houborg, N. E. C. Verhoest, T. E. Franz, J. Shi, H. Gao, and E. F. Wood (2017), The future of earth observation in hydrology, *Hydrology and Earth System Sciences*, *21*(7), 3879–3914, doi:10.5194/hess-21-3879-2017.
- Michel, D., C. Jiménez, D. G. Miralles, M. Jung, M. Hirschi, A. Ershadi, B. Martens, M. F. McCabe, J. B. Fisher, Q. Mu, S. I. Seneviratne, E. F. Wood, and D. Fernández-Prieto (2016), The wacmos-et project – part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms, *Hydrology and Earth System Sciences*, *20*(2), 803–822, doi:10.5194/hess-20-803-2016.
- Milewski, A., M. Sultan, S. M. Jayaprakash, R. Balekai, and R. Becker (2009), RESDEM, a tool for integrating temporal remote sensing data for use in hydrogeologic investigations, *Computers & Geosciences*, *35*(10), 2001 – 2010, doi:<https://doi.org/10.1016/j.cageo.2009.02.010>.

- Miralles, D. G., J. H. Gash, T. R. H. Holmes, R. A. M. de Jeu, and A. J. Dolman (2010), Global canopy interception from satellite observations, *Journal of Geophysical Research: Atmospheres*, *115*(D16), D16,122, doi:10.1029/2009JD013530, d16122.
- Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman (2011a), Global land-surface evaporation estimated from satellite-based observations, *Hydrology and Earth System Sciences*, *15*(2), 453–469, doi:10.5194/hess-15-453-2011.
- Miralles, D. G., R. A. M. De Jeu, J. H. Gash, T. R. H. Holmes, and A. J. Dolman (2011b), Magnitude and variability of land evaporation and its components at the global scale, *Hydrology and Earth System Sciences*, *15*(3), 967–981, doi:10.5194/hess-15-967-2011.
- Miralles, D. G., C. Jiménez, M. Jung, D. Michel, A. Ershadi, M. F. McCabe, M. Hirschi, B. Martens, A. J. Dolman, J. B. Fisher, Q. Mu, S. I. Seneviratne, E. F. Wood, and D. Fernández-Prieto (2016), The WACMOS-ET project – part 2: Evaluation of global terrestrial evaporation data sets, *Hydrology and Earth System Sciences*, *20*(2), 823–842, doi:10.5194/hess-20-823-2016.
- Missimer, T. M., J. E. Drewes, G. Amy, R. G. Maliva, and S. Keller (2012), Restoration of wadi aquifers by artificial recharge with treated waste water, *Groundwater*, *50*(4), 514–527, doi:10.1111/j.1745-6584.2012.00941.x.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough (1997), Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16,663–16,682, doi:10.1029/97JD00237.
- Molod, A., L. Takacs, M. Suarez, and J. Bacmeister (2015), Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, *Geoscientific Model Development*, *8*(5), 1339–1356, doi:10.5194/gmd-8-1339-2015.
- Mu, Q., F. A. Heinsch, M. Zhao, and S. W. Running (2007), Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, *Remote Sensing of Environment*, *111*(4), 519 – 536, doi:https://doi.org/10.1016/j.rse.2007.04.015.

- Mu, Q., M. Zhao, and S. W. Running (2011), Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sensing of Environment*, 115(8), 1781 – 1800, doi:<https://doi.org/10.1016/j.rse.2011.02.019>.
- Mu, Q., M. Zhao, J. S. Kimball, N. G. McDowell, and S. W. Running (2013), A remotely sensed global terrestrial drought severity index, *Bulletin of the American Meteorological Society*, 94(1), 83–98, doi:10.1175/BAMS-D-11-00213.1.
- Mueller, B., S. I. Seneviratne, C. Jimenez, T. Corti, M. Hirschi, G. Balsamo, P. Ciais, P. Dirmeyer, J. B. Fisher, Z. Guo, M. Jung, F. Maignan, M. F. McCabe, R. Reichle, M. Reichstein, M. Rodell, J. Sheffield, A. J. Teuling, K. Wang, E. F. Wood, and Y. Zhang (2011), Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations, *Geophysical Research Letters*, 38(6), L06,402, doi:10.1029/2010GL046230, 106402.
