

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/327618852>

Report on the improved Water Resources Reanalysis

Technical Report · June 2017

DOI: 10.13140/RG.2.2.14523.67369

CITATIONS

0

READS

30

12 authors, including:



Emanuel Dutra

University of Lisbon

135 PUBLICATIONS 2,328 CITATIONS

[SEE PROFILE](#)



Gianpaolo Balsamo

European Center For Medium Range Weather Forecasts

170 PUBLICATIONS 14,348 CITATIONS

[SEE PROFILE](#)



Jean-Christophe Calvet

Centre National de Recherches Météorologiques

298 PUBLICATIONS 7,512 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Copernicus Global Land Service [View project](#)



GloFAS [View project](#)



Global Earth Observation for integrated water resource assessment

Report on the improved Water Resources Reanalysis

Deliverable No: D.5.2 – Report on the improved Water Resources Reanalysis (WRR
tier 2) Ref.: WP5 - Task 5.1
Date: May 2017

This project has received funding from the European Union's Seventh Programme for
research, technological development and demonstration under grant agreement No.
603608



Deliverable Title	D.5.2 – Report on the improved Water Resources Reanalysis
Filename	E2O_D52_v1.docx
Authors	Emanuel, Dutra (ECMWF)
Contributors	Gianpaolo, Balsamo (ECMWF) Jean-Christophe, Calvet (Meteo-France) Simon Munier (Meteo-France) Sophia Burke (AmbioTEK) Gabriel Fink (Uni Kassel) Albert Van Dijk (Australian National University/CSIRO) Alberto Martinez-de la Torre (CEH) Rens van Beek (Uni. Utrecht) Ad De Roo (JRC) Jan Polcher (CNRS)
Reviewer	Wouter Dorigo (TUW)
Date	6/06/2017

Prepared under contract from the European Commission
 Grant Agreement No. 603608
 Directorate-General for Research & Innovation (DG Research), Collaborative project, FP7-ENV-2013-two-stage

Start of the project: 01/01/2014
 Duration: 48 months
 Project coordinator: Stichting Deltares, NL

Dissemination level

<input checked="" type="checkbox"/> PU	Public
<input type="checkbox"/> PP	Restricted to other programme participants (including the Commission Services)
<input type="checkbox"/> RE	Restricted to a group specified by the consortium (including the Commission Services)
<input type="checkbox"/> CO	Confidential, only for members of the consortium (including the Commission Services)

Deliverable status version control

Version	Date	Author	Reviewer	Singed-off by
0.0	03/05/2017	Emanuel, Dutra	-	-
0.1	6/05/2017	Emanuel Dutra	Wouter Dorigo	



Table of Contents

Table of Contents	i
List of Figures.....	iii
List of Tables	vii
1 Executive Summary	1
2 Introduction	3
3 Modelling protocol.....	4
3.1 Meteorological forcing.....	4
3.1.1 MSWEP precipitation	5
3.1.2 Topographic adjustment	7
3.2 Simulations setup	11
3.2.1 WRR2 ensemble	11
4 Modelling systems	12
4.1 HTESSEL-CaMa	13
4.1.1 Model description	13
4.1.2 Data assimilation	14
4.1.3 Evaluation	15
(a) WRR2	15
(b) Data assimilation.....	21
4.2 Jules	26
4.2.1 Model description	26
4.2.2 Evaluation	27
4.3 LISFLOOD	31
4.3.1 Model description	31
4.3.2 Evaluation	32
(a) Impact of the model improvements	32
(b) Impact of the new forcing.....	33
(c) Impact of the precipitation ensemble forcing.	40
4.4 ORCHIDEE.....	41
4.4.1 Model description	41
(a) New ancillary data.....	41
(b) New formulation for surface roughness	42
(c) New snow scheme	42
(d) Soil freezing scheme	42
(e) New routing scheme	43
4.4.2 Data assimilation	43
4.4.3 Evaluation	43

(a)	The Kura basin	43
(b)	The Rhone and Po	45
4.5	PCR-GLOBWB	46
4.5.1	Model description	46
4.5.2	Data assimilation	48
4.5.3	Evaluation	49
4.6	SURFEX-TRIP.....	50
4.6.1	Model description	50
4.6.2	Data assimilation	51
4.6.3	Evaluation	52
(a)	Impact of the new forcing.....	52
(b)	Impact of the model improvements	56
(c)	Analysis of the SURFEX LDAS-Monde	58
4.7	W3	64
4.7.1	Model description	64
4.7.2	Data assimilation	65
4.7.3	Evaluation	66
4.8	WaterGAP3	66
4.8.1	Model description	66
4.8.2	Model improvements with respect to WRR1	68
4.8.3	Evaluation	68
(a)	The global average water balance	69
(b)	The global spatial distribution of major water balance components.....	69
(c)	Testing of WaterGAP 3.2 reservoir management scheme with EO data	72
(d)	Effects on the improvement on water availability and discharge	76
4.9	WaterWorld.....	77
5	Data	80
5.1	Data access.....	80
5.1.1	Meteorological forcing	80
5.1.2	Model output convention	80
6	Summary	87
7	References	88

List of Figures

Figure 3-1: Flow chart summarizing the main steps of the MSWEP methodology. The different colours represent different time scales. P_n refers to normalized precipitation (i.e. P anomalies)	6
Figure 3-2: Example of the different weights for the 25 th of April 2006: (a) total weight, (b) relative gauge component, (c) relative satellite component and (d) relative reanalysis component.....	7
Figure 3-3: Mean topography of (a) ERAI, (b) WRR1, (c) WRR2, (d) ERAI-WRR2 and (e) WRR1-WRR2.....	8
Figure 3-4: ELR estimates from ERAI in January (left) and July (right): multi-year mean (top), inter-annual variability (middle) and uncertainty (bottom).	9
Figure 3-5: Stations altitude difference with respect to ERAI (a) and evolution of the number of stations (b). The legend in (b) denotes the total number of unique stations with more than 5 years of data and the standard deviation of the elevation error for all stations (black), stations below the model orography (blue) ad stations above the model orography (red).	10
Figure 3-6: Mean statistics of the root mean square error (RMSE) for the stations below the WRR2 orography (a,b), above the WRR2 orography (c,d) and all stations (e,f). The top panels (a,c,e) denote the actual RMSE, while the bottom panels (b,d,f) the normalized difference with respect to WFDEI025. The error bars denote the 95% confidence interval of the mean over the stations. In all panels the statistics are shown for the full period (ALL), and the 4 seasons on the horizontal axis. WFDEI denotes the forcing used in WRR1, WFDEI025 is a bilinear interpolation of WFDEI to a 0.25° grid, EI05 and EI025 are the ERAI data with a bilinear interpolation to 0.5° and 0.25° grids, respectively, and E20 is the WRR2 forcing (based in EI025 with the ELR corrections).	10
Figure 4-1: Mean annual cycle of precipitation (top left), evapotranspiration (top right), runoff (bottom left) and discharge (bottom right) over the Amazon basin in WRR1, WRR1_M2, WRR2_F2 and WRR2. On the discharge panel, the observations from GRDC are the red line.....	16
Figure 4-2: Mean annual cycle of evapotranspiration (top left), soil evaporation (top right), vegetation transpiration (bottom left) and canopy interception evaporation (bottom right) over the Amazon basin in WRR1, WRR1_M2, WRR2_F2 and WRR2.	17
Figure 4-3: Mean evapotranspiration (top left) and canopy interception evaporation (top right) in WRR1 and differences between WRR2 and WRR1 (bottom panels)	18
Figure 4-4: As Figure 4-1 but for the Ob basin.	19
Figure 4-5: Mean annual discharge in Yukon (top left), Mackenzie (top right), Danube (bottom left) and Mississippi (bottom right). Simulations are compared against the GRDC observations (red) and the mean annual cycle computed only for the period with available observations.	20
Figure 4-6: (a) temporal correlation of between daily surface soil moisture and EO ESA CCI soil moisture for the period 1988 to 2012 in WRR1. Correlation differences of different model/forcing configurations and WRR1: b) WRR1_M2 – WRR1, c) WRR1_F2 – WRR1 and d) WRR2-WRR1.	21
Figure 4-7: As Figure 4-6 but for the anomalies after removing a 35 days moving window in both the simulations and EO CCI.	21
Figure 4-8: Mean annual cycle of snow depth verification over Eurasia using GHCN stations for the 2000-2010 period (grey line with the number of stations: right	

vertical axis; dotted symbols with the mean observations of snow depth (m)]. The dashed lines represent the simulation bias and the solid lines the standard deviation of the error for: WRR1 (black), WRR1_M2 (grey), WRR1_F2 (dark magenta), WRR2 (magenta) and WRR2_DA (blue)	22
Figure 4-9: As Figure 4-8 but for Canada.....	22
Figure 4-10: As Figure 4-8 but for United States.....	22
Figure 4-11: Mean annual discharge in Yukon (top left), Mackenzie (top right), Danube (centre left), Mississippi (centre right), Ob (bottom left), and Yenisei (bottom right). Simulations are compared against the GRDC observations (red) and the mean annual cycle computed only for the period with available observations between 2000-2012 (available WRR2_DA).....	24
Figure 4-12: Mean annual snow mass increments in the WRR2_DA data assimilation simulation in Yukon (top left), Mackenzie (top right), Danube (centre left), Mississippi (centre right), Ob (bottom left), and Yenisei (bottom right).....	25
Figure 4-13: Saturation fraction calculated by the PDM scheme in JULES as a function of soil water storage to generate saturation excess runoff. Left: Variability obtained by the b parameter, for a fixed $S_0/S_{max}=0$. Right: Variability of the S_0 parameter (green curves) and variability of b for a fixed $S_0/S_{max}=0.5$	26
Figure 4-14: River flow at Kingston (UK), outlet of the Thames catchment simulated. Green lines are flow observation at the gauge station. Blue lines are river flow simulated by JULES with the best possible parameterization before introducing the S_0 linear dependency on terrain slope (left) and with the new parameters adopted for WRR2 (right). Total Bias and Nash-Sutcliffe efficiency metrics are given on top of the plots.....	27
Figure 4-15: Mean Global Runoff ($kg\ m^{-2}\ d^{-1}$) from the GCRF-OBS product (Fekete et al., 2002). White areas are not represented by the dataset (based on catchments)....	28
Figure 4-16: Global shift (days) between the maximum values in the WRR simulations (WRR1_F1, right; WRR2_F1, left) and the GCRF-OBS runoff benchmarking dataset.....	29
Figure 4-17: RMSE Score (0-1) over the Europe region obtained by the ILAMB evaluation against the GCRF-OBS runoff benchmarking dataset (WRR1_F1, right; WRR2_F1, left).....	29
Figure 4-18: Monthly climatology for runoff ($kg\ m^{-2}\ d^{-1}$) averaged over the Europe region (Figure 4-18) by the GCRF-OBS product (grey), WRR1_F1 (red) and WRR2_F1 (cyan).....	30
Figure 4-19: Observed and LISFLOOD simulated monthly discharge (m^3/s) for the Rhine basin.....	35
Figure 4-20 Observed and LISFLOOD simulated monthly discharge (m^3/s) for the Amazon river	35
Figure 4-21 Estimated evapotranspiration deficit (LISFLOOD WRR2 with WFDEI forcing).....	37
Figure 4-22 Estimated evapotranspiration deficit (LISFLOOD WRR2 with MSWEP forcing).....	37
Figure 4-23 Estimated average water availability per capita (LISFLOOD WRR2 with WFDEI forcing).....	38
Figure 4-24 Estimated average water availability per capita (LISFLOOD WRR2 with MSWEP forcing).....	38
Figure 4-25 Estimated average water exploitation index (LISFLOOD WRR2 with WFDEI forcing).....	39

Figure 4-26 Estimated average water exploitation index (LISFLOOD WRR2 with MSWEP forcing).....	39
Figure 4-27 Estimated average water dependency (LISFLOOD WRR2 with WFDEI forcing).....	40
Figure 4-28 Estimated average water dependency (LISFLOOD WRR2 with MSWEP forcing).....	40
Figure 4-29: Precipitation and simulated evaporation averaged over the Kura basin. The dashed lines in the lower panel are potential evaporation estimated by ORCHIDEE with the forcing. The 3 simulations presented in these figures are : WRF, E2OFD and WFDEI forcing.	44
Figure 4-30: Discharge simulated by ORCHIDEE and observed for the Kura river at the Surra	45
Figure 4-31: Observed and simulated discharge of the Rhone at the Beaucaire station.	46
Figure 4-32: Observed and simulated discharge of the Po at the Pontelagoscuro station.....	46
Figure 4-33: Scheme of the WRR2 model version, including the representation of floodplains and groundwater reservoir.	51
Figure 4-34: Comparison of Nash-Sutcliffe Efficiency for experiments WRR1_F1 (model WRR1, forcings WRR1) and WRR1_F2 (model WRR1, forcings WRR2).	53
Figure 4-35: Comparison of the accumulated distribution of NSE over different regions of the globe, for both experiments WRR1_F1 and WRR1_F2.....	54
Figure 4-36: Seasonal cycle of the runoff for both WRR1_F1 and WRR1_F2 experiments and comparison with in situ observations.	55
Figure 4-37: Comparison of Nash-Sutcliffe Efficiency for experiments WRR1_F2 (model WRR1, forcings WRR2) and WRR2_F2 (model WRR2, forcings WRR2).	56
Figure 4-38: Accumulated distribution of NSE over different regions of the globe, for both experiments WRR1_F2 and WRR2_F2.....	57
Figure 4-39: Seasonal cycle of the runoff for both WRR1_F2 and WRR2_F2 experiments and comparison with in situ observations.	58
Figure 4-40: Correlation and Root Mean Square Difference (RMSD, in m^2/m^2) of LAI for experiments without data assimilation (Model) and with data assimilation (Analysis).	59
Figure 4-41: Correlation and Root Mean Square Difference (RMSD, in m^3/m^3) of SSM for experiments without data assimilation (Model) and with data assimilation (Analysis).	60
Figure 4-42: Comparison of LAI from open-loop, analysis and observation over Europe.	61
Figure 4-43: LAI seasonal scores of the open-loop and analysis over Europe: Standard Deviation (SDD), Bias, Root Mean Square Difference (RMSD) and correlation.....	62
Figure 4-44: Time series of analysis increments averaged over Europe for both control variables. WG2 is the water content in the second soil layer (1cm-4cm). Units are m^2/m^2 for LAI increments and m^3/m^3 for WG2 increments.	62
Figure 4-45: Average over the whole period of Runoff, Drainage and Evapotranspiration for the open-loop and analysis experiments, and impact of the assimilation. Units are $kg/m^2/day$	63
Figure 4-46: Simplified conception structure of W3 v2.0	64

Figure 4-47: Example of additional evaporation from non-rainfall sources estimated for 03 July 2012 over Pakistan and NW India. Values range from 0 (white) to 5 mm/d (blue).....	66
Figure 4-48: Global distribution of mean annual precipitation: a) WRR2 1979-2014 b) difference to WRR1 (WRR2-WRR1) 1979-2012.	70
Figure 4-49: Global distribution of mean annual evapotranspiration: a) WRR2 1979-2014 b) difference to WRR1 (WRR2-WRR1) 1979-2012.	71
Figure 4-50: Global distribution of mean annual runoff: a) WRR2 1979-2014 b) difference to WRR1 (WRR2-WRR1) 1979-2012.	72
Figure 4-51: Calculated Water volumes and observed water levels of the Chora Bassa Reservoir, Mozambique, for the period 2002 - 2014.	73
Figure 4-52: Calculated Water volumes and observed water levels of the Kariba Reservoir for the period 1992-2004.	74
Figure 4-53: Calculated Water volumes and observed water levels of Lake Mead for the period 2002 to 2014.....	75
Figure 4-54: Calculated Water volumes and observed water levels of Lake Victoria for the period 1992 to 2014.....	76
Figure 4-55: Modelled Nile River Discharges at Al Akhsas gauging station (GRDC ID: 1362100) for the period 2000-2010.	77
Figure 4-56: WorldClim (right) and TRMM (left) rainfall data in the Magdalena catchment.....	78
Figure 4-57: Slope angle (from Hydrosheds data) and areas of bare ground (from MODIS data) in the Magdalena catchment.	79
Figure 4-58: Water stress by global basin modelled at 10Km resolution.....	79

List of Tables

Table 3-1 Overview of the meteorological forcing and differences between WRR1 and WRR2.....	4
Table 3-2 Precipitation datasets used in the WRR2 ensemble.	11
Table 4-1: Modelling systems main changes from WRR1 to WRR2.....	12
Table 4-2: HTESSEL simulations details	15
Table 4-3: Summary table obtained applying the ILAMB system to the WRR1_F1 and WRR2-F2 climatology runoff fields. The overall score is a composition of the rest of 0-1 scores: Bias Score, RMSE Score, Seasonal Cycle Score and Spatial Distribution Score.	28
Table 4-4: Summary table obtained applying the ILAMB system to the WRR1_F1 and WRR2-F2 surface soil moisture (top 10 cm) monthly fields. The overall score is a composition of the rest of 0-1 scores: Bias Score, RMSE Score, Seasonal Cycle Score, Spatial Distribution Score and Interannual Variability Score.....	30
Table 4-5: Summary table obtained applying the ILAMB system to the WRR1_F1 and WRR2-F2 latent heat monthly fields. The overall score is a composition of the rest of 0-1 scores: Bias Score, RMSE Score, Seasonal Cycle Score, Spatial Distribution Score and Interannual Variability Score.....	31
Table 4-6: Kling-Gupta Efficiency (KGE) for monthly discharge, for the Rhine and Amazon basin, showing the effects of the LISFLOOD model changes.	33
Table 4-7: Kling-Gupta Efficiency (KGE) for daily discharge for selected gauging stations, showing the effects of the LISFLOOD model changes, for the period 1980-2008.....	33
Table 4-8: Kling-Gupta Efficiency (KGE) for daily discharge, for the Rhine and Amazon basin, showing the effects of the TIER2 forcing changes. For the Rhine, the comparison is also shown with the 5km resolution LISFLOOD setup, forced with the JRC meteo data.	34
Table 4-9: Kling-Gupta Efficiency (KGE) for daily discharge for selected gauging stations, showing the effects of the TIER2 forcing changes, for the period 1980-2008.	36
Table 4-10: Kling Gupta Efficiencies (KGE) for various rivers for the various WRR2-ENS simulations with various precipitation forcings.	41
Table 4-11: Parameter combinations for the brute force calibration. 24 parameter combinations are used ($3 \times 4 \times 2$). The default parameter setting (Run 10) has been omitted in favour of Run 24.	48
Table 4-12: Model setups evaluated in WRR2. All model runs cover the period 1979-2014 with the exception of WRR1_Nat30 which covers the period of the WRR1 forcing 1979-2012	49
Table 4-13: Global average water balance calculated by WaterGAP3.1 (WRR1) and WaterGAP3.2 (WRR2)	69
Table 4-14: The WaterGAP3 modelled Nile discharges statistics (2000-2012) of different WRR settings.....	77
Table 5-1: Forcing data	80
Table 5-2: Institution identifiers in the model output.....	81
Table 5-3: Experiment version names, associated start and end year and domain....	81
Table 5-4: List of output variables and conventions	82
Table 5-5: Model output locations.....	85

1 Executive Summary

This report describes the tier-2 water resources reanalysis (WRR2) development, implementation and data structure and storage. The report is organized in 5 sections and their main highlights are described in the following paragraphs.