- Nash, J., and J. Sutcliffe (1970), River flow forecasting through conceptual models part i – a discussion of principles, *Journal of Hydrology*, 10(3), 282 – 290, doi:[https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- New, M., M. Hulme, and P. Jones (2000), Representing twentieth-Century spacetime climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate, *Journal of Climate*, 13(13), 2217–2238, doi:10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2.
- Nikolopoulos, E. I., E. N. Anagnostou, M. Borga, E. R. Vivoni, and A. Papadopoulos (2011), Sensitivity of a mountain basin flash flood to initial wetness condition and rainfall variability, *Journal of Hydrology*, 402(3), 165 – 178, doi:<https://doi.org/10.1016/j.jhydrol.2010.12.020>.
- Niu, G.-Y., Z.-L. Yang, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, A. Kumar, K. Manning, D. Niyogi, E. Rosero, M. Tewari, and Y. Xia (2011), The community NOAA land surface model with multiparameterization options (NOAH-MP): 1. model description and evaluation with local-scale measurements, *Journal of Geophysical Research: Atmospheres*, 116(D12), D12,109, doi:10.1029/2010JD015139, d12109.
- Norman, J., W. Kustas, and K. Humes (1995), Source approach for estimating soil and vegetation energy fluxes in observations of directional radiometric surface temperature, *Agricultural and Forest Meteorology*, 77(3), 263 – 293, doi:

- [https://doi.org/10.1016/0168-1923\(95\)02265-Y](https://doi.org/10.1016/0168-1923(95)02265-Y), thermal Remote Sensing of the Energy and Water Balance over Vegetation.
- Odhiambo, G. O. (2017), Water scarcity in the Arabian peninsula and socio-economic implications, *Applied Water Science*, 7(5), 2479–2492, doi:10.1007/s13201-016-0440-1.
- Pan, M., and E. F. Wood (2006), Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter, *Journal of Hydrometeorology*, 7(3), 534–547, doi:10.1175/JHM495.1.
- Pan, M., E. F. Wood, R. Wjck, and M. F. McCabe (2008), Estimation of regional terrestrial water cycle using multi-sensor remote sensing observations and data assimilation, *Remote Sensing of Environment*, 112(4), 1282 – 1294, doi:<https://doi.org/10.1016/j.rse.2007.02.039>, remote Sensing Data Assimilation Special Issue.
- Pan, M., A. K. Sahoo, T. J. Troy, R. K. Vinukollu, J. Sheffield, and E. F. Wood (2012), Multisource estimation of long-term terrestrial water budget for major global river basins, *Journal of Climate*, 25(9), 3191–3206, doi:10.1175/JCLI-D-11-00300.1.
- Petropoulos, G. P., G. Ireland, and B. Barrett (2015), Surface soil moisture retrievals from remote sensing: Current status, products & future trends, *Physics and Chemistry of the Earth, Parts A/B/C*, 83-84(Supplement C), 36 – 56, doi:<https://doi.org/10.1016/j.pce.2015.02.009>, emerging science and applications with microwave remote sensing data.
- Pitman, A. J. (2003), The evolution of, and revolution in, land surface schemes designed for climate models, *International Journal of Climatology*, 23(5), 479–510, doi:10.1002/joc.893.
- Qadir, M., T. Boers, S. Schubert, A. Ghafoor, and G. Murtaza (2003), Agricultural water management in water-starved countries: challenges and opportunities, *Agricultural Water Management*, 62(3), 165 – 185, doi:[https://doi.org/10.1016/S0378-3774\(03\)00146-X](https://doi.org/10.1016/S0378-3774(03)00146-X).
- Quinlan, J. R. (1996), Improved use of continuous attributes in c4.5, *J. Artif. Int. Res.*, 4(1), 77–90.
- Quinlan, J. R., and R. L. Rivest (1989), Inferring decision trees using the minimum description length principle, *Information and Computation*, 80(3), 227 – 248, doi:[https://doi.org/10.1016/0890-5401\(89\)90010-2](https://doi.org/10.1016/0890-5401(89)90010-2).