The Earth2Observe WRR1 dataset consisted of an ensemble of ten global hydrological and land surface models for the 1979-2012 period using a meteorological forcing dataset based on reanalysis. The WRR1 dataset was developed using state-of-the-art modelling systems and provides a robust benchmark for further improvements in the coming years. WRR2 was based on WRR1 by exploring several potential sources of improvements regarding: (i) meteorological forcing data; (ii) modelling systems processes description; and (iii) assimilation of Earth Observations.

The modelling protocol of WRR2 followed WRR1 closely in terms of the data structure, storage and simulations setup. The main difference was the increased resolution from $0.5^\circ \times 0.5^\circ$ in WRR1 to $0.25^\circ \times 0.25^\circ$ in WRR2 and a new meteorological forcing. The meteorological forcing is based on ERA-Interim, extending from 1979 to 2014, with a topographic adjustment of temperature and using a new precipitation dataset. The new precipitation dataset, MSWEP, is a combined product of atmospheric reanalysis, satellite estimates and in-situ observations. In addition to the 1979-2014 baseline WRR2 simulations, each modelling system also carried out simulations for a shorter period: 2000-2013 using four different satellite precipitation products, which are not available prior to 2000.

WRR2 consists of eight global land surface and hydrological models with several modifications with respect to WRR1 (two of the voluntary models that participated in WRR1 did not take part in WRR2). The model changes include a variety of process oriented developments (e.g. snowpack, runoff generation), model parameters calibration, and representation of water use (e.g. withdrawal, consumption, management). The changes were not systematic, i.e. each model addresses different aspects, mainly driven by the expertise of the group developing the respective model. While this approach does not allow for an exact evaluation of which model improvement is more important, it does allow for a variety of approaches and methodologies, which are expected to partially encompass the intrinsic uncertainty of large-scale water resources estimates.

Following the technical developments of WRR1, the WRR2 data structure and storage were designed to facilitate, as much as possible, user access. The total model output is about 5 TB, and all efforts have been made to harmonize the data format and access via the water cycle integrator (WCI¹) data portal.

This report focuses on the technical details of the simulations and the description of the modelling system. Each modelling system presents an evaluation of its developments along with the impact of the meteorological forcing. Due to the different nature of each modelling and data assimilation system changes, a consistent evaluation among models was not pursued in this report. Each modelling system was evaluated independently, and this report serves as a detailed description of the

¹ <https://wci.earth2observe.eu/>

dataset and not an integrated quality assessment of WRR2 dataset. Such analysis will be presented in the following D5.3 project report. Additionally, the different precipitation forcing will be used to generate an ensemble, WRR-ENS which will be described and evaluated in D5.4. However, among the several evaluation examples presented in this report there is a clear signal of the improvements due to the new meteorological forcing used in WRR2. This is mainly attributed to the new precipitation dataset (MSWEP) which leads to improvements of the water cycle simulations. Each modelling system change also shows improvements, but these tend to be restricted to the main processes addressed by these improvements. Overall, WRR2 brings a significant improvement in terms of the land water cycle representation when compared with WRR1. This new dataset will have an important impact on the user community and applications due to its higher resolution and state-of-the art modelling systems and data assimilation.

2 Introduction

Water resources concern local to global economies independently of their level of development. Pressures from demographic, economic and climatic changes further reinforce the importance of an accurate and reliable estimate of the different components of the land water cycle. On global scale, there is a limited number of global reanalysis datasets that can support water resources (e.g. Rodell et al., 2004; Dirmeyer et al., 2006; Reichle et al., 2011; Balsamo et al., 2015; Haddeland et al., 2011; Reichle et al., 2017; van Dijk et al., 2014). However, only Haddeland et al. (2011) combined in WATCH both global land surface and hydrological models into a single multi-model ensemble. In Earth2Observe, this has been further explored with the release of the water resources reanalysis tier-1 dataset (WRR1, see report D5.1 and Schellekens et al., 2016).

The Earth2Observe WRR1 dataset consists of an ensemble of ten global hydrological and land surface model outputs for the period 1979-2012 using a reanalysis-based meteorological forcing dataset. This dataset was developed using state-of-the-art modelling systems and provides a robust benchmark for further improvements in the coming years. With respect to WRR1, there are several potential sources of improvement including: (i) meteorological forcing data; (ii) modelling systems processes description; and (iii) assimilation of new Earth Observations.

Following the achievements of the WRR1 datasets, this report presents the tier-2 version (WRR2) which aims at exploring several potential sources of improvements of the water resources multi-model dataset. The main developments include: (i) increased resolution from $0.5^\circ \times 0.5^\circ$ to $0.25^\circ \times 0.25^\circ$ and a new meteorological forcing extending from 1979 to 2014; (ii) numerous modelling system updates (from process description to water use) and (iii) data assimilation of different Earth observations. As in WRR1, all simulations have a common meteorological forcing. Each modelling system and data assimilation application was developed independently by each group to combine and represent, as much as possible, the intrinsic uncertainty of these estimates. This last point is further explored by providing a special ensemble of model simulations for the period 2000-2013 using four different satellite precipitation products.

The main goal of this report is to provide a detailed description of the new WRR2 dataset, including the simulations setup, modelling systems description and output data structure and storage. While each of the modelling systems description includes an evaluation of the independent model and data assimilation enhancements, an integrated evaluation of WRR2 vs WRR1 is not presented as this will be further explored in upcoming project reports.

The report is organized in three main sections. The next section describes the modelling protocol, including the meteorological forcing and simulations setup. Section 4 presents the detailed description of each participating modelling system along with an initial evaluation of the new simulations. Section 5 describes the models output structure and location on the data server and a final overview is presented on the last summary section.

3 Modelling protocol

The modelling protocol for the tier-2 water resources reanalysis (WRR2) follows closely the tier-1 reanalysis and definition regarding the model setup and output formats (WRR1, see the D5.1 report² and Schellekens et al. (2016)). This simplifies the comparison between the two WRR datasets, the data storage and user access. Compared to WRR1, the main changes of WRR2 include i) increased horizontal resolution ii) a new meteorological forcing and iii) the use of multi-precipitation product ensemble. To accommodate the different requirements of WRR2, the comparison with WRR1 and assessment of the uncertainty in precipitation and role of Earth observations (EO), WRR2 contains two data streams: i) baseline from 1979 to 2014 and ii) multi-precipitation ensemble from 2000 to 2013. In the following sections the details of the meteorological forcing, simulations setup and data format are presented in detail.

3.1 Meteorological forcing

The meteorological forcing used in WRR1, the Watch Forcing Dataset ERA-Interim (hereafter WFDEI, Weedon et al., 2014) was not suitable for WRR2 due to the horizontal resolution limitation. Furthermore, the methodology used in WFDEI to correct the ERA-Interim reanalysis (ERA1, Dee et al., 2011) could not be directly applied to the higher resolution proposed for WRR2 as some of the observational datasets used in WFDEI are not available at higher resolutions (e.g. Climate Research Unit, CRU, number of wet day or diurnal temperature range).

The WRR2 forcing is based on the original ERA-Interim meteorological data bilinearly interpolated to 0.25°. It includes a topographic adjustment to temperature using a spatially-temporally varying environmental lapse rate (ELR) and a newly developed precipitation dataset the Multi-Source Weighted-Ensemble Precipitation (MSWEP, Beck et al., 2017) which are both described in the following sub-sections, and the data access in section 5.1.1. Hence both WRR1 and WRR2 are based on ERA1 3-hourly data differing in:

- Resolution from 0.5° in WRR1 to 0.25° in WRR2;
- Temporal extension: 1979-2012 in WRR1 to 1979-2014 in WRR2;
- Corrections applied to ERA1 detailed in Table 3-1.

Table 3-1 Overview of the meteorological forcing and differences between WRR1 and WRR2.

Variable	Definition	Units	Corrections in WRR1	Corrections in WRR2
Wind	Wind speed at 10 meters	m s ⁻¹	None	None

²

http://earth2observe.eu/files/Public%20Deliverables/D5.1_Report%20on%20the%20WRR1%20tier1.pdf

Variable	Definition	Units	Corrections in WRR1	Corrections in WRR2
Tair	Temperature 2 meters	K	Elevation using a constant ELR, CRU average and CRU average diurnal temperature range	Elevation using a variable ELR
Qair	Specific air humidity at 2 meters	kg kg ⁻¹	Via changes in Tair and Psurf	Via changes in Tair and Psurf
Psurf	Pressure at the surface	Pa	Via changes in Psurf	Via changes in Tair and Psurf
Swdown	Surface incident shortwave radiation	W m ⁻²	CRU average cloud cover and interannual changes in atmospheric aerosol loading	None
Lwdown	Surface incident longwave radiation	W m ⁻²	Via fixed relative humidity and changes in Tair, Psurf and Qair	None
Rainf	Rainfall rate (only liquid phase)	kg m ⁻² s ⁻¹	CRU number of wet days and precipitation totals	See MSWEP
Snow	Snowfall rate (only solid phase)	kg m ⁻² s ⁻¹	CRU number of wet days and precipitation totals	See MSWEP

3.1.1 MSWEP precipitation

The Multi-Source Weighted-Ensemble Precipitation (MSWEP) is a global precipitation dataset for the period 1979-2015 with a 3-hourly temporal resolution and 0.25° spatial resolution. The datasets were specifically designed for hydrological modelling and described in detail by Beck et al. (2017). MSWEP was designed to optimally merge the highest quality precipitation sources available as a function of time scale and location (see Figure 3-1).

The long-term mean of MSWEP is based on the CHPclim dataset (Funk et al., 2015) but replaced with more accurate regional datasets where available (over USA and New Zealand). Additionally, a correction for gauge under-catch and orographic effects was also applied by inferring catchment-average precipitation from streamflow observations at 13762 stations across the globe.

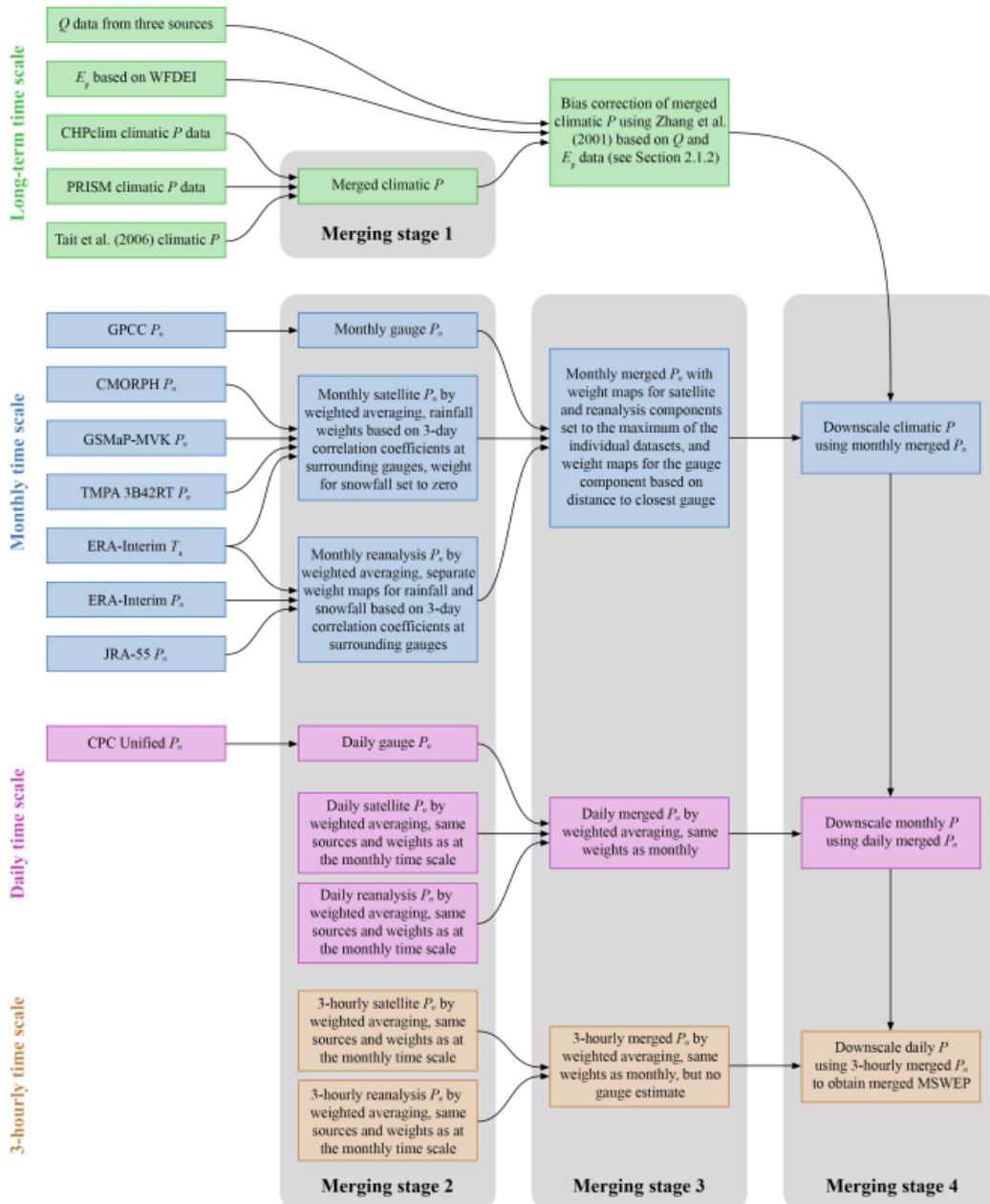


Figure 3-1: Flow chart summarizing the main steps of the MSWEP methodology. The different colours represent different time scales. P_n refers to normalized precipitation (i.e. P anomalies).

The temporal variability of MSWEP was determined by a weighted averaging of precipitation anomalies from seven datasets: two based solely on interpolation of gauge observations, three on satellite remote sensing, and two on atmospheric reanalysis. For each grid cell, the weight assigned to the gauge-based estimates was calculated from the gauge network density, while the weights assigned to the satellite- and reanalysis-based estimates were calculated from their comparative performance at the surrounding gauges. Figure 3-2 shows the sum of the weights used to merge the gauge, satellite and reanalysis component of MSWEP during the stage 3 (see Figure 3-1) at the daily time scale on an example day (25th April 2016), along with their relative contributions.

Beck et al. (2017) compared MSWEP against other state-of-the-art gauge adjusted precipitation datasets (including the data used in WRR1) using independent precipitation data from FLUXNET tower stations attaining the highest daily correlation coefficient among all the tested datasets. The performance of the HBV hydrological model was also improved when using MSWEP (Beck et al. 2017).

Considering the novelty of this dataset, the encouraging results, and use of Earth observations in this dataset, it was decided to use MSWEP as baseline precipitation in WRR2.

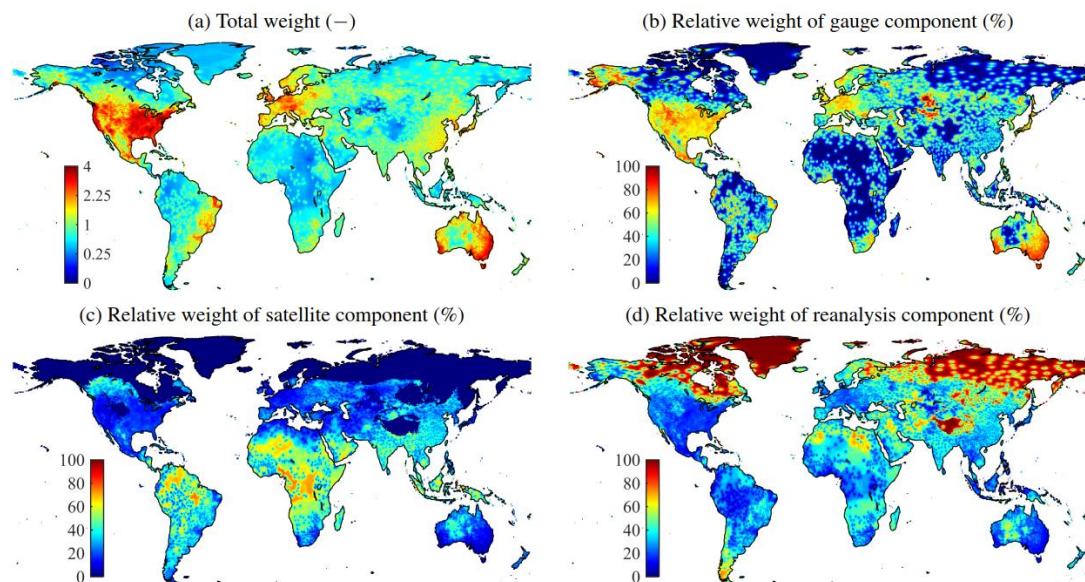


Figure 3-2: Example of the different weights for the 25th of April 2006: (a) total weight, (b) relative gauge component, (c) relative satellite component and (d) relative reanalysis component.

3.1.2 Topographic adjustment

The ERAI horizontal resolution of about 0.7° together with the required smoothing of the topography used by the model due to numerical constraints results in a very smooth topography with large differences compared to an average 0.25° topography. Therefore, just a simple bilinear interpolation of ERAI resolution to the 0.25° would not provide the spatial detail expected, in particular over complex terrain regions (see Figure 3-3).