- Raupach, M. R., K. Finkele, and L. Zhang (1997), SCAM (Soil-Canopy-Atmosphere Model): Description and comparison with field data, *Tech. Rep. 132*, CSIRO Center for Environmental Mechanics Technical Report, Canberra, ACT, Australia.
- Renzullo, L., A. van Dijk, J.-M. Perraud, D. Collins, B. Henderson, H. Jin, A. Smith, and D. McJannet (2014), Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment, *Journal of Hydrology*, 519(Part D), 2747 – 2762, doi:<https://doi.org/10.1016/j.jhydrol.2014.08.008>.
- Rieser, D., M. Kuhn, R. Pail, I. M. Anjasmara, and J. Awange (2010), Relation between GRACE-derived surface mass variations and precipitation over Australia, *Australian Journal of Earth Sciences*, 57(7), 887–900, doi:[10.1080/08120099.2010.512645](https://doi.org/10.1080/08120099.2010.512645).
- Rodell, M., J. S. Famiglietti, J. Chen, S. I. Seneviratne, P. Viterbo, S. Holl, and C. R. Wilson (2004a), Basin scale estimates of evapotranspiration using GRACE and other observations, *Geophysical Research Letters*, 31(20), L20,504, doi:[10.1029/2004GL020873](https://doi.org/10.1029/2004GL020873), L20504.
- Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin\*, J. P. Walker, D. Lohmann, and D. Toll (2004b), The global land data assimilation system, *Bulletin of the American Meteorological Society*, 85(3), 381–394, doi:[10.1175/BAMS-85-3-381](https://doi.org/10.1175/BAMS-85-3-381).
- Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, 460(7258), 999–1002, doi:[10.1038/nature08238](https://doi.org/10.1038/nature08238).
- Rojas, R., L. Feyen, A. Bianchi, and A. Dosio (2012), Assessment of future flood hazard in Europe using a large ensemble of bias-corrected regional climate simulations, *Journal of Geophysical Research: Atmospheres*, 117(D17), D17,109, doi:[10.1029/2012JD017461](https://doi.org/10.1029/2012JD017461), d17109.
- Rosas, J., R. Houborg, and M. F. McCabe (2017), Sensitivity of Landsat 8 surface temperature estimates to atmospheric profile data: A study using MODTRAN in dry-land irrigated systems, *Remote Sensing*, 9(10), doi:[10.3390/rs9100988](https://doi.org/10.3390/rs9100988).

- Rossow, W. B., and E. N. Dueñas (2004), The international satellite cloud climatology project (ISCCP) web site: An online resource for research, *Bulletin of the American Meteorological Society*, *85*(2), 167–172, doi:10.1175/BAMS-85-2-167.
- Rowlands, D. D., S. B. Luthcke, S. M. Klosko, F. G. R. Lemoine, D. S. Chinn, J. J. McCarthy, C. M. Cox, and O. B. Anderson (2005), Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements, *Geophysical Research Letters*, *32*(4), L04,310, doi:10.1029/2004GL021908, 104310.
- Rowlands, D. D., S. B. Luthcke, J. J. McCarthy, S. M. Klosko, D. S. Chinn, F. G. Lemoine, J.-P. Boy, and T. J. Sabaka (2010), Global mass flux solutions from GRACE: A comparison of parameter estimation strategies-mass concentrations versus stokes coefficients, *Journal of Geophysical Research: Solid Earth*, *115*(B1), B01,403, doi:10.1029/2009JB006546, b01403.
- Sadeghi, S.-H., T. Peters, B. Shafii, M. Amini, and C. Stöckle (2017), Continuous variation of wind drift and evaporation losses under a linear move irrigation system, *Agricultural Water Management*, *182*(Supplement C), 39 – 54, doi:https://doi.org/10.1016/j.agwat.2016.12.009.
- Sahoo, A. K., M. Pan, T. J. Troy, R. K. Vinukollu, J. Sheffield, and E. F. Wood (2011), Reconciling the global terrestrial water budget using satellite remote sensing, *Remote Sensing of Environment*, *115*(8), 1850 – 1865, doi:https://doi.org/10.1016/j.rse.2011.03.009.
- Santanello Jr., J. A., and M. A. Friedl (2003), Diurnal covariation in soil heat flux and net radiation, *Journal of Applied Meteorology*, *42*(6), 851–862, doi:10.1175/1520-0450(2003)042<0851:DCISHF>2.0.CO;2.
- Scanlon, B. R., Z. Zhang, R. C. Reedy, D. R. Pool, H. Save, D. Long, J. Chen, D. M. Wolock, B. D. Conway, and D. Winester (2015), Hydrologic implications of GRACE satellite data in the Colorado river basin, *Water Resources Research*, *51*(12), 9891–9903, doi:10.1002/2015WR018090.
- Scanlon, B. R., Z. Zhang, H. Save, A. Y. Sun, H. Müller Schmied, L. P. H. van Beek, D. N. Wiese, Y. Wada, D. Long, R. C. Reedy, L. Longuevergne, P. Döll, and M. F. P. Bierkens (2018), Global models underestimate large decadal declining and rising water storage trends relative to grace satellite data, *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1704665115.

- Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf, and M. Ziese (2011), GPCP monitoring product: Near real-time monthly land-surface precipitation from rain-gauges based on synop and climat data, doi:10.5676/DWD\_GPCP/MP\_M\_V4\_100.
- Sheffield, J., G. Goteti, and E. F. Wood (2006), Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *Journal of Climate*, *19*(13), 3088–3111, doi:10.1175/JCLI3790.1.
- Sheffield, J., C. R. Ferguson, T. J. Troy, E. F. Wood, and M. F. McCabe (2009), Closing the terrestrial water budget from satellite remote sensing, *Geophysical Research Letters*, *36*(7), L07,403, doi:10.1029/2009GL037338, 107403.
- Shi, M., Z.-L. Yang, D. M. Lawrence, R. E. Dickinson, and Z. M. Subin (2013), Spin-up processes in the Community Land Model version 4 with explicit carbon and nitrogen components, *Ecological Modelling*, *263*(Supplement C), 308 – 325, doi:https://doi.org/10.1016/j.ecolmodel.2013.04.008.
- Shi, W., M. Wang, and W. Guo (2014), Long-term hydrological changes of the aral sea observed by satellites, *Journal of Geophysical Research: Oceans*, *119*(6), 3313–3326, doi:10.1002/2014JC009988.
- Singh, R. S., J. T. Reager, N. L. Miller, and J. S. Famiglietti (2015), Toward hyper-resolution land-surface modeling: The effects of fine-scale topography and soil texture on CLM4.0 simulations over the southwestern u.s., *Water Resources Research*, *51*(4), 2648–2667, doi:10.1002/2014WR015686.
- Skamarock, W., W. Skamarock, J. Klemp, J. Dudhia, D. Gill, D. Barker, W. Wang, and J. Powers (2008), A description of the advanced research WRF version 3, *Tech. rep.*, NCAR Technical Note - 475 + STR.
- Song, L., W. P. Kustas, S. Liu, P. D. Colaizzi, H. Nieto, Z. Xu, Y. Ma, M. Li, T. Xu, N. Agam, J. A. Tolk, and S. R. Evett (2016), Applications of a thermal-based two-source energy balance model using Priestley-Taylor approach for surface temperature partitioning under advective conditions, *Journal of Hydrology*, *540*(Supplement C), 574 – 587, doi:https://doi.org/10.1016/j.jhydrol.2016.06.034.
- Spall, J. C. (1998), Implementation of the simultaneous perturbation algorithm for stochastic optimization, *IEEE Transactions on Aerospace and Electronic Systems*, *34*(3), 817–823, doi:10.1109/7.705889.

- Sridhar, V. (2013), Tracking the influence of irrigation on land surface fluxes and boundary layer climatology, *Journal of Contemporary Water Research & Education*, 152(1), 79–93, doi:10.1111/j.1936-704X.2013.03170.x.
- Steiner, J., E. Kanemasu, and R. Clark (1983), Spray losses and partitioning of water under a center pivot sprinkler system, *Transactions of the American Society of Agricultural Engineering*, 26(4), 1128 – 1134, doi:10.13031/2013.34090.