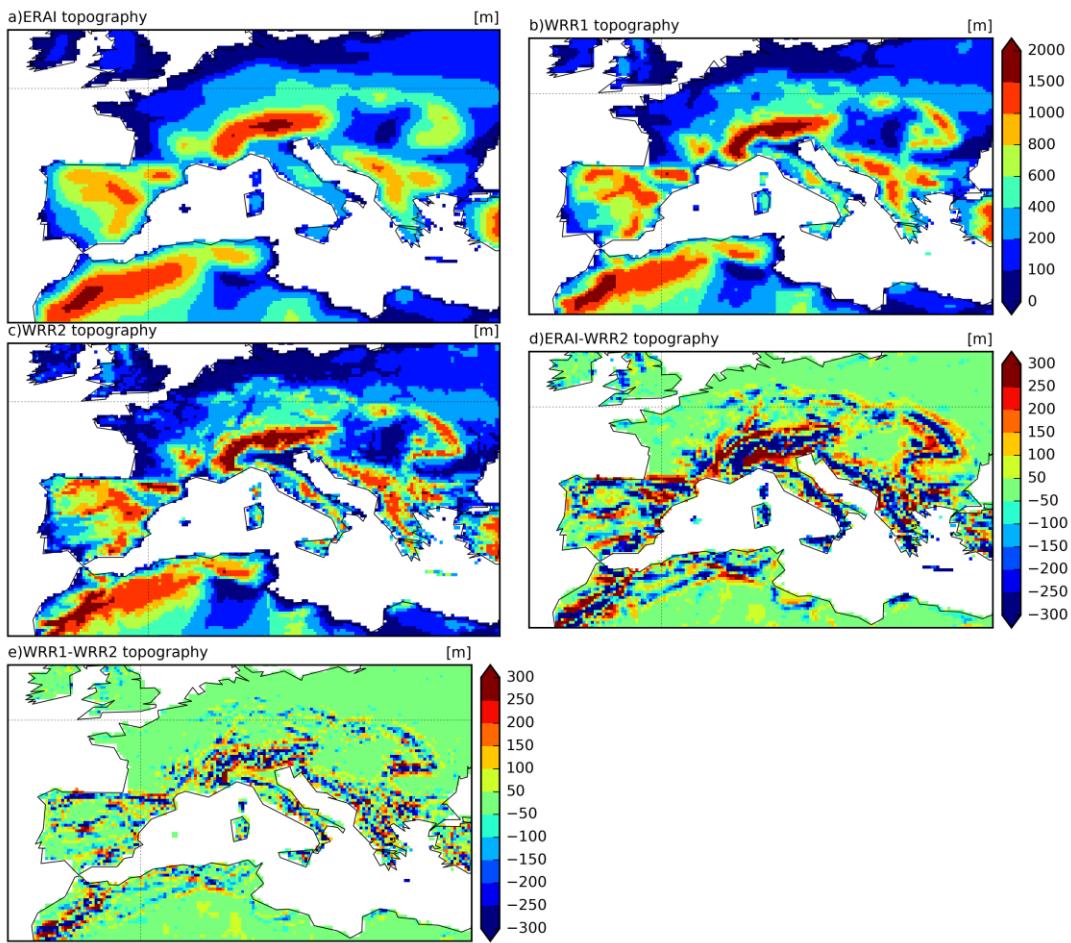


Figure 3-3: Mean topography of (a) ERAI, (b) WRR1, (c) WRR2, (d) ERAI-WRR2 and (e) WRR1-WRR2.

Spatial downscaling of temperature is a complex problem that can be addressed via dynamical downscaling, but which requires high resolution numerical simulations that are not computationally feasible on the global scale and for such a long period. The simplest approach is to apply a topographic adjustment to correct for elevation differences between the two topographies. This adjustment requires the definition of the environmental lapse rate (ELR), the rate at which temperature decreases with increasing altitude. It is common practice to use a spatially and temporally constant ELR of -6.5 K km^{-1} (as was in the case in WFDEI for WRR1). However, the ELR depends on the predominant large-scale circulation and other local features.

Gao et al. (2012) proposed that the ELR could be estimated from the vertical profiles of ERAI and tested the approach over the Swiss Alps. For WRR2, a similar methodology was used where the ELR was computed from the ERAI model levels mean monthly vertical profiles between 300 and 1300 meters above the surface. The first 300 meters near the surface are not considered to avoid local inversions. The ELR is computed as a temperature difference between several pairs of model levels in ERAI resulting in a mean ELR along with its standard deviation, which provides a measure of the calculation uncertainty. Figure 3-4 displays the mean ELR derived from ERAI using this methodology for January and July along with the inter-annual variability and uncertainty. This methodology leads to a clear annual cycle of the ELR with large values during the summer in the northern hemisphere and lower values during winter

in the northern hemisphere. Areas of large inter-annual variability and uncertainty are mainly associated with large-scale subsidence areas occurring mainly over the oceans and over the Northern continents during winter.

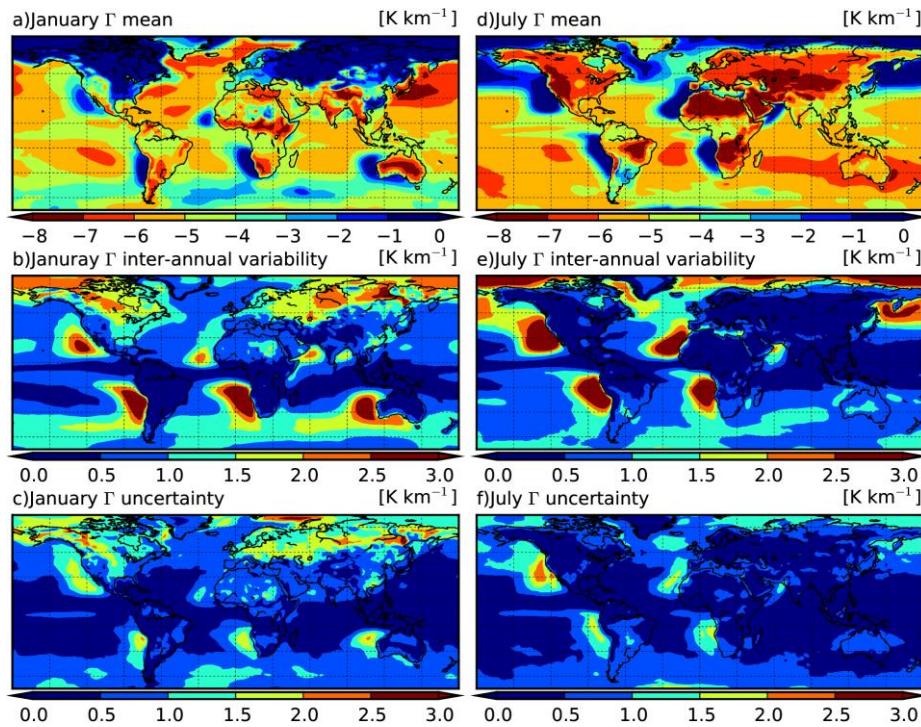


Figure 3-4: ELR estimates from ERAI in January (left) and July (right): multi-year mean (top), inter-annual variability (middle) and uncertainty (bottom).

This methodology was evaluated by comparing the temperature forcing with synoptic observations available on the Global Telecommunication System (GTS). Figure 3-5 displays the station locations and their altitude differences with respect to ERAI. Although GTS data is available globally, a European domain was selected due to its reasonably high station density. The root-mean-square-errors (RMSE) are displayed in Figure 3-6 comparing WRR1 and WRR2 forcing along with other intermediate approaches. The RMSE differences with respect to WRR1 with a bilinear interpolation to the 0.25 grid is also presented as this would have been the default option to prepare the WRR2 forcing. The results are aggregated by stations below the model orography (left panel), above the model orography (centre panels) and all stations (right panels). The ELR approach to correct for the topographic differences has a clear positive impact for the stations above the model orography with an average reduction of 15% of the RMSE. For the stations below the model orography the mean impact is small, and mainly visible during summer. This can be explained by local effects in valleys that are not possible to correct with this simple approach. Overall, considering all stations the selected methodology corrections improved the air temperature RMSE in about 6% comparing with a simple bilinear interpolation of WRR1.

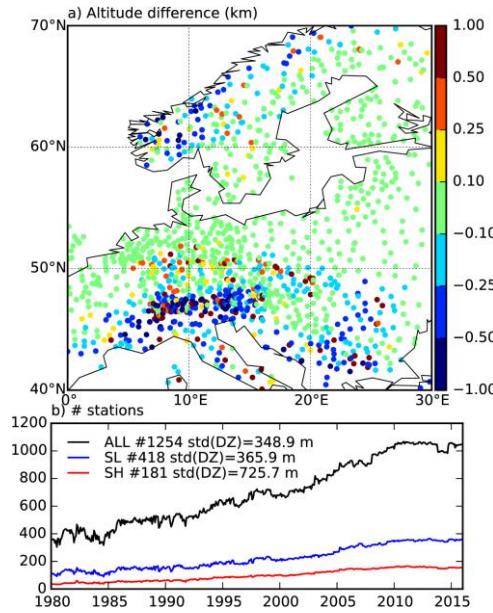


Figure 3-5: Stations altitude difference with respect to ERAI (a) and evolution of the number of stations (b). The legend in (b) denotes the total number of unique stations with more than 5 years of data and the standard deviation of the elevation error for all stations (black), stations below the model orography (blue) ad stations above the model orography (red).

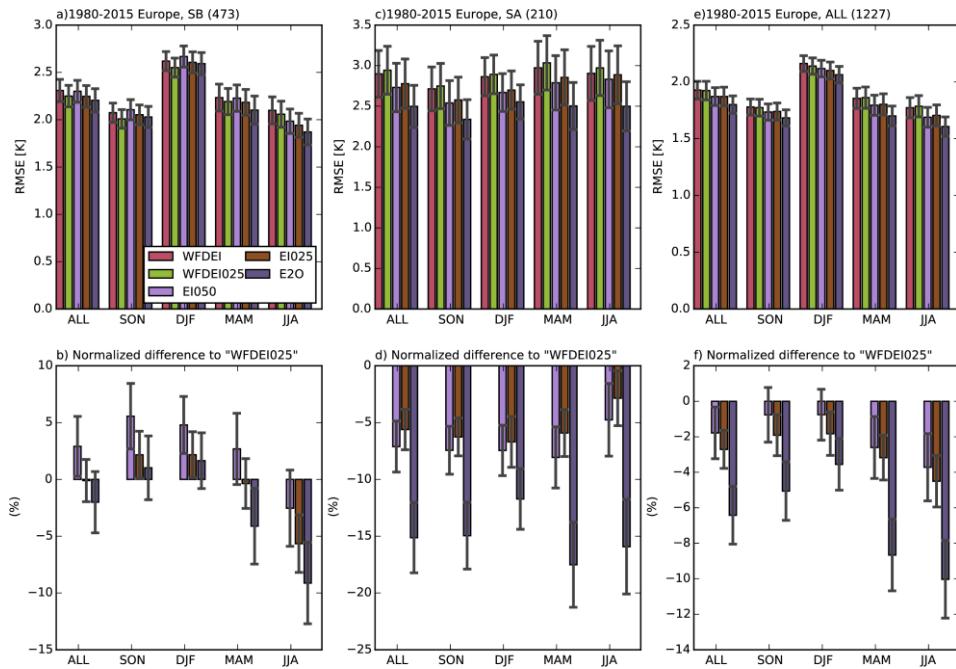


Figure 3-6: Mean statistics of the root mean square error (RMSE) for the stations below the WRR2 orography (a,b), above the WRR2 orography (c,d) and all stations (e,f). The top panels (a,c,e) denote the actual RMSE, while the bottom panels (b,d,f) the normalized difference with respect to WFDEI025. The error bars denote the 95% confidence interval of the mean over the stations. In all panels the statistics are shown for the full period (ALL), and the 4 seasons on the horizontal axis. WFDEI denotes the forcing used in WRR1, WFDEI025 is a bilinear interpolation of WFDEI to a 0.25° grid, EI05 and EI025 are the ERAI

data with a bilinear interpolation to 0.5° and 0.25° grids, respectively, and E20 is the WRR2 forcing (based in EI025 with the ELR corrections).

3.2 Simulations setup

The simulations protocol closely followed WRR1. For the baseline experiments, the simulations were performed from the 1st of January 1979 to the 31st of December 2014 in a continuous simulation without any kind of data assimilation. The initialization method of the model on the 1st of January 1979 is left to each modelling group as different models require different spin-up periods. To achieve a common spin-up period, the 1st year of simulation (1979) is considered as spin-up and was not used in any evaluation. Regarding static fields (e.g. land cover) each group used their own datasets as this is an integral component of each modelling system.

3.2.1 WRR2 ensemble

The WRR2 ensemble component follows the same protocol as the WRR2 baseline simulation, including the meteorological forcing except for the Rainfall data, which is replaced by four different satellite-derived products detailed in Table 3-2. The satellite products were processed to the same horizontal and temporal resolutions as WRR2 and replace the Rainfall data when there is no Snowfall. When WRR2 has Snowfall, the Rainfall was not replaced due to limitations of the satellite data to detect and quantify snowfall events. WRR2 ensemble is restricted to the 2000-2013 period, and the initial conditions on the 1st of January 2000 are taken from the WRR2 baseline simulation.

Table 3-2 Precipitation datasets used in the WRR2 ensemble.

Name	Product	Spatial resol.	Temporal resol.	Spatial domain	Temporal coverage	Reference
TRMM	TRMM Multi-Satellite precipitation Analysis (TMPA) 3B42 V7	0.25°	3-hourly	50°S-50°N	1998-2014	(Huffman et al., 2007)
TRMMRT	TRMM Multi-Satellite precipitation Analysis (TMPA) 3B42 RT	0.25°	3-hourly	50°S-50°N	03/2000-12/2013	(Huffman et al., 2007)
GSMAP	Global satellite Mapping of Precipitation Gauge calibrated reanalysis product	0.1°	hourly	60°S-60°N	03/2000-12/2013	(Ushio et al., 2009)
CMORPH	CPC MORPHing Technique V1 Bias-corrected	0.25°	3-hourly	60°S-60°N	2000-2013	(Joyce et al., 2004)

4 Modelling systems

This section presents the detailed description of the modelling systems participating in WRR2 along with an evaluation of their individual performance. Table 4-1 presents an overview of the main changes of the modelling systems between WRR1 and WRR2, which are detailed in the following sub-sections.

Table 4-1: Modelling systems main changes from WRR1 to WRR2

Modeling systems	Model changes	Data assimilation
HTESEL-CaMa	Improved representation of the snowpack with a multi-layer snow scheme; increased number of soil layers.	Optimal interpolation snow analysis using in-situ snow depth observations
JULES	Rainfall-runoff processes	N.A
LISFLOOD	Global scale at 0.1°, increased number of soil layers, groundwater abstraction	N.A
ORCHIDEE	Revision of the ancillary data, surface roughness, snow scheme, soil freezing and routing.	Derivation of background albedo and assimilation of river discharge
PCR-GLOBWB	Improvements to river routing reservoir schemes and water withdrawal and consumption	Assimilation of discharge and soil moisture
SURFEX-TRIP	Improvements in ground water, flood plains, land use, plant growth, surface energy, and snow.	Global LDAS
W3 (version 2.0)	Modified soil and groundwater hydrology equations, resolution increased to 0.05°, improved parameter estimates, dynamic data assimilation, evaporation of water not derived from rainfall (e.g. irrigation)	Some calibration with global streamflow, in situ soil moisture and FLUXNET data sets. Dynamic assimilation of MODIS-derived vegetation properties, albedo, snow cover, water dynamics, land surface temperature.
WaterGAP3	5 arc minute, assimilation of soil water estimates and improved reservoir management	N.A

4.1 HTESSEL-CaMa

4.1.1 Model description

The land surface model (LSM) HTESSEL (Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land) computes the land surface response to atmospheric conditions, and estimates the surface water and energy fluxes and the temporal evolution of soil temperature, soil moisture content, vegetation interception and snowpack conditions. At the interface to the atmosphere each grid box is divided into fractions (tiles), with up to six fractions over land (bare ground, low vegetation, high vegetation, intercepted water, shaded snow and exposed snow). Vegetation types and cover fractions are derived from an external climate database, based on the Global Land Cover Characteristic (Loveland et al., 2000). The grid box surface fluxes are calculated separately for each tile, leading to a separate solution of the surface energy balance equation and skin temperature.

The interception reservoir is a thin layer on top of the soil/vegetation, collecting liquid water by the interception of rain and the collection of dew, and evaporating at the potential rate. The maximum capacity of this reservoir is a function of the grid-box Leaf Area Index (LAI). A resistance parameterization is used to calculate the turbulent fluxes, and in particular for evaporation the aerodynamic resistance is added to the canopy resistance. The canopy resistance is a function of downward short-wave radiation, leaf area index, soil moisture, water vapour deficit and a minimum stomatal resistance. For the soil moisture dependence, the fraction of roots in each soil layer is also considered.

In respect to WRR1 model version, two main changes have been introduced:

- Replacement of the bulk snow scheme by a multi-layer model;
- Increased the number of vertical soil layers from 4 to 9.

In WRR1, the snow was represented as an additional layer on top of the upper soil layer, with independent prognostic, thermal and mass contents (Dutra et al. 2010). It was represented by a single layer with an evolution of snow temperature, snow mass, snow density, snow albedo, and a diagnostic formulation for the snow liquid water content. For WRR2, a multi-layer snow-scheme was developed and implemented following Dutra et al. (2012). The multi-layer scheme carries the prognostics of snow mass, density and liquid water content for each layer and a single snow albedo. The number of snow layers is flexible, dependent on the total snow depth with a maximum number of layers that is user defined. In the current simulations 5 layers were used as the maximum number of snow layers. In addition to the new snow scheme, several changes were also done to the snow albedo parameterization and surface energy balance solver.