- Stephens, G. L., and C. D. Kummerow (2007), The remote sensing of clouds and precipitation from space: A review, *Journal of the Atmospheric Sciences*, 64(11), 3742–3765, doi:10.1175/2006JAS2375.1.
- Stisen, S., M. F. McCabe, J. C. Refsgaard, S. Lerer, and M. B. Butts (2011), Model parameter analysis using remotely sensed pattern information in a multi-constraint framework, *Journal of Hydrology*, 409(1), 337 – 349, doi:https://doi.org/10.1016/j.jhydrol.2011.08.030.
- Sultan, M., N. Sturchio, S. A. Sefry, A. Milewski, R. Becker, I. Nasr, and Z. Sagintayev (2008), Geochemical, isotopic, and remote sensing constraints on the origin and evolution of the Rub Al Khali aquifer system, Arabian peninsula, *Journal of Hydrology*, 356(1), 70 – 83, doi:https://doi.org/10.1016/j.jhydrol.2008.04.001.
- Sultan, M., M. Ahmed, J. Wahr, E. Yan, and M. K. Emil (2014), *Monitoring Aquifer Depletion from Space*, pp. 347–366, John Wiley & Sons, Inc, doi: 10.1002/9781118872086.ch21.
- Sun, A. Y. (2013), Predicting groundwater level changes using GRACE data, *Water Resources Research*, 49(9), 5900–5912, doi:10.1002/wrcr.20421.
- Sun, A. Y., R. Green, S. Swenson, and M. Rodell (2012), Toward calibration of regional groundwater models using GRACE data, *Journal of Hydrology*, 422-423(Supplement C), 1 – 9, doi:https://doi.org/10.1016/j.jhydrol.2011.10.025.
- Sun, L., S. Liang, and Z. Chen (2014), Comparison of ET models over different land cover, *IOP Conference Series: Earth and Environmental Science*, 17(1), 012,128.
- Swenson, S. (2010), Assessing high-latitude winter precipitation from global precipitation analyses using GRACE, *Journal of Hydrometeorology*, 11(2), 405–420, doi: 10.1175/2009JHM1194.1.

- Swenson, S., and J. Wahr (2002), Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity, *Journal of Geophysical Research: Solid Earth*, *107*(B9), ETG 3–1–ETG 3–13, doi:10.1029/2001JB000576, 2193.
- Swenson, S., and J. Wahr (2006), Post-processing removal of correlated errors in GRACE data, *Geophysical Research Letters*, *33*(8), L08,402, doi:10.1029/2005GL025285, 108402.
- Swenson, S., and J. Wahr (2009), Monitoring the water balance of lake Victoria, east Africa, from space, *Journal of Hydrology*, *370*(1), 163 – 176, doi:https://doi.org/10.1016/j.jhydrol.2009.03.008.
- Swenson, S., J. Famiglietti, J. Basara, and J. Wahr (2008), Estimating profile soil moisture and groundwater variations using GRACE and Oklahoma mesonet soil moisture data, *Water Resources Research*, *44*(1), w01413, doi:10.1029/2007WR006057, w01413.
- Takala, M., K. Luojus, J. Pulliainen, C. Derksen, J. Lemmetyinen, J.-P. Krn, J. Koskinen, and B. Bojkov (2011), Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements, *Remote Sensing of Environment*, *115*(12), 3517 – 3529, doi:https://doi.org/10.1016/j.rse.2011.08.014.
- Tang, R., Z.-L. Li, Y. Jia, C. Li, X. Sun, W. P. Kustas, and M. C. Anderson (2011), An intercomparison of three remote sensing-based energy balance models using large aperture scintillometer measurements over a wheat-corn production region, *Remote Sensing of Environment*, *115*(12), 3187 – 3202, doi:https://doi.org/10.1016/j.rse.2011.07.004.
- Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004a), The Gravity Recovery and Climate Experiment: Mission overview and early results, *Geophysical Research Letters*, *31*(9), L09,607, doi:10.1029/2004GL019920, 109607.
- Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins (2004b), Grace measurements of mass variability in the Earth system, *Science*, *305*(5683), 503–505, doi:10.1126/science.1099192.
- Thomas, A. C., J. T. Reager, J. S. Famiglietti, and M. Rodell (2014), A GRACE-based water storage deficit approach for hydrological drought characteriza-

- tion, *Geophysical Research Letters*, 41(5), 1537–1545, doi:10.1002/2014GL059323, 2014GL059323.
- Vahmani, P., and T. S. Hogue (2014), High-resolution land surface modeling utilizing remote sensing parameters and the NOAA UCM: a case study in the los angeles basin, *Hydrology and Earth System Sciences*, 18(12), 4791–4806, doi:10.5194/hess-18-4791-2014.
- van Beek, L. P. H., Y. Wada, and M. F. P. Bierkens (2011), Global monthly water stress: 1. Water balance and water availability, *Water Resources Research*, 47(7), W07,517, doi:10.1029/2010WR009791, w07517.
- Voss, K. A., J. S. Famiglietti, M. Lo, C. de Linage, M. Rodell, and S. C. Swenson (2013), Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the tigris-euphrates-western iran region, *Water Resources Research*, 49(2), 904–914, doi:10.1002/wrcr.20078.
- Wahr, J., M. Molenaar, and F. Bryan (1998), Time variability of the Earth’s gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *Journal of Geophysical Research: Solid Earth*, 103(B12), 30,205–30,229, doi:10.1029/98JB02844.
- Wahr, J., S. Swenson, and I. Velicogna (2006), Accuracy of GRACE mass estimates, *Geophysical Research Letters*, 33(6), L06,401, doi:10.1029/2005GL025305, 106401.
- Wang, H., P. Wu, and Z. Wang (2006), An approach for spherical harmonic analysis of non-smooth data, *Computers & Geosciences*, 32(10), 1654 – 1668, doi:https://doi.org/10.1016/j.cageo.2006.03.004.
- Wang, H., H. Guan, H. A. Gutiérrez-Jurado, and C. T. Simmons (2014), Examination of water budget using satellite products over Australia, *Journal of Hydrology*, 511(Supplement C), 546 – 554, doi:https://doi.org/10.1016/j.jhydrol.2014.01.076.
- Wang, K., and R. E. Dickinson (2012), A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, *Reviews of Geophysics*, 50(2), RG2005, doi:10.1029/2011RG000373, rG2005.
- Wang, L., P. D’Odorico, J. P. Evans, D. J. Eldridge, M. F. McCabe, K. K. Caylor, and E. G. King (2012), Dryland ecohydrology and climate change: critical issues and technical advances, *Hydrology and Earth System Sciences*, 16(8), 2585–2603, doi:10.5194/hess-16-2585-2012.

- Wang, Y. P. (2000), A refinement to the two-leaf model for calculating canopy photosynthesis, *Agricultural and Forest Meteorology*, *101*(2), 143 – 150, doi:[https://doi.org/10.1016/S0168-1923\(99\)00165-3](https://doi.org/10.1016/S0168-1923(99)00165-3).
- Wang, Y.-P., and R. Leuning (1998), A two-leaf model for canopy conductance, photosynthesis and partitioning of available energy I: Model description and comparison with a multi-layered model, *Agricultural and Forest Meteorology*, *91*(1), 89 – 111, doi:[https://doi.org/10.1016/S0168-1923\(98\)00061-6](https://doi.org/10.1016/S0168-1923(98)00061-6).
- Wang, Y. P., E. Kowalczyk, R. Leuning, G. Abramowitz, M. R. Raupach, B. Pak, E. van Gorsel, and A. Luhar (2011), Diagnosing errors in a land surface model (CABLE) in the time and frequency domains, *Journal of Geophysical Research: Biogeosciences*, *116*(G1), G01,034, doi:10.1029/2010JG001385, g01034.