Below the surface, in WRR1 the soil was discretized in four layers (0.07, 0.21, 0.72 and 1.89 m layer thickness) and this was changed to 9 layers in WRR2 (0.01, 0.02, 0.04, 0.08, 0.1, 0.25, 0.5, 1 and 1 m layer thickness). The total soil depth in WRR2 (3 m) is similar to that of WRR1 (2.89 m), while having a higher vertical resolution closer to the surface to enhance fast water/thermal processes. Soil heat transfer follows a Fourier law of diffusion, modified to consider soil water freezing/melting (Viterbo et al. 1999). The vertical movement of water in the unsaturated zone of the soil matrix

follows Richards's equation and Darcy's law. Hydraulic conductivity and diffusivity are derived using van Genuchten formulation while 6 soil textures are used globally. Water movement is limited in the case of partially frozen soil by reducing the hydraulic conductivity and diffusivity. The top boundary condition is precipitation minus evaporation minus runoff, and the bottom boundary condition assumes free drainage.

The increased vertical resolution in both the soil and snow required some changes in the surface energy balance. In WRR1 the coupling between the skin layer and the underlying snow or soil was fixed with a constant (temporal/spatial) parameter and in WRR2 this was changed to be a diagnostic based on the underlying layer thickness and thermal characteristics. In the snow case, an iterative method was also introduced to limit the skin temperature over exposed snow to the freezing point.

Water leaves the soil column in the bottom layer as free drainage, and this is denoted as sub-surface runoff. At the surface, a variable infiltration rate that accounts for the sub-grid variability related to orography is used to compute the surface runoff (Balsamo et al. 2009).

The surface and sub-surface runoff generated by HTESSEL are fed to the Catchment-based Macro-scale Floodplain model (CaMa-Flood, Yamazaki et al. (2011). CaMa-Flood simulates the hydrodynamics in continental-scale rivers. The entire river network of the world are discretized to hydrological units named unit-catchments for achieving efficient flow computation at the global scale (Yamazaki et al. 2009). The water level and flooded area are diagnosed from the water storage at each unit-catchment using the sub-grid topographic parameters of the river channel and floodplains. The river discharge and flow velocity are calculated with the local inertial equation along the river network map which prescribes the upstream-downstream relationship of unit-catchments. The time evolution of the water storage, the only prognostic variable, is solved by the water balance equation which considers inflow from the upstream cells, outflow to the downstream cell and input from runoff forcing at each unit-catchment.

The HTESSEL-CaMA simulations were carried out at the default $0.25^\circ \times 0.25^\circ$ resolution. Land cover (e.g. type of vegetation and cover), mean climatologies of leaf area index and snow-free surface albedo were prescribed from the data used currently in the operational weather forecasts at ECMWF. The simulations were performed only over land points and a 10 year spin-up was carried out: initial simulation from 1st Jan 1979 to 1st Jan 1989, and the land surface state in Jan 1989 was used to initialize the main simulation starting on the 1st January 1979. In both HTESSEL and CaMa-Flood the default set of parameters were used, with no calibration performed for these simulations.

4.1.2 Data assimilation

For the data assimilations experiments (wrr2da) a 2D optimal interpolation (OI) methodology (de Rosnay et al., 2014; Brasnett, 1999) was implemented for snow depth using in-situ snow depth observations from the Global Historical Climatology Network (GHCN). The 2D OI scheme expresses the observations weighting functions from vertical and horizontal structure functions. It accounts for the covariance matrices of background and observations errors which enable to optimally combine model background and observations. Since most of the GHCN station information does not contain the exact time of the day of the observations, the data assimilation was performed with a 24h window using as model background the model state at 00UTC.

The model background and the observations of a particular day are integrated into the 2D OI that produces an analysis of snow depth. The following step is to propagate the snow depth analysis increments into the model prognostics. Snow depth increments could be applied to either snow mass or snow density or both. In the current implementation, it is assumed that snow density does not change, and all the snow depth increments are applied to snow mass. In addition, the snow mass increments are distributed to the different snow layers weighted by the thickness, i.e. thicker snow layers will have higher increments than shallower layers. Finally, once the snow mass increments are computed, these are applied to the model prognostic vector at 00UTC and the next day is simulated. This sequential data assimilation of the snow depth allows a smooth temporal correction of the model state and consistent evolution of all the remaining surface conditions. However, the snow mass increments become a new water source/sink leading to a water budget imbalance if only precipitation, evaporation and runoff are considered.

4.1.3 Evaluation

To evaluate the impact of the different components in WRR2 (meteorological forcing and model changes) a set of simulations, described in Table 4-2, were performed with WRR1/WRR2 forcing and the model versions used in WRR1 and WRR2, respectively. The comparison between WRR1_M2 vs WRR1 and WRR2 vs WRR1_F2 will show the impact of the model changes while WRR1_F2 vs WRR1 and WRR vs WRR1_M2 will address the impact of the meteorological forcing.

The evaluation mainly focuses on the mean annual cycle of key components of the water cycle, including the comparison with observed river discharge from the Global Runoff Data Center (GRDC). The surface soil moisture is also evaluated against the ESA CCI soil moisture (Dorigo et al., 2015), and point snow depth with the GHCN in-situ data which was also used in the data assimilation simulation.

Table 4-2: HTESEL simulations details

Simulations	Details
WRR1	WRR1 simulations (black)
WRR1_M2	WRR1 forcing with WRR2 model (grey)
WRR1_F2	WRR1 model with WRR2 forcing (dark magenta)
WRR2	WRR2 simulation (magenta)
WRR2_DA	WRR2 with snow data assimilation (blue, wrr2da)

(a) WRR2

The mean annual cycle of precipitation, evaporation, runoff and discharge for the Amazon basin (Figure 4-1) shows a very small impact of the model changes from WRR1 to WRR2 when compared with the forcing changes. While precipitation is reasonably similar between WRR1 and WRR2, there is a reduction of evapotranspiration in WRR2 which is compensated with an increased runoff. Since these changes cannot be attributed to model changes (WRR1_F2 is almost identical to

WRR2), it must be due to changes in the meteorological forcing. The mean annual cycle of the different components of evapotranspiration (Figure 4-2) shows that the main differences between WRR1 and WRR2 occur in the canopy evaporation. This is further illustrated on the global maps of differences between WRR2-WRR1 for evapotranspiration and canopy evaporation (see Figure 4-3) indicating that most of the decrease of evaporation occurring in the tropics results from a reduction of canopy evaporation.

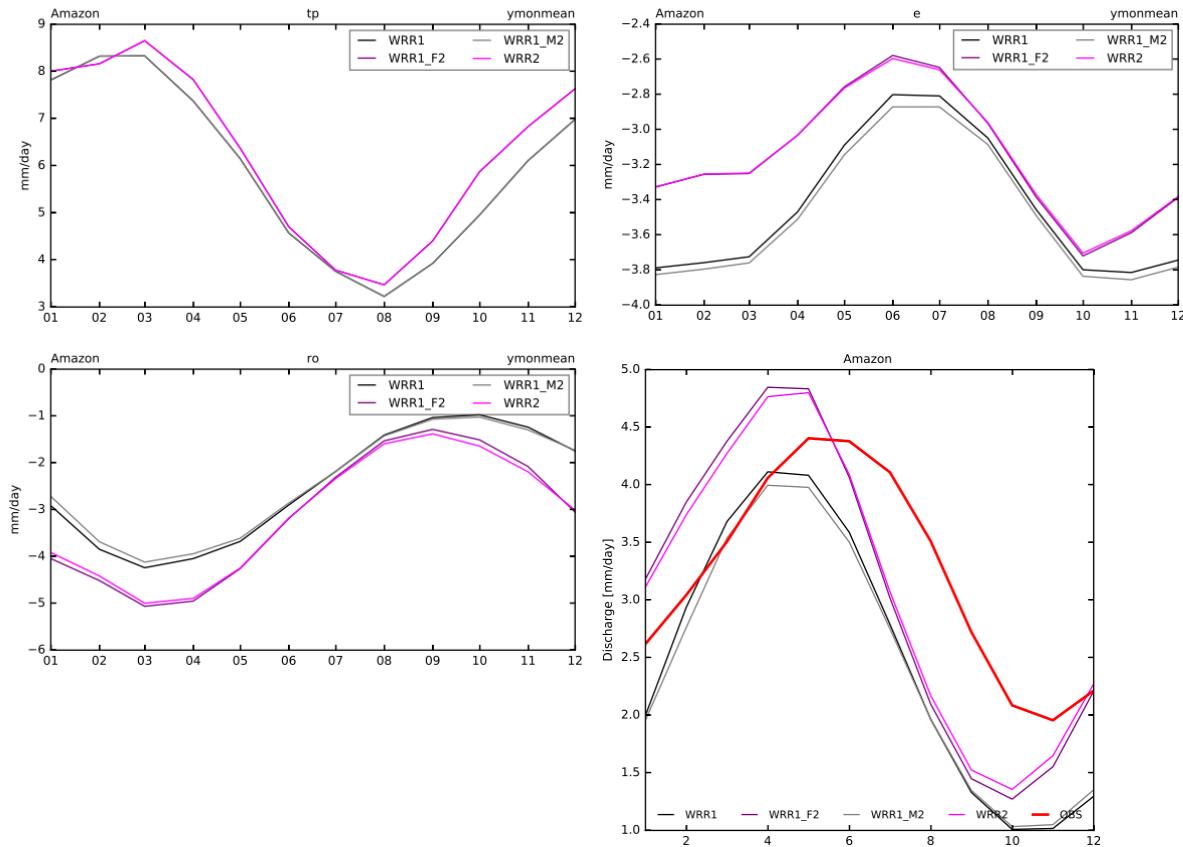


Figure 4-1: Mean annual cycle of precipitation (top left), evapotranspiration (top right), runoff (bottom left) and discharge (bottom right) over the Amazon basin in WRR1, WRR1_M2, WRR2_F2 and WRR2. On the discharge panel, the observations from GRDC are the red line.

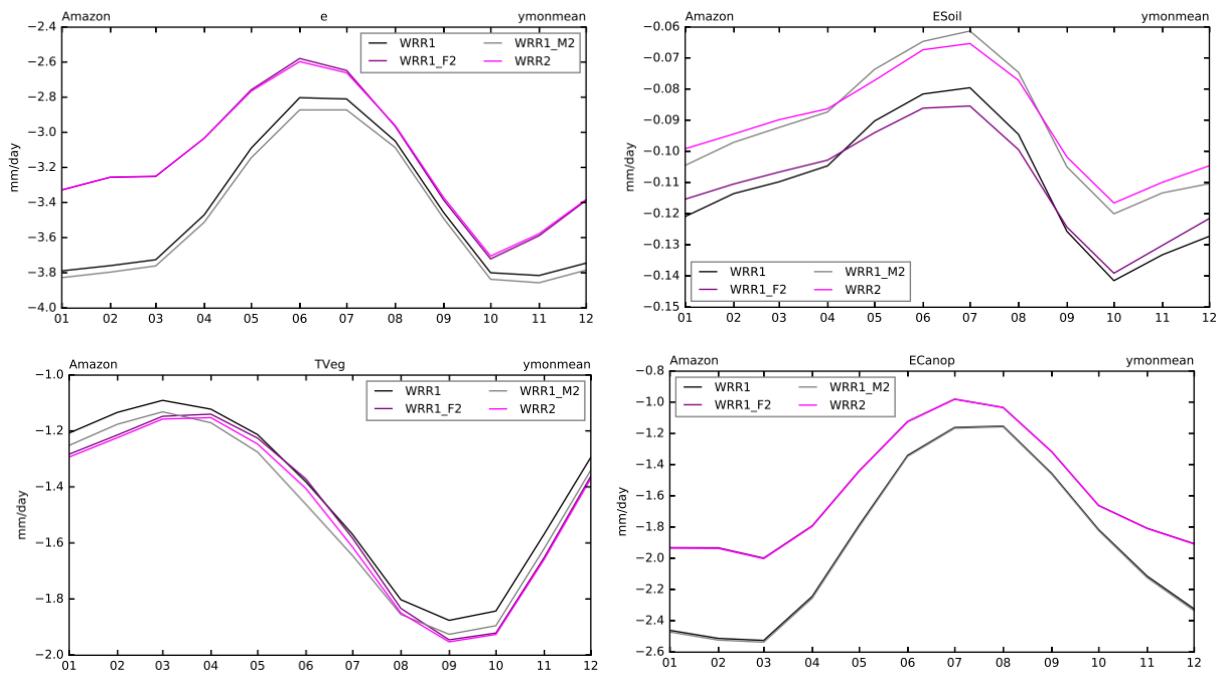


Figure 4-2: Mean annual cycle of evapotranspiration (top left), soil evaporation (top right), vegetation transpiration (bottom left) and canopy interception evaporation (bottom right) over the Amazon basin in WRR1, WRR1_M2, WRR2_F2 and WRR2.

Canopy evaporation is a very fast process that occurs after rainfall events evaporating the intercepted water over the canopy. This is an important process in the tropics due to the large interception reservoir size (large LAI) and energy availability. There was no change in the interception reservoir capacity from WRR1 to WRR2, and the changes in the available energy were small (not shown). One of the differences in WRR2 was the precipitation data that came from MSWEP that took advantage of satellite data. Besides the changes on total precipitation, the sub-daily distribution was also corrected with a much clearer diurnal cycle of precipitation over tropics compared with WRR1 (not shown). This results in less water being intercepted by the canopy in WRR2 and partially explaining the reduction of evaporation, and increase in runoff. These results highlight, that in the case of HTESSEL, the daily precipitation distribution impacts the water balance components. Furthermore, these results bring some light on two contradictory results, one suggesting that HTESSEL had too much evaporation (Haddeland et al., 2011) and another suggesting that this is too little (Getirana et al., 2014). In fact the differences seen in these studies are consistent with WRR1 vs WRR2 results that had a different precipitation daily distribution.

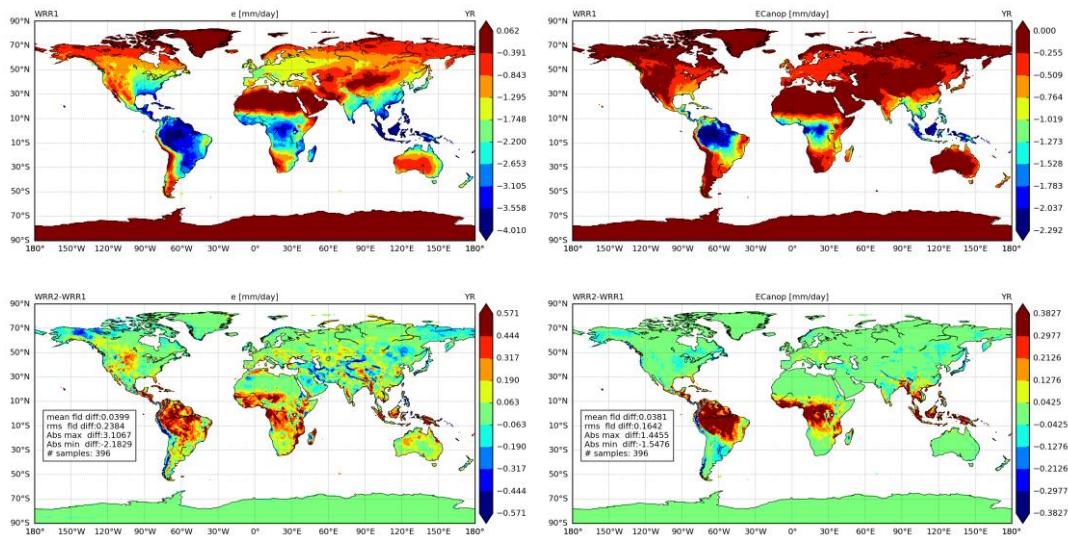


Figure 4-3: Mean evapotranspiration (top left) and canopy interception evaporation (top right) in WRR1 and differences between WRR2 and WRR1 (bottom panels)

Moving from the tropics to the snow dominated Ob basin (see Figure 4-4), the impact of the model changes from WRR1 to WRR2 are still limited compared to the impact of the meteorological changes. However, there is a slight increase in runoff as well as a delay of the peak discharge, due to the model changes, which moves modelled discharge closer to the observations. This can be primarily attributed to the enhanced snow thermal insulation with the multi-layer snow scheme that reduced soil freezing allowing for more water to infiltrate during the melting season.

The mean annual cycle of discharge in Yukon, Mackenzie, Danube and Mississippi (see Figure 4-5) shows a mixed impact of WRR2 v WRR1, with clear improvements in the northern basins (Yukon and Mackenzie), similar to the Ob basin. On the other hand, the two mid-latitude basins (Danube and Mississippi) show a deterioration with a reduction of discharge from WRR1 to WRR2, further away from the observations. This is primarily attributed to changes in the precipitation from WRR1 to WRR2. However, considering that runoff is much smaller than evaporation in these basins, errors in evapotranspiration cannot be discarded and a deterioration in discharge cannot be uniquely associated with the changes in the meteorological forcing.

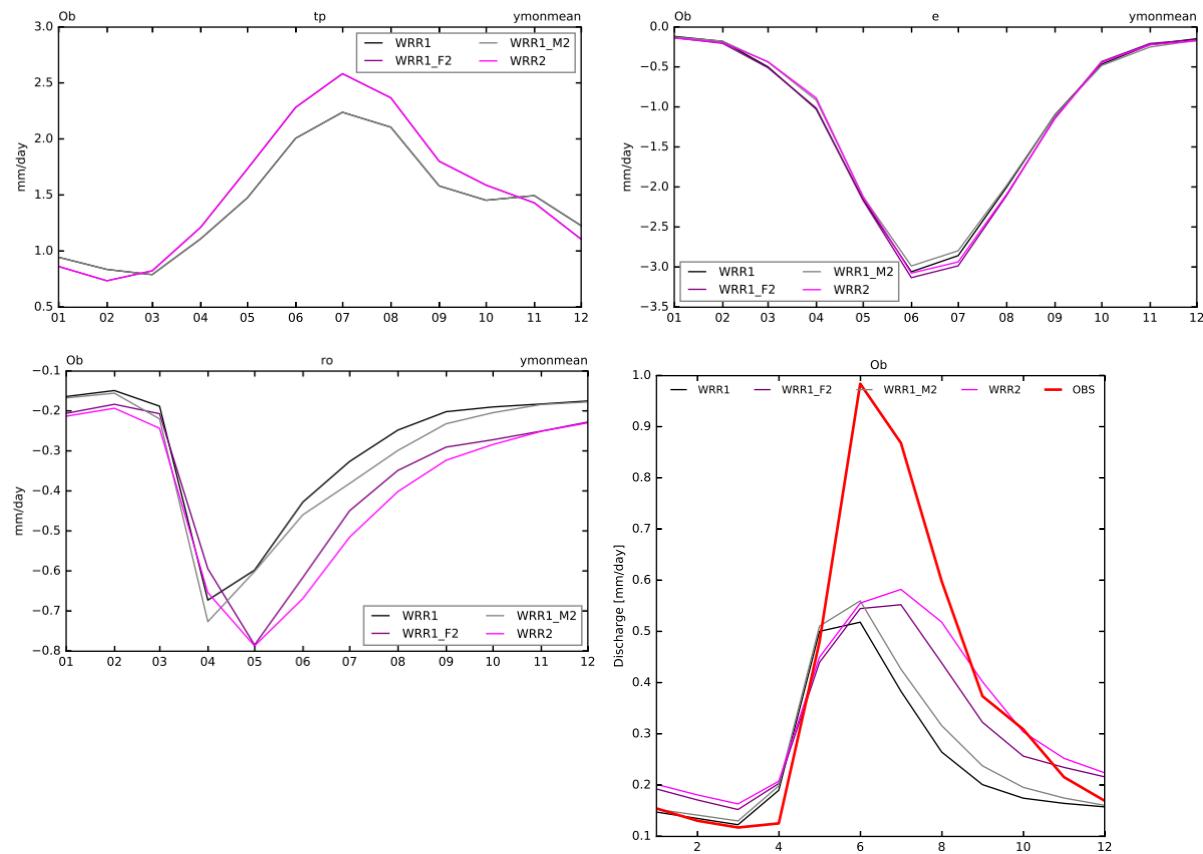
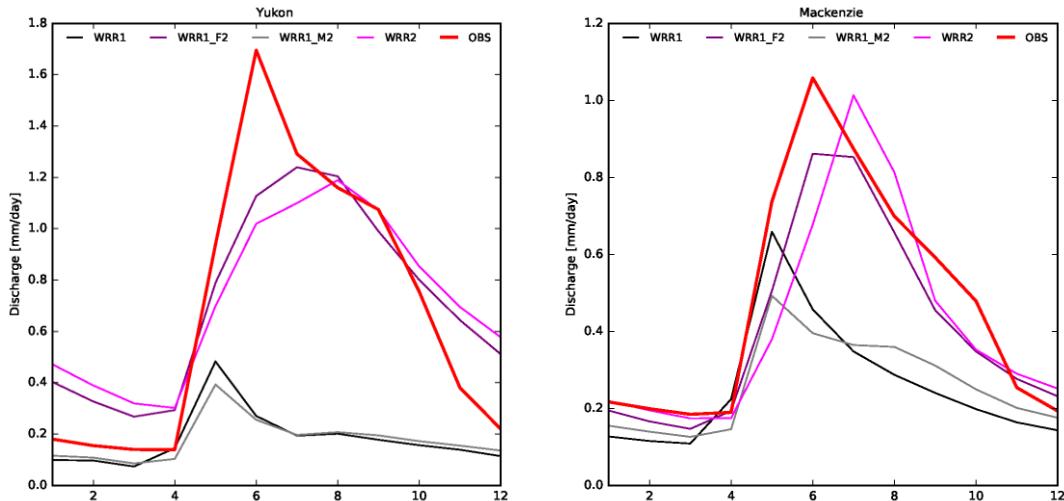


Figure 4-4: As Figure 4-1 but for the Ob basin.



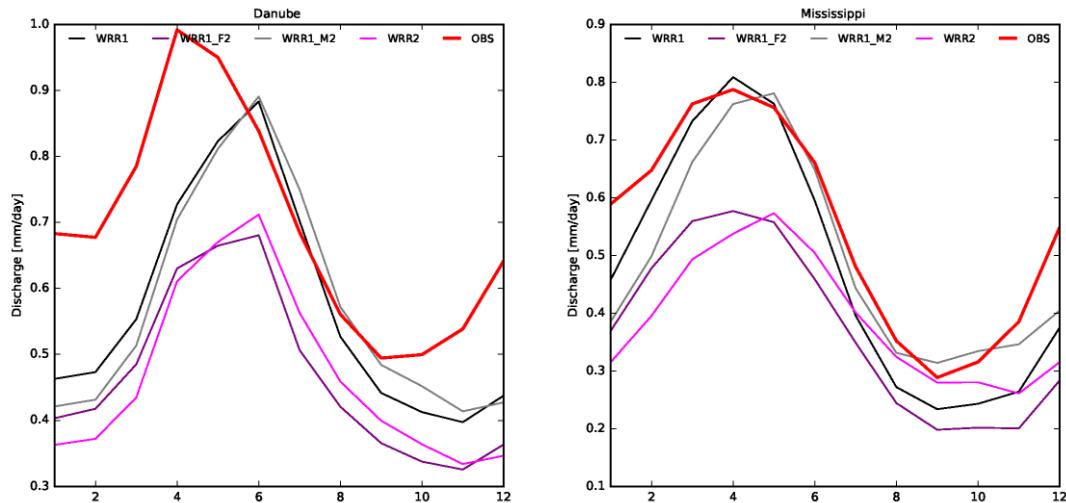


Figure 4-5: Mean annual discharge in Yukon (top left), Mackenzie (top right), Danube (bottom left) and Mississippi (bottom right). Simulations are compared against the GRDC observations (red) and the mean annual cycle computed only for the period with available observations.

Finally, we compared the surface soil moisture (top 7 cm) with the ESA CCI COMBINED product (version 2.2). For each grid-point the time series of surface soil moisture from the simulations were correlated with the ESA CCI COMBINED product for the 1988 to 2012 period. The correlations of the daily time series are shown in Figure 4-6 while the correlation of the anomalies, after removing a 35 days moving average, are shown in Figure 4-7. The full correlations show high values in semi-arid regions, and low values in deserts and northern latitudes. The anomaly correlations are generally lower, as expected, but the higher values also tend to cluster in semi-arid regions. The higher agreement over semi-arid regions is expected as in these regions the surface soil moisture is expected to have a clear response to precipitation variability and satellite retrievals have a high skill, while in other regions other factors dominate (e.g. northern areas: snow, soil freezing; tropics: missing ESA CCI data, and effects of vegetation). The impact of the model changes from WRR1 to WRR2 is small but tends to be positive, while the impact of the new forcing in WRR2 is clearly positive and dominated the marked improvement from WRR1 to WRR2 in terms of the agreement with the satellite data.

Overall, there is a clear difference between the model and forcing changes from WRR1 to WRR2, with the forcing changes bringing a much larger impact than the model, which is also apparent in other modelling system results (see for example SURFEX-TRIP results in section 4.6).

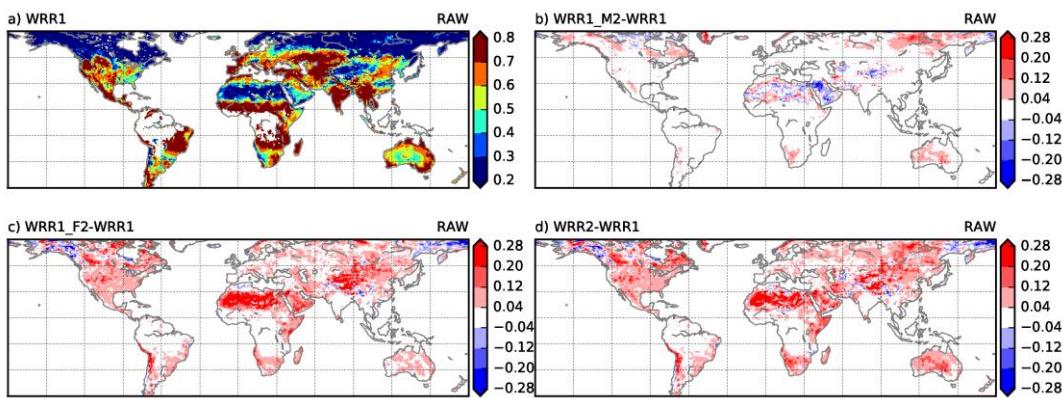


Figure 4-6: (a) temporal correlation of between daily surface soil moisture and EO ESA CCI soil moisture for the period 1988 to 2012 in WRR1. Correlation differences of different model/forcing configurations and WRR1: b) WRR1_M2 – WRR1, c) WRR1_F2 – WRR1 and d) WRR2-WRR1.

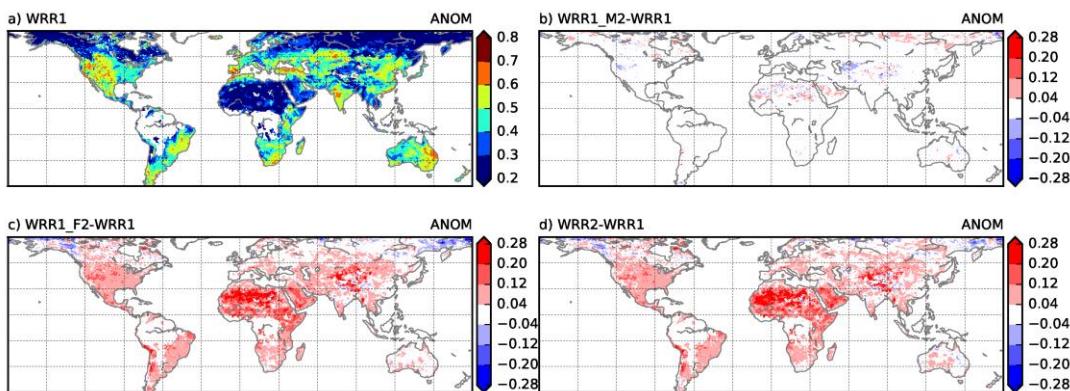


Figure 4-7: As Figure 4-6 but for the anomalies after removing a 35 days moving window in both the simulations and EO CCI.

(b) Data assimilation

The data assimilation simulations (wrr2da) were evaluated along with wrr1 and wrr2 by comparing the grid-point evolution of snow depth with the GHCN in-situ observations. The statistics are presented in terms of the mean annual cycle for the period 2000-2010 in three domains: Eurasia (Figure 4-8), Canada (Figure 4-9) and United States (Figure 4-10), including the mean bias and standard deviation of the errors. In all domains the model changes were positive during the accumulation to peak winter time snow depth, while there is some deterioration during the spring melt season in the Canada and United States domains. In the latter two regions, there is also a clear benefit of the new meteorological forcing, while over Eurasia the impact of the meteorological forcing is less evident, but still positive. The simulations with the data assimilation (wrr2da, blue lines in Figure 4-8 to Figure 4-10) show a significant improvement with almost zero mean bias and a clear reduction of the standard deviation of the errors. This comparison has a caveat since most of the data used for the evaluation was also used during the data assimilation. A clear evaluation should exclude some of the station data from the data assimilation to allow an independent verification. However, since the main goal of wrr2da was to produce the best estimate possible, all available stations were used by the data assimilation system.

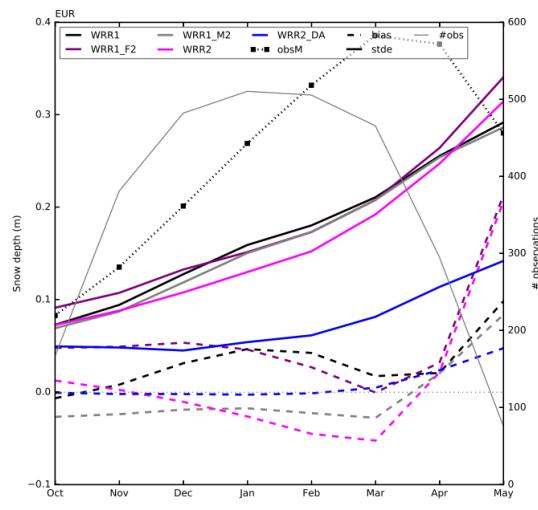


Figure 4-8: Mean annual cycle of snow depth verification over Eurasia using GHCN stations for the 2000-2010 period (grey line with the number of stations: right vertical axis; dotted symbols with the mean observations of snow depth (m)). The dashed lines represent the simulation bias and the solid lines the standard deviation of the error for: WRR1 (black), WRR1_M2 (grey), WRR1_F2 (dark magenta), WRR2 (magenta) and WRR2_DA (blue)

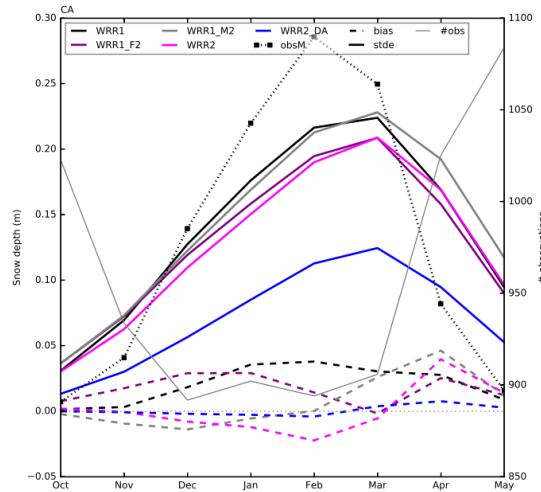


Figure 4-9: As Figure 4-8 but for Canada.

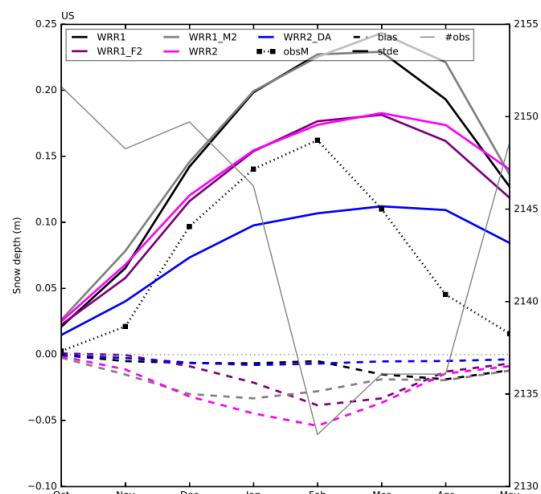


Figure 4-10: As Figure 4-8 but for United States.

Since wrr2da was sequential data assimilation, it is expected that the snow mass increments will impact other model components. The mean annual cycle of discharge in 6 selected large-scale basins, of which 4 snow-dominated (Yukon, Mackenzie, Ob and Yenisei) and 2 mid-latitude basins (Danube and Mississippi) illustrate the impact of the data assimilation on the simulated river discharge (see Figure 4-11). On the snow-dominated basins, the data assimilation had a negative effect while on the mid-latitude basins there was a slightly positive effect. The deterioration over the snow dominated regions is contra-intuitive sine wrr2da had a better snow depth evolution as shown in the previous results. To further understand these results, the snow mass increments from the data assimilation are shown in Figure 4-12. Over Yukon, Mackenzie, Ob and Yenisei, there is a clear negative snow mass increment during the melting seasons. This suggests that the model has a systematic melting bias, melting too late which is corrected by the data assimilation. However, the negative snow mass increments are removed from the surface, and are not available for infiltration and later runoff. This explains the deterioration of wrr2da simulation in terms of river discharge. These results indicate the 2D OI scheme used is very effective in correcting the snow depth evolution but it is not capable to trigger the model snow melt and, as a consequence, the negative increments are removed from the surface. Further developments are required in this area to better use the observations to constrain the model evolution so that the data assimilation can improve not only the snow depth evolution but also the melting.

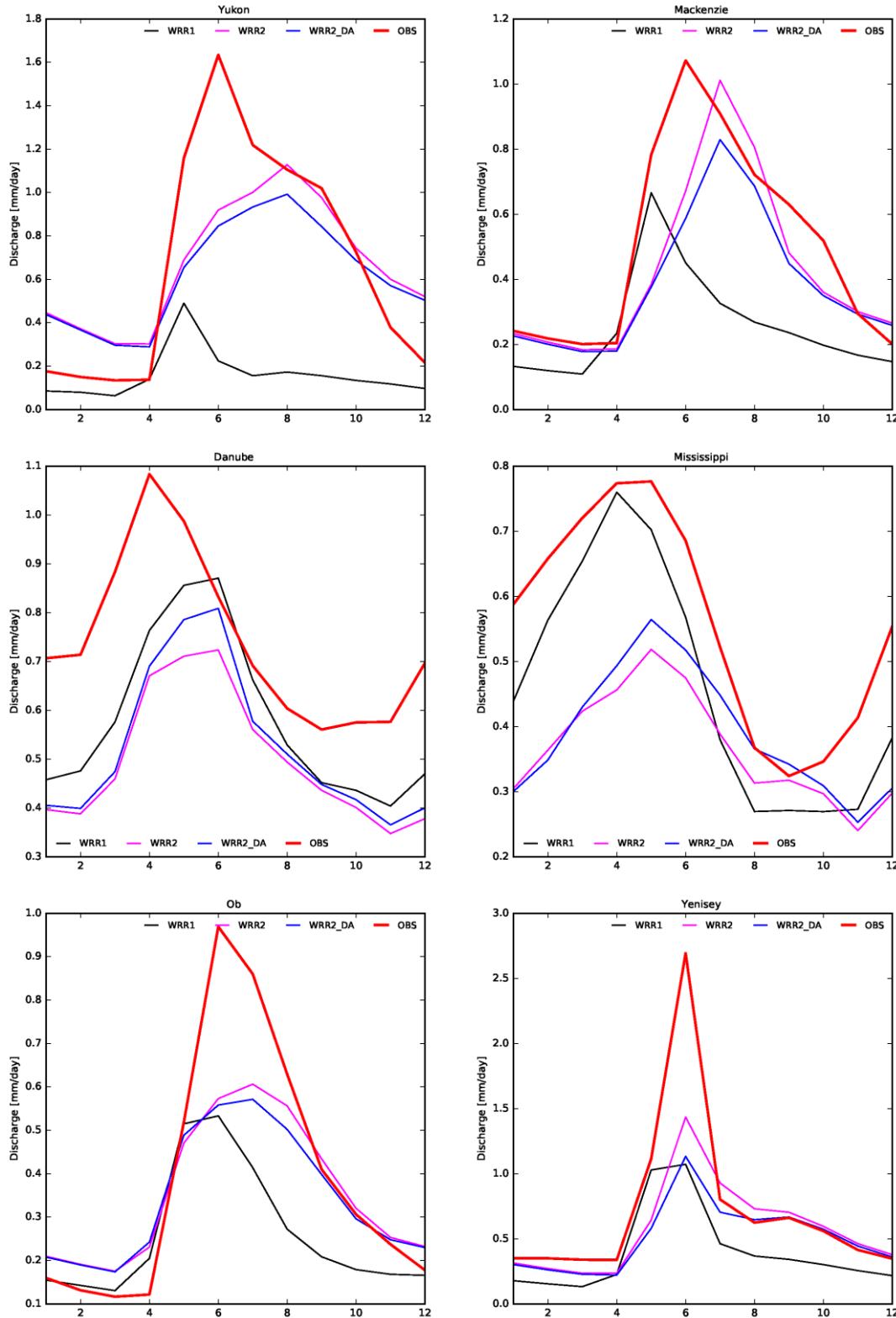


Figure 4-11: Mean annual discharge in Yukon (top left), Mackenzie (top right), Danube (centre left), Mississippi (centre right), Ob (bottom left), and Yenisey (bottom right). Simulations are compared against the GRDC observations (red) and the mean annual cycle computed only for the period with available observations between 2000-2012 (available WRR2_DA).