- Wood, E. F., J. K. Roundy, T. J. Troy, L. P. H. van Beek, M. F. P. Bierkens, E. Blyth, A. de Roo, P. Döll, M. Ek, J. Famiglietti, D. Gochis, N. van de Giesen, P. Houser, P. R. Jaffé, S. Kollet, B. Lehner, D. P. Lettenmaier, C. Peters-Lidard, M. Sivapalan, J. Sheffield, A. Wade, and P. Whitehead (2011), Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth’s terrestrial water, *Water Resources Research*, *47*(5), W05,301, doi:10.1029/2010WR010090, w05301.
- Xing, X. Q., and M. Damodaran (2005), Application of simultaneous perturbation stochastic approximation method for aerodynamic shape design optimization, *AIAA Journal*, *43*(2), 284–294, doi:10.2514/1.9484.
- Yang, K., L. Zhu, Y. Chen, L. Zhao, J. Qin, H. Lu, W. Tang, M. Han, B. Ding, and N. Fang (2016), Land surface model calibration through microwave data assimilation for improving soil moisture simulations, *Journal of Hydrology*, *533*(Supplement C), 266 – 276, doi:<https://doi.org/10.1016/j.jhydrol.2015.12.018>.
- Yang, Y., D. Long, H. Guan, W. Liang, C. Simmons, and O. Batelaan (2015), Comparison of three dual-source remote sensing evapotranspiration models during the musoexe-12 campaign: Revisit of model physics, *Water Resources Research*, *51*(5), 3145–3165, doi:10.1002/2014WR015619.
- Yang, Z.-L., G.-Y. Niu, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, L. Longueveergne, K. Manning, D. Niyogi, M. Tewari, and Y. Xia (2011), The community NOAA land surface model with multiparameterization options (NOAH-MP): 2. Evaluation over global river basins, *Journal of Geophysical Research: Atmospheres*, *116*(D12), D12,110, doi:10.1029/2010JD015140, d12110.

- Yesubabu, V., C. V. Srinivas, S. Langodan, and I. Hoteit (2016), Predicting extreme rainfall events over Jeddah, Saudi Arabia: impact of data assimilation with conventional and satellite observations, *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 327–348, doi:10.1002/qj.2654.
- Yesubabu, V., H. P. Dasari, S. Langodan, V. S. Challa, and I. Hoteit (2017), Climatic features of the red sea from a regional assimilative model, *International Journal of Climatology*, *37*(5), 2563–2581, doi:10.1002/joc.4865.
- Zeng, Y., Z. Xie, Y. Yu, S. Liu, L. Wang, J. Zou, P. Qin, and B. Jia (2016), Effects of anthropogenic water regulation and groundwater lateral flow on land processes, *Journal of Advances in Modeling Earth Systems*, *8*(3), 1106–1131, doi:10.1002/2016MS000646.
- Zhang, H., L. Zhang, and B. Pak (2011), Comparing surface energy, water and carbon cycle in dry and wet regions simulated by a land-surface model, *Theoretical and Applied Climatology*, *104*(3), 511–527, doi:10.1007/s00704-010-0364-x.
- Zhang, Y., R. Leuning, L. B. Hutley, J. Beringer, I. McHugh, and J. P. Walker (2010), Using long-term water balances to parameterize surface conductances and calculate evaporation at 0.05 spatial resolution, *Water Resources Research*, *46*(5), W05,512, doi:10.1029/2009WR008716, w05512.
- Zhang, Y., R. Leuning, F. H. S. Chiew, E. Wang, L. Zhang, C. Liu, F. Sun, M. C. Peel, Y. Shen, and M. Jung (2012), Decadal trends in evaporation from global energy and water balances, *Journal of Hydrometeorology*, *13*(1), 379–391, doi:10.1175/JHM-D-11-012.1.
- Zhuang, Q., and B. Wu (2015), Estimating evapotranspiration from an improved two-source energy balance model using aster satellite imagery, *Water*, *7*(12), 66736688, doi:10.3390/w7126653.
- Zmijewski, K., and R. Becker (2014), Estimating the effects of anthropogenic modification on water balance in the Aral sea watershed using GRACE: 200312, *Earth Interactions*, *18*(3), 1–16, doi:10.1175/2013EI000537.1.
- Zobler, L. (1999), Global soil types, 1-degree grid (Zobler), doi:10.3334/ornldaac/418.