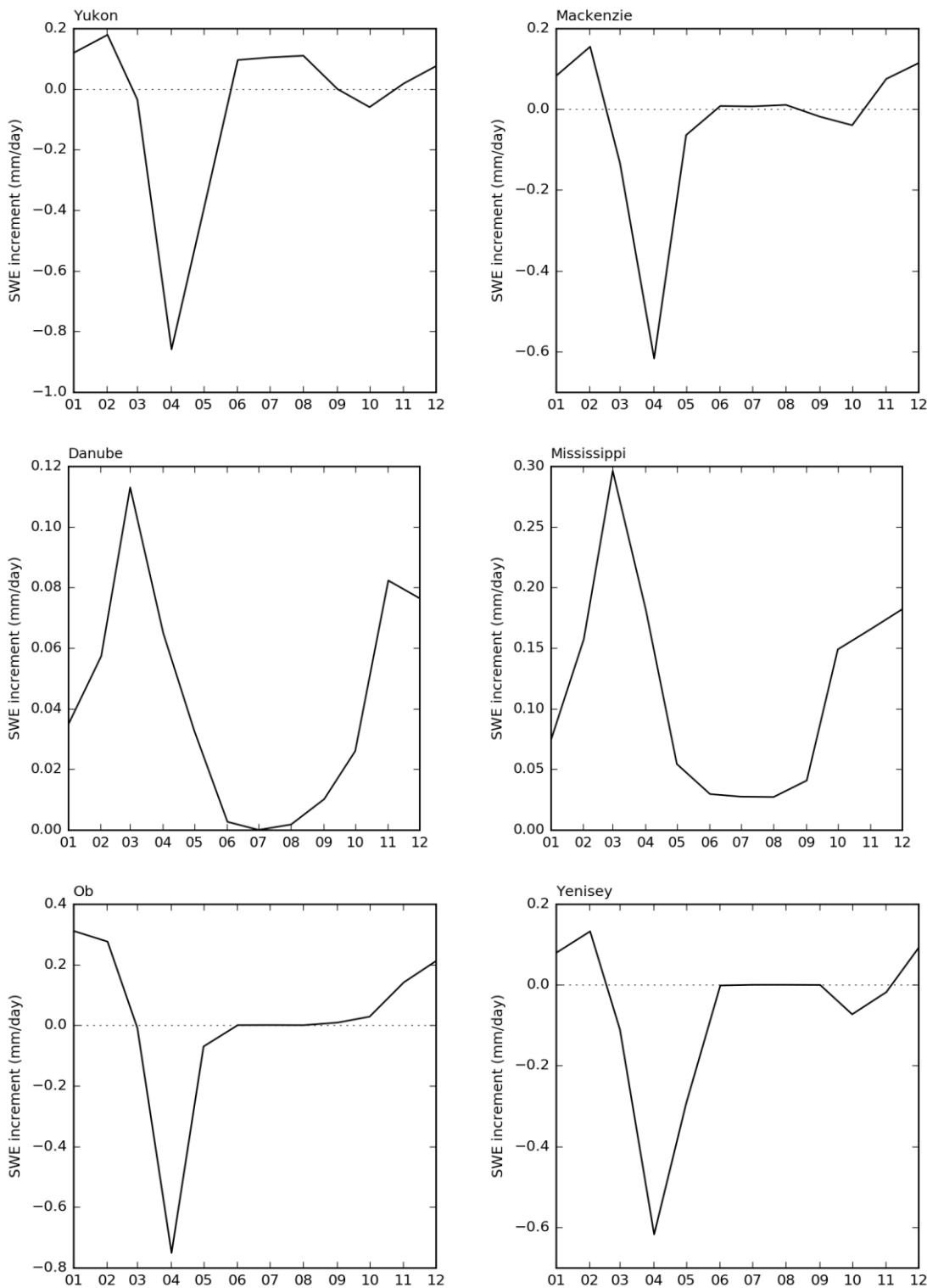


Figure 4-12: Mean annual snow mass increments in the WRR2_DA data assimilation simulation in Yukon (top left), Mackenzie (top right), Danube (centre left), Mississippi (centre right), Ob (bottom left), and Yenisei (bottom right).

4.2 Jules

4.2.1 Model description

General JULES model description is available on the D5.1 report WRR1. The main model development adapted for WRR2 was the inclusion of a terrain slope dependency in the saturation excess runoff scheme (PDM, Moore, 1985). The WRR1 version used PDM, which calculates a saturation fraction for each model grid cell as:

$$F_{sat} = 1 - \left(1 - \frac{S - S_0}{S_{max} - S_0}\right)^{\frac{b}{b+1}}$$

where S is the grid cell soil water storage, S_0 the minimum storage below which there is no surface saturation, S_{max} is the maximum grid cell storage and b is a shape parameter. Any saturation excess over the saturated area then generates surface runoff (Clark and Gedney, 2008). Plots in Figure 4-13 show how the b parameter shape up the curve for saturation fraction with soil water content in the x-axis, and how the S_0 parameter constrains the saturation fraction and the b variability. The sub-surface runoff is obtained as free drainage at the bottom of the soil column (at 3 m depth).

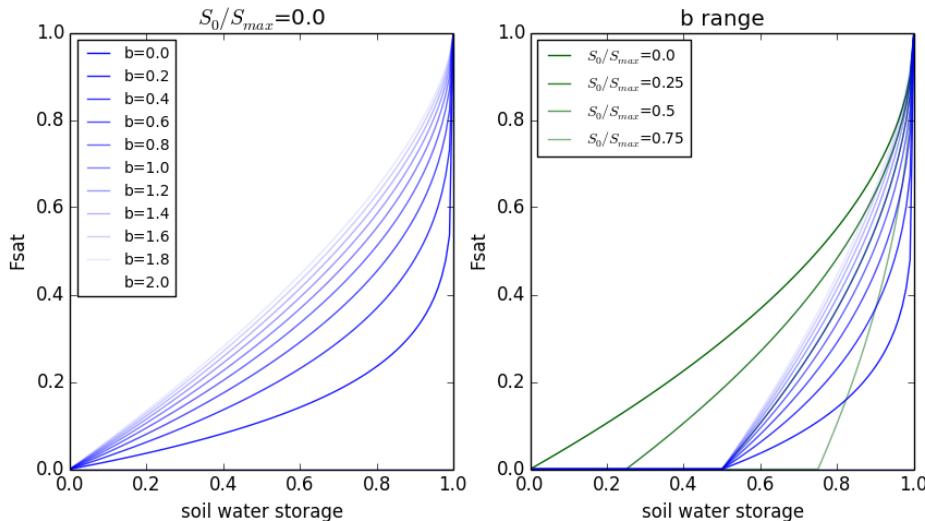


Figure 4-13: Saturation fraction calculated by the PDM scheme in JULES as a function of soil water storage to generate saturation excess runoff. Left: Variability obtained by the b parameter, for a fixed $S_0/S_{max}=0$. Right: Variability of the S_0 parameter (green curves) and variability of b for a fixed $S_0/S_{max}=0.5$.

JULES in the WRR1 version had the S_0 parameter hardwired to be zero. This together with a fixed value of b ($b=0.35$ in WRR1) give no room for sub-grid variability in the runoff generation of the model. After a deep study of the b and S_0 variabilities for a set of catchments in Great Britain using the CHESS 1 km meteorological driving dataset (Robinson et al., 2017), we adopted a new approach for WRR2, using a higher value of $b=2$ and a variable value for S_0/S_{sat} , that depends on the terrain slope as

$$S_0/S_{max} = 1 - \text{slope} * (1/\text{slope}_{max})$$

where $slope$ is the mean terrain slope for the given grid-cell and $slope_{max}$ is the maximum slope to apply the equation, taking $S_0/S_{max}=0$ for higher slopes. This way flatter regions will obtain higher values of S_0 and therefore constraint the saturation excess runoff production as compared to WRR1, while regions of high slope will produce higher runoff than WRR1 due to the high value of b . Figure 4-14 shows the improvement in Nush-Sutcliffe efficiency due to a better representation of the base-flow using the new WRR2 approach for the flat Thames catchment in the UK (using the aforementioned 1km driving data).

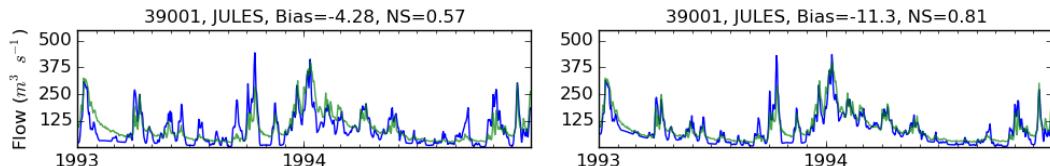


Figure 4-14: River flow at Kingston (UK), outlet of the Thames catchment simulated. Green lines are flow observation at the gauge station. Blue lines are river flow simulated by JULES with the best possible parameterization before introducing the S_0 linear dependency on terrain slope (left) and with the new parameters adopted for WRR2 (right). Total Bias and Nash-Sutcliffe efficiency metrics are given on top of the plots.

Finally, WRR2 also has a climatology for LAI, whereas WRR1 used a fixed annual LAI for each PFT (plant functional type).

4.2.2 Evaluation

To evaluate the impact of model development we have carried out an analysis comparing WRR1_F1 and WRR2_F2 to a series of hydrological benchmark datasets using the ILAMB system (<https://bitbucket.org/ncollier/ilamb>, Mu et al, 2016). The system has already been used in the project to evaluate the WRR1 runs for all models (Schellekens et al., 2016).

To focus on the runoff production, which is the more direct impact of WRR2 developments, we use the GCRF (Global Composite Runoff Fields) product (Fekete et al., 2002), particularly the OBS field that uses a 30 min river network to extrapolate gauging stations river flows from the GRDC (Global Runoff Data Center) database to produce a gridded 0.5x0.5 resolution global climatology for runoff. To use this benchmark dataset, we have created the runoff climatologies from WRR1_F1 and WRR2_F1 runs.

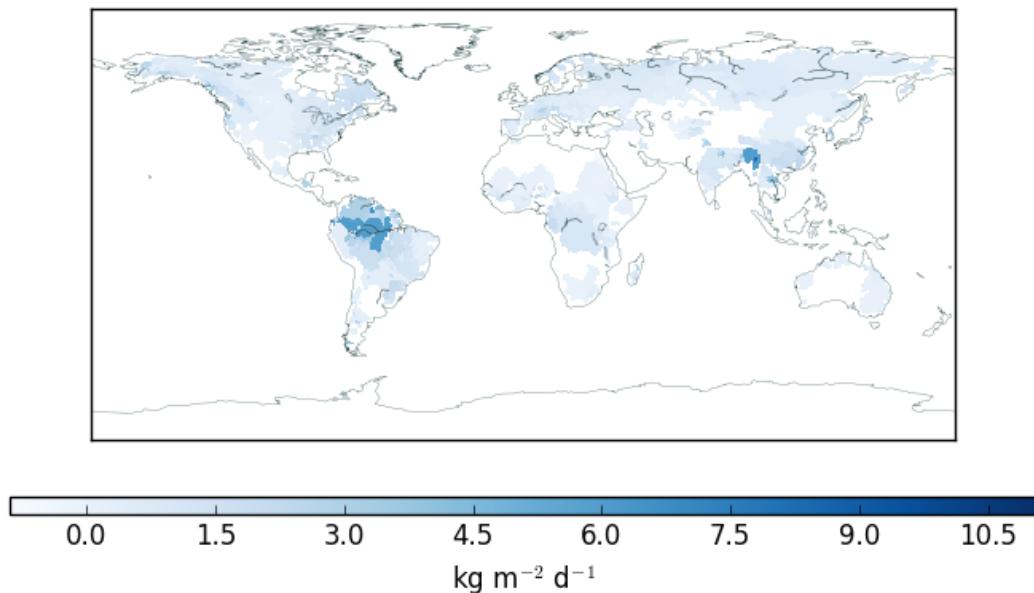


Figure 4-15: Mean Global Runoff (kg m⁻² d⁻¹) from the GCRF-OBS product (Fekete et al., 2002). White areas are not represented by the dataset (based on catchments).

The resultant scores applying the ILAMB benchmarking system to compare with GCRF-OBS are summarized in Table 4-3. The total global mean runoff has been reduced with the new WRR2 developments, and the overall score has improved slightly, from 0.64 to 0.656. The more significant changes in score from WRR1 to WRR2 are the improvements in seasonal score and RMSE score. A better represented seasonal cycle is therefore the first conclusion for the model development detailed on the previous section. Figure 4-16 shows the shift in days between the maximum values in the WRR simulations and the benchmark dataset, and it is clear that the shift has been significantly improved in west Europe and Asia, where the maximum values come too early in WRR1.

Table 4-3: Summary table obtained applying the ILAMB system to the WRR1_F1 and WRR2-F2 climatology runoff fields. The overall score is a composition of the rest of 0-1 scores: Bias Score, RMSE Score, Seasonal Cycle Score and Spatial Distribution Score.

Runoff	Period Mean [kg m ⁻² d ⁻¹]	Bias [kg m ⁻² d ⁻¹]	RMSE [kg m ⁻² d ⁻¹]	Phase Shift [d]	Bias Score [1]	RMSE Score [1]	Seasonal Cycle Score [1]	Spatial Distribution Score [-]	Overall Score [1]
GCRF-OBS	0.865								
JULES-WRR1_F1	0.949	0.083	0.935	-23.394	0.8	0.387	0.738	0.89	0.64
JULES-WRR2-F1	0.921	0.056	0.913	-14.654	0.802	0.413	0.766	0.887	0.656

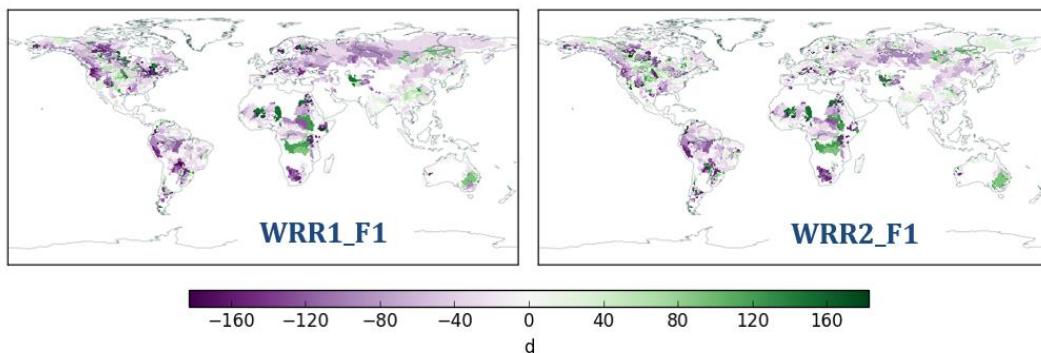


Figure 4-16: Global shift (days) between the maximum values in the WRR simulations (WRR1_F1, right; WRR2_F1, left) and the GCRF-OBS runoff benchmarking dataset.

The improvement in skill by WRR2 can be closely assessed using regional analysis with the ILAMB system. A series of regions were included in this study and in Figure 4-17 we show the RMSE score maps for Europe (WRR1 and WRR2). The red regions in west Europe have been significantly improved, and again this improvement can be seen in the seasonal cycle represented by both simulations as a space integration over the Europe region (together with the GCRF-OBS product) in Figure 4-18. It is from metrics in flatter regions like west Europe that WRR2 has reached an overall global skill improvement.

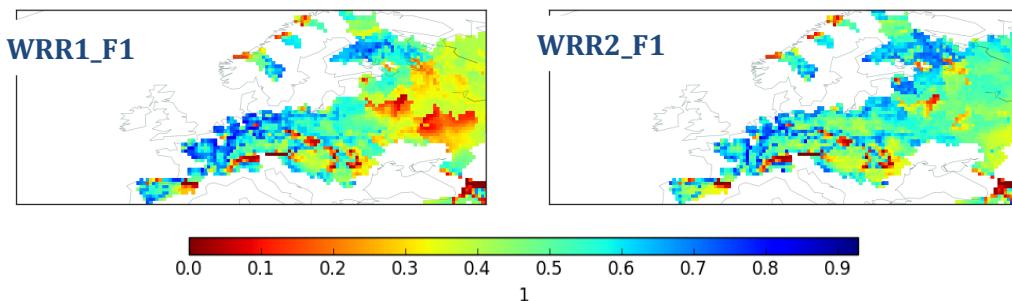


Figure 4-17: RMSE Score (0-1) over the Europe region obtained by the ILAMB evaluation against the GCRF-OBS runoff benchmarking dataset (WRR1_F1, right; WRR2_F1, left).

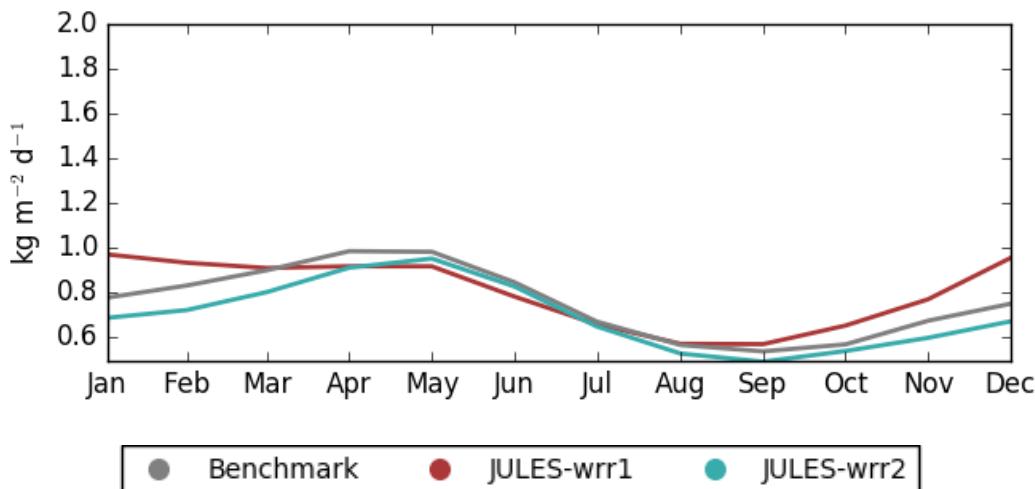


Figure 4-18: Monthly climatology for runoff ($\text{kg m}^{-2} \text{d}^{-1}$) averaged over the Europe region (Figure 4-18) by the GCRF-OBS product (grey), WRR1_F1 (red) and WRR2_F1 (cyan).

Other benchmarking datasets for water resources has been included in our ILAMB evaluation, particularly the ESA-CCI surface soil moisture product (Dorigo et al., 2015) and the FLUXNET-MTE latent heat product derived from empirical up-scaling eddy covariance measurements (Jung et al., 2009). The improvement in skill of runoff production does not affect negatively other components of the water cycle as we can see from Table 4-4 and Table 4-5 summarizing the global ILAMB evaluation of surface soil moisture and latent heat. Particularly for soil moisture, WRR2 improves the skill from a better characterization of the spatial distribution, as now high and low altitude regions do behave differently in the runoff production and therefore the water availability for the top soil layers.

From previous analysis (Deliverable D4.2) we knew that the JULES evaporation simulation was high as compared to other sources (GLEAM satellite based evapotranspiration product). Hence it is not surprising that an increase in the global latent heat by WRR2 (1% from 44.0 W m^{-2} with WRR1 to 44.44 W m^{-2}) provokes lower skill metrics in Table 4-5. The overall decrease, however, is not as high as the increases in soil moisture and runoff.

Table 4-4: Summary table obtained applying the ILAMB system to the WRR1_F1 and WRR2_F2 surface soil moisture (top 10 cm) monthly fields. The overall score is a composition of the rest of 0-1 scores: Bias Score, RMSE Score, Seasonal Cycle Score, Spatial Distribution Score and Interannual Variability Score.

Surface Soil Moisture	Period Mean [$\text{m}^3 \text{m}^{-3}$]	Bias [$\text{m}^3 \text{m}^{-3}$]	RMSE [$\text{m}^3 \text{m}^{-3}$]	Phase Shift [d]	Bias Score [1]	RMSE Score [1]	Seasonal Cycle Score [1]	Spatial Distribution Score [-]	Interannual Variability Score [1]	Overall Score [1]
ESA-CCI	0.196									
JULES-WRR1_F1	0.166	-0.03	0.09	-9.11	0.714	0.622	0.723	0.765	0.62	0.678

JULES-WRR2-F1	0.164	-0.03	0.088	-11.0	0.724	0.63	0.727	0.826	0.634	0.695
---------------	-------	-------	-------	-------	-------	------	-------	-------	-------	-------

Table 4-5: Summary table obtained applying the ILAMB system to the WRR1_F1 and WRR2-F2 latent heat monthly fields. The overall score is a composition of the rest of 0-1 scores: Bias Score, RMSE Score, Seasonal Cycle Score, Spatial Distribution Score and Interannual Variability Score.

Latent Heat	Period Mean [W m ⁻²]	Bias [W m ⁻²]	RMSE [W m ⁻²]	Phase Shift [d]	Bias Score [1]	RMSE Score [1]	Seasonal Cycle Score [1]	Spatial Distribution Score [-]	Interannual Variability Score [1]	Overall Score [1]
FLUXNET-MTE	39.23									
JULES-WRR1_F1	44.00	4.283	16.01	0.599	0.748	0.638	0.894	0.976	0.603	0.749
JULES-WRR2-F1	44.44	4.717	16.70	0.082	0.728	0.628	0.882	0.97	0.609	0.741

We have focused in this chapter in the evaluation of model improvements, rather than driving data improvements, as we will be evaluating that on future reports, namely D4.3 and D5.3.

4.3 LISFLOOD

Contributor: JRC

4.3.1 Model description

Between WRR1 and WRR2, the LISFLOOD model has been modified quite substantially. While still running at 0.1 degree, the following changes were implemented

- LISFLOOD has been fully recoded into Python-PCRaster, and is using netcdf as input and output file formats, next to the PCRaster map format;
- The number of computational soil layers has been increased from 2 to 3, with layer 1 representing the top 5cm of the soil, allowing for data assimilation of satellite soil moisture;
- Paddy rice and crop-irrigation water demand calculations are now embedded in LISFLOOD, including abstraction, conveyance losses, efficiencies, and net consumption and return flow. This has the advantage that irrigation demand will now change with a changed forcing, or a change in irrigated area. Or a change in efficiency;
- The simulation of groundwater has been improved by implementing groundwater abstraction and including a threshold value, underneath which no base-flow occurs to the river system. Thus, the aquifer might temporarily lose its connection to the river due to climate forcing or over-abstraction;

- Water allocation scheme is introduced whereby water is first made available for the public sector, followed by energy sector, manufacturing industry, livestock, crop irrigation. Also, an optional ‘environmental flow’ threshold of river discharge is implemented. If the actual local daily river discharge falls below the environmental flow threshold, any water abstraction is temporarily stopped until the minimum is exceeded again; this may lead to a ‘water shortage’ per sector, which is recorded and available as model output;
- LISFLOOD produces several additional water resources indicators – when set in the settings file -, such as the Water Exploitation Index (WEI) for demand and consumption, the Falkenmark indicator showing the amount of freshwater available per capita per year (both including and excluding upstream inflow), and the new Water Dependency Index, showing the dependency of a region on upstream freshwater inflow to fulfil its local water demand;
- The 0.1° global LISFLOOD setups now includes an increased number of lakes and reservoirs and their estimated parameterisation;
- Assimilation of discharge, soil moisture and snow cover is possible, but not executed yet.

For WRR2, LISFLOOD was not calibrated.

4.3.2 Evaluation

(a) Impact of the model improvements

To evaluate the impact of the model changes between WRR1 and WRR2 setup, a simulation has been carried using the model setups of WRR1 and WRR2 but the same forcing of WRR1 for both runs. Results of these simulations are compared to observed monthly and daily discharges at selected gauges (see Table 4-6 and Table 4-7).

It can be concluded that the updated LISFLOOD version at 0.1 degree for WRR2 simulates river discharge less well than the previous version. This may be partially related to the increased number of reservoirs and their uncertain parameterisation. But most likely, it is due to the fact that for WRR2 the same parameterisation was used as for the calibrated WRR1 version. Since LISFLOOD has been significantly changed with automated irrigation, enhanced abstraction and groundwater models, an updated calibration is likely needed to improve the results compared to WRR1.

Some basins show improvements from WRR1 to WRR2, such as the Oder river or the Amazon. Some are unchanged, such as Danube or Thames. But for several rivers such as Ebro, Elbe, Rhine, Zambezi and others, the calibrated WRR1 LISFLOOD version outperforms the uncalibrated WRR2 LISFLOOD version.

Table 4-6: Kling-Gupta Efficiency (KGE) for monthly discharge, for the Rhine and Amazon basin, showing the effects of the LISFLOOD model changes.

RHINE			
1990-2009		WRR1-WFDEI-100	WRR2_WFDEI-100
KGE		0.559	0.438
AMAZON			
1980-2008		WRR1-WFDEI-100	WRR2_WFDEI-100
KGE		0.902	0.894

Table 4-7: Kling-Gupta Efficiency (KGE) for daily discharge for selected gauging stations, showing the effects of the LISFLOOD model changes, for the period 1980-2008.

RiverName	StationName	TIER1_WFDEI_100	TIER2_WFDEI_100
LENA	KYUSYUR (KUSUR)	0.664	0.112
YENISEY	IGARKA	0.662	0.288
MACKENZIE	ARCTIC RED RIVER	0.829	0.759
OB	SALEKHARD	0.254	0.299
Rhone	Le_Rhone_a_Beaucaire	0.780	0.700
Ebro	EbroEnTortosa	0.756	0.271
Po	Lagoscuro	0.664	0.471
Elbe	NeuDarchau	0.634	0.416
Oder	Hohensaaten	0.568	0.636
RHINE	REES	0.834	0.690
VOLGA	VOLGOGRAD POWER PLANT	0.758	0.000
DANUBE	CEATAL IZMAIL	0.697	0.680
THAMES	KINGSTON_ON_THAMES	0.845	0.842
Sava	S.Mitrovica	0.812	0.710
Tisza	Senta / Szeged	0.871	0.809
MISSISSIPPI	VICKSBURG, MS	0.398	0.361
AMAZONAS	OBIDOS - LINIGRAFO	0.802	0.846
CONGO	KINSHASA	0.779	0.721
ZAMBEZI	KATIMA MULILO (64370001)	0.146	-0.035
ORANGE	VIOOLSDRIF (27811003)	-0.006	-0.555
MURRAY	BELOW WAKOOL JUNCTION	-6.002	-5.564
TOTAL KGE		6.744	3.456

(b) Impact of the new forcing

To evaluate the impact of the new forcing used in the WRR2 setup, simulations have been carried out using the WRR2 model setup with various forcings from WRR2 as well as the WRR1 WFDEI forcing. Results of these simulations are compared to observed monthly and daily discharges at selected gauges.

The impacts of different forcing datasets on LISFLOOD model output is quite significant. Differences vary from basin to basin, and some forcing data produce better results in one basin, but not necessarily in other basins.

In general, it can be concluded that the MSWEP precipitation forcing outperforms the WFDEI forcing. Some individual basins show the opposite results, the majority of investigated basins performs significantly better with MSWEP. Especially for the Amazon basin, extremely good results are obtained using MSWEP (see Table 4-8 and Figure 4-19, Figure 4-20). In the Rhine (Figure 4-19), the results of the 0.1 degree LISFLOOD setup have been compared with the calibrated 5km setup forced with the JRC gridded observed meteorological dataset. The latter outperforms the 0.1 degree run with MSWEP, but this may well be related to resolution and calibration as well, and not only to forcing.

In addition, several simulation experiments were done with various evapotranspiration settings. As in Tier1 it was found that LISFLOOD potential evapotranspiration estimates were somewhat below other models, we experimented with including a 'gain' factor of 1.118, so an 11.8 % bias correction. This increase does bring the overall simulated discharges closer to the observed ones. The summed up KGE for the stations in Table 4-9 is higher (5.3) than before (4.7). But there are regional differences, and in various river basins this correction shows opposite effects. Using the evapotranspiration dataset provided by the VU University did yield a slightly less performing model with respect to daily river discharge, with a summed KGE of 3.6 compared to 4.7.

Table 4-8: Kling-Gupta Efficiency (KGE) for daily discharge, for the Rhine and Amazon basin, showing the effects of the TIER2 forcing changes. For the Rhine, the comparison is also shown with the 5km resolution LISFLOOD setup, forced with the JRC meteo data.

RHINE						
1990-2009	WRR2_WFDEI-100	WRR2_MSWEP-100	WRR2-MSWEP-PET	Europe5kObs	Europe5kObsNw	
KGE	0.438	0.701	0.682	0.774	0.776	
AMAZON						
1980-2008	WRR2_WFDEI-100	WRR2_MSWEP-100	WRR2-MSWEP-PET			
KGE	0.894	0.976	0.866			

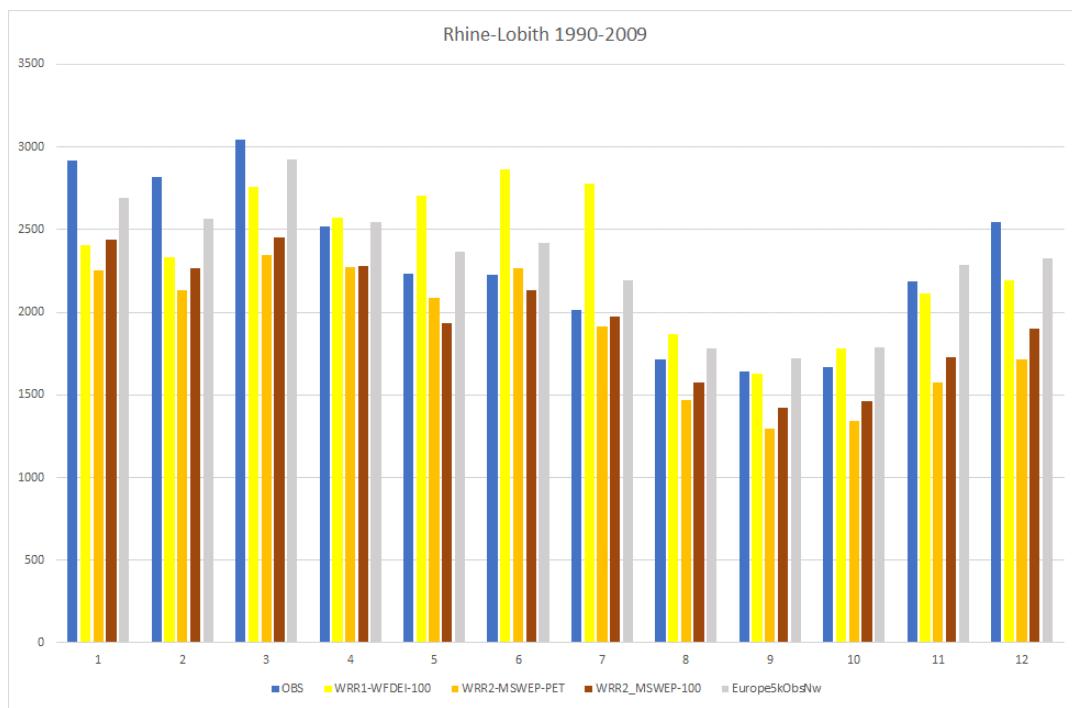


Figure 4-19: Observed and LISFLOOD simulated monthly discharge (m³/s) for the Rhine basin.

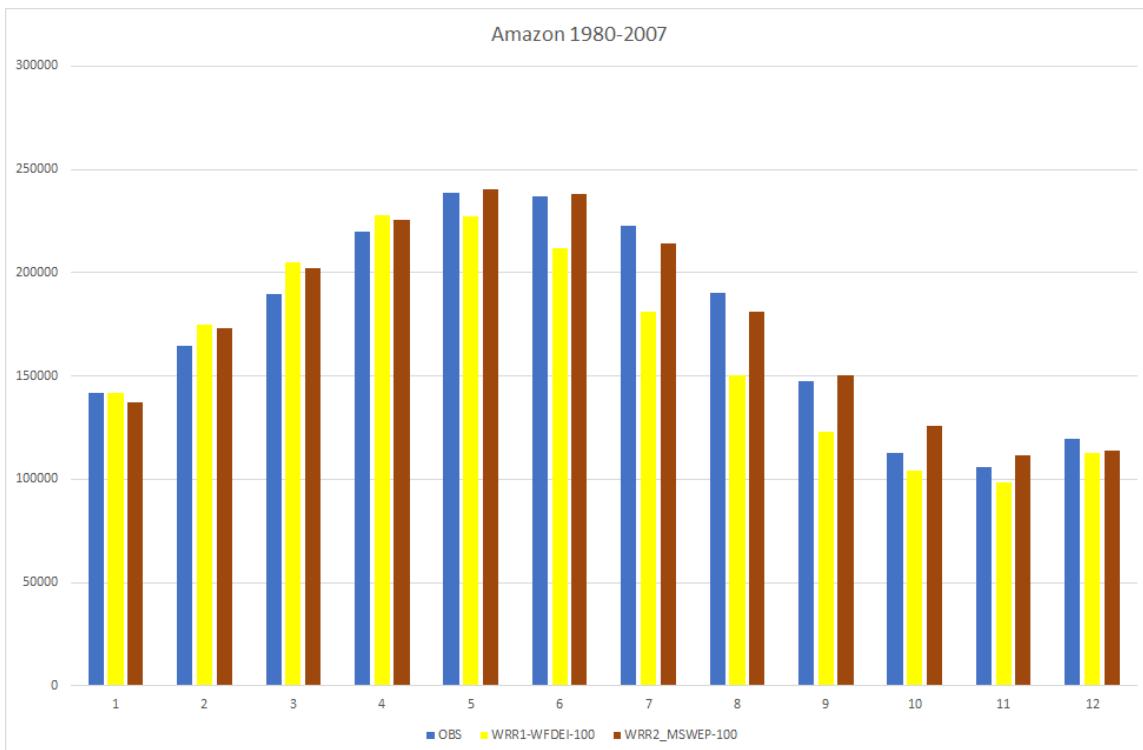


Figure 4-20 Observed and LISFLOOD simulated monthly discharge (m³/s) for the Amazon river.

Table 4-9: Kling-Gupta Efficiency (KGE) for daily discharge for selected gauging stations, showing the effects of the TIER2 forcing changes, for the period 1980-2008.

RiverName	StationName	TIER2_WFDEI_100	TIER2_MSWEPE_100	TIER2_MSWEPE_EVAP	TIER2_MSWEPE_118
LENA	KYUSYUR (KUSUR)	0.112	0.060	0.086	0.063
YENISEY	IGARKA	0.288	0.178	0.199	0.193
MACKENZIE	ARCTIC RED RIVER	0.759	0.707	0.678	0.751
OB	SALEKHARD	0.299	0.345	0.388	0.374
Rhone	Le_Rhone_a_Beaucaire	0.700	0.650	0.615	0.603
Ebro	EbroEnTortosa	0.271	-0.076	-0.113	0.116
Po	Lagoscuro	0.471	0.469	0.431	0.449
Elbe	NeuDarchau	0.416	0.486	0.462	0.468
Oder	Hohensaaten	0.636	0.491	0.438	0.358
RHINE	REES	0.690	0.715	0.673	0.639
VOLGA	VOLGOGRAD POWER PLANT	0.000	-0.239	-0.466	-0.467
DANUBE	CEATAL IZMAIL	0.680	0.753	0.800	0.786
THAMES	KINGSTON_ON_THAMES	0.842	0.665	0.652	0.642
Sava	S.Mitrovica	0.710	0.733	0.746	0.739
Tisza	Senta / Szeged	0.809	0.640	0.576	0.537
MISSISSIPPI	VICKSBURG MS	0.361	0.503	0.503	0.563
AMAZONAS	OBIDOS - LINIGRAFO	0.846	0.956	0.851	0.944
CONGO	KINSHASA	0.721	0.519	0.449	0.457
ZAMBEZI	KATIMA MULILO	-0.035	-0.412	-0.874	-0.088
ORANGE	VIOOLSDRIF	-0.555	-0.087	-0.263	-0.162
MURRAY	BELOW WAKOOL JUNCTION	-5.564	-3.290	-3.206	-2.646
TOTAL KGE		3.456	4.765	3.623	5.317

Summary water indicator outputs of the LISFLOOD model, such as the evaporative deficit, the amount of freshwater available per capita do change as well based on the different forcings, but the overall patterns and key messages stay the same (see Figure 4-21 to Figure 4-28).

The multiannual average evapotranspiration deficit (Figure 4-21 and Figure 4-22) seems not to be different from WFDEI to MSWEP forcings, but there are differences for individual months.

The freshwater availability per capita maps (Figure 4-23, Figure 4-24) are different in several water scarce areas such as the Middle East, northern Africa, but the main areas stay the same. 1700 m³ per capita per year is typically used as a critical threshold to identify water scarce regions.

The Water Exploitation Index 'plus' (WEI+) is defined as net water consumption as a fraction of available freshwater, including upstream inflow. Net consumption is including all sectors: irrigated agriculture, livestock, manufacturing industry, cooling for energy production, public/domestic water consumption. Return flows are accounted for. This is a core indicator used in European water policy. Values above 0.2 are critical, and above 0.4 indicate severely water stressed areas. At the global scale, and for the average WEI+, only marginal differences occur when using WRR1 or WRR2 forcing, so the key message stays the same (Figure 4-25, Figure 4-26).

The Water Dependency Index is defined as the fraction of water demand that cannot be met by local freshwater supply and is not met by desalination. This is typically the demand part for which freshwater from upstream inflowing rivers is used. The Nile in Egypt is a typical example here, where Egypt depends for a large part on this upstream,

cross border water. Changes in WDI may be an indicator of risk of water conflicts and the need for basin wide integrated water policies. Also for the WDI, forcing changes from WRR1 to WRR2 lead to changes, but the overall critical areas remain the same (see Figure 4-27, Figure 4-28).

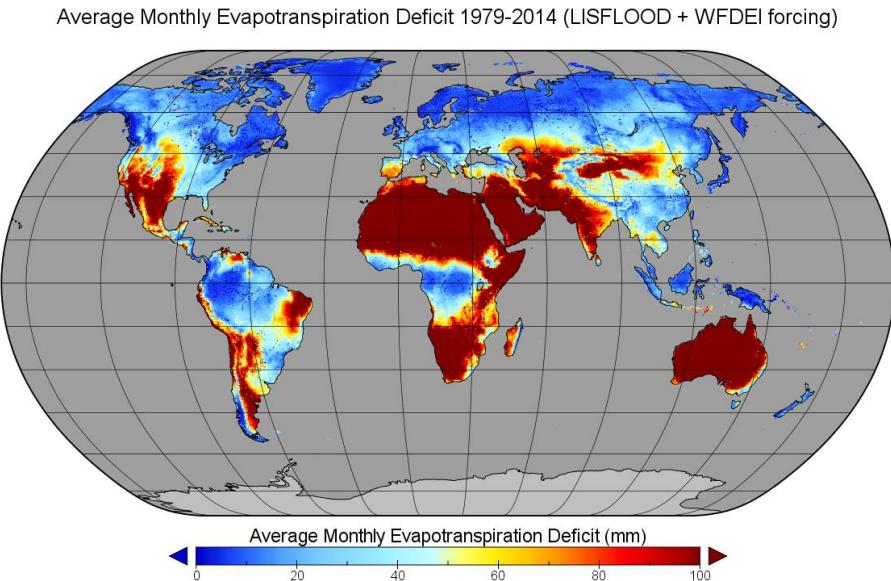


Figure 4-21 Estimated evapotranspiration deficit (LISFLOOD WRR2 with WFDEI forcing).

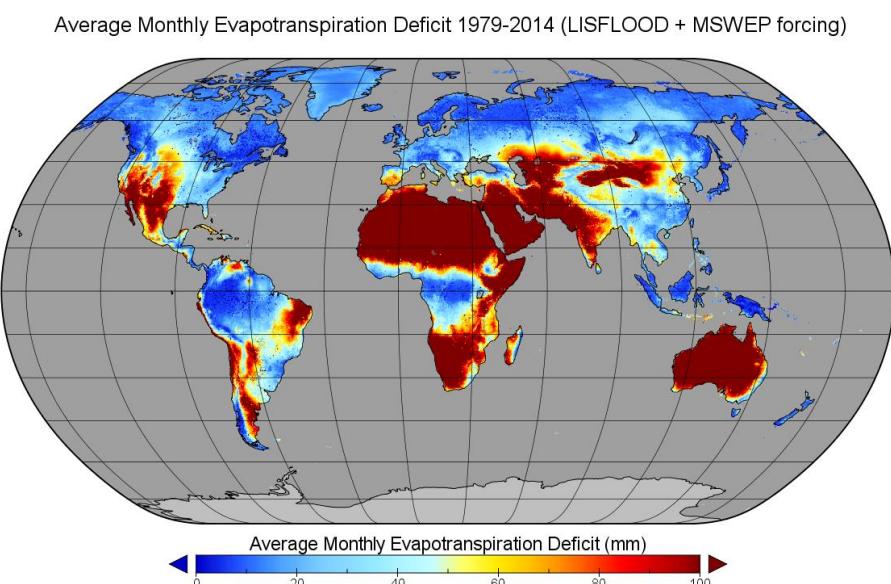
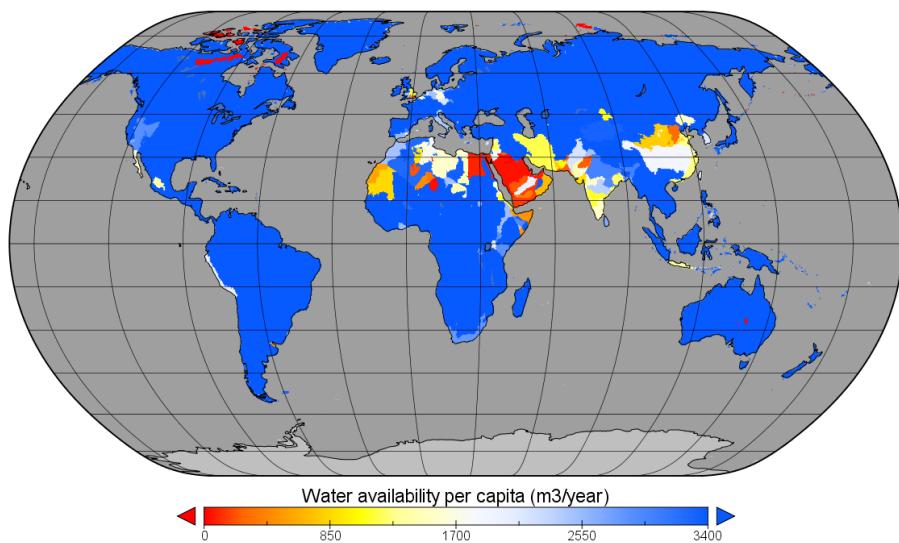
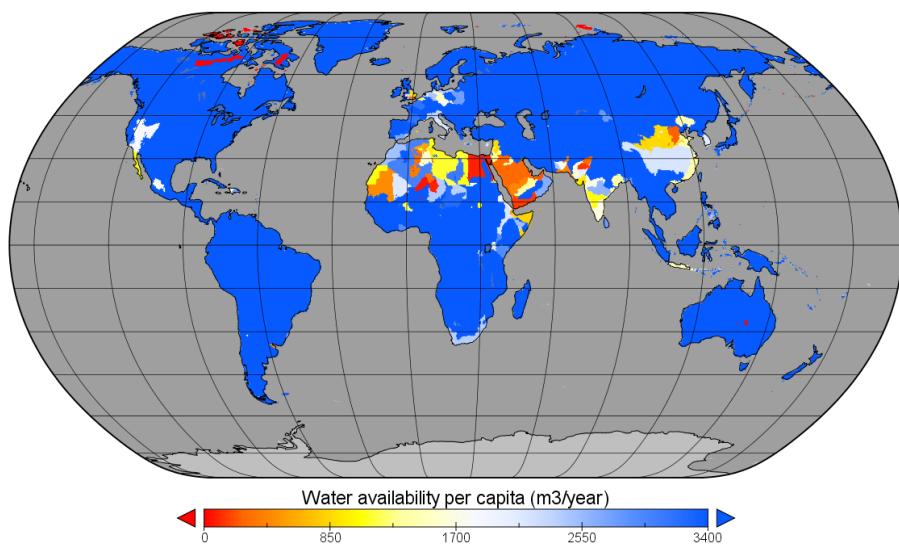


Figure 4-22 Estimated evapotranspiration deficit (LISFLOOD WRR2 with MSWEP forcing).

Average Water Availability per Capita 1979-2014 (LISFLOOD + WFDEI forcing)


Figure 4-23 Estimated average water availability per capita (LISFLOOD WRR2 with WFDEI forcing).

Average Water Availability per Capita 1979-2014 (LISFLOOD + MSWEP forcing)


Figure 4-24 Estimated average water availability per capita (LISFLOOD WRR2 with MSWEP forcing).

Average Monthly Water Exploitation Index 1979-2014 (LISFLOOD + WFDEI forcing)

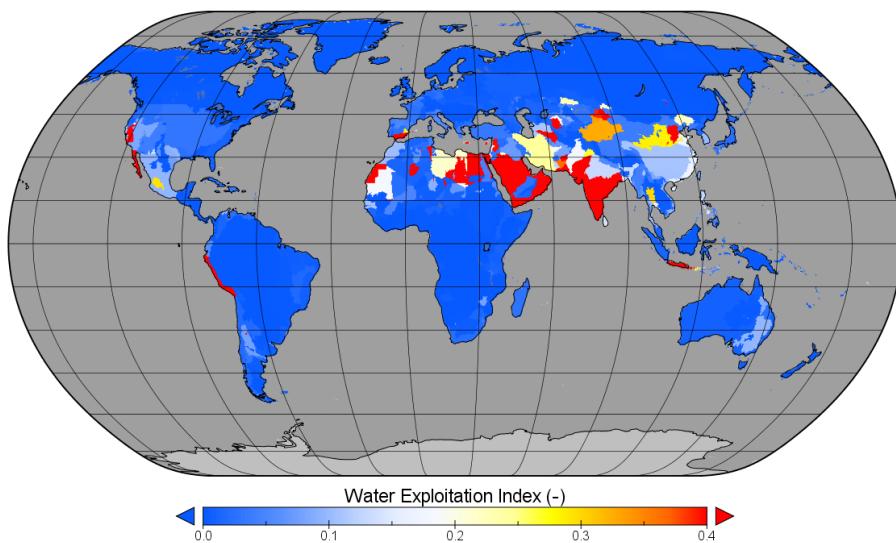


Figure 4-25 Estimated average water exploitation index (LISFLOOD WRR2 with WFDEI forcing).

Average Monthly Water Exploitation Index 1979-2014 (LISFLOOD + MSWEP forcing)

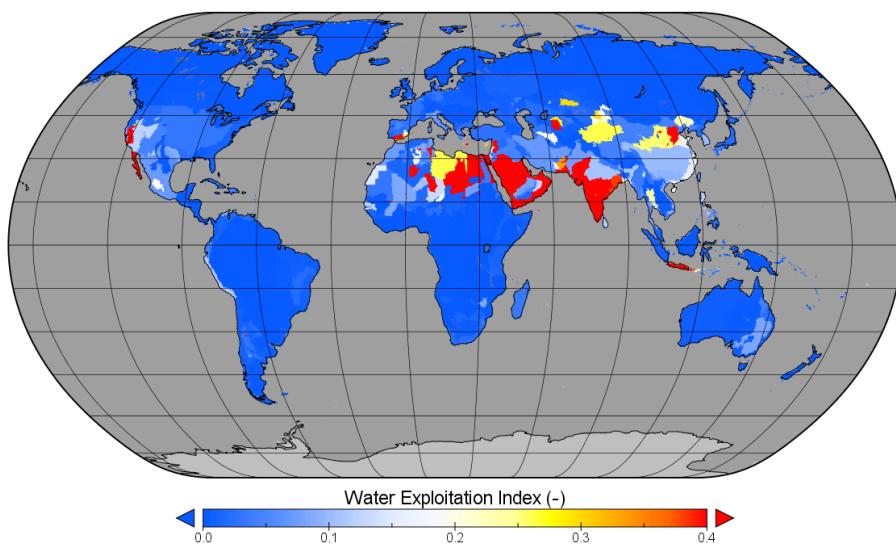


Figure 4-26 Estimated average water exploitation index (LISFLOOD WRR2 with MSWEP forcing).

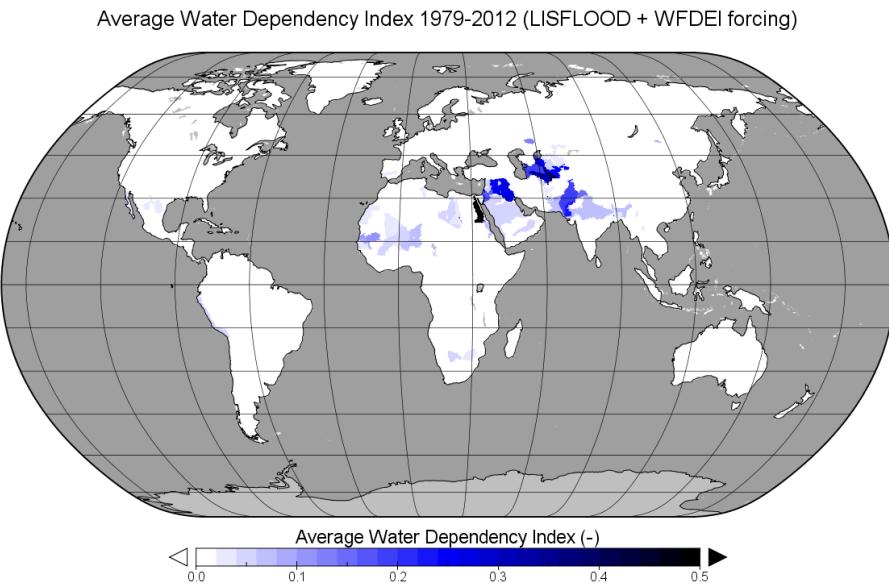


Figure 4-27 Estimated average water dependency (LISFLOOD WRR2 with WFDEI forcing).

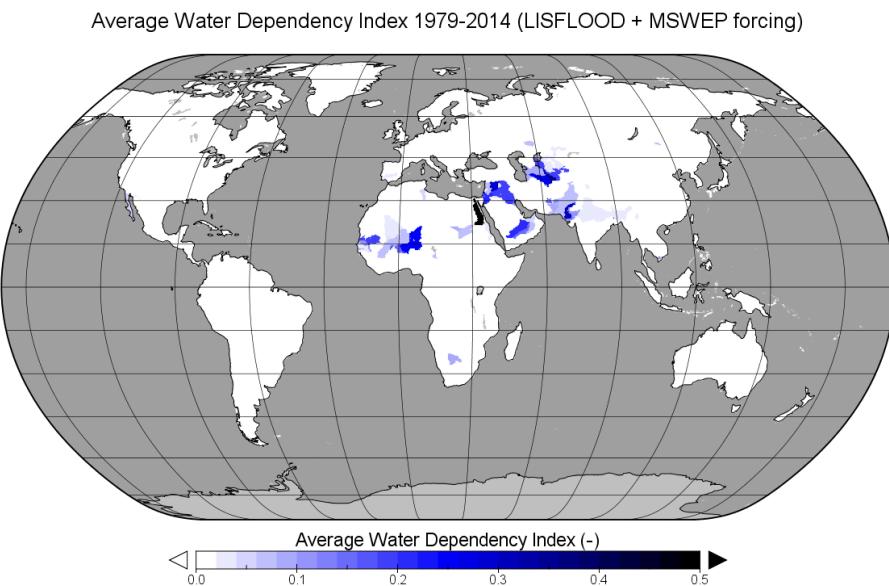


Figure 4-28 Estimated average water dependency (LISFLOOD WRR2 with MSWEP forcing).

(c) Impact of the precipitation ensemble forcing

The LISFLOOD ensemble runs with the various precipitation forcings while keeping all other input identical - analysed for the period 2000-2012 – shows that the LISFLOOD run with MSWEP is close to the observed discharge. The summed KGE for the stations in Table 4-10 is 3.5. WFDEI is the second-best performer, while the real-time and uncorrected TRMMRT shows the worst performance. For some basins, though – Ob, Ebro, Thames, Amazon, Congo -, TRMMRT is producing reasonable results.

Table 4-10: Kling Gupta Efficiencies (KGE) for various rivers for the various WRR2-ENS simulations with various precipitation forcings.

RiverName	StationName	ENS_WFDEI	ENS_MSWEPE	ENS_TRMMRRT	ENS_TRMM	ENS_CMORPH	ENS_GSMAP
LENA	KYUSYUR (KUSUR)	0.147	0.089	-0.253	-0.252	-0.481	-0.437
YENISEY	IGARKA	0.286	0.221	-0.513	-0.493	-0.502	-0.451
MACKENZIE	ARCTIC RED RIVER	0.649	0.683	0.078	0.078	0.013	0.083
OB	SALEKHARD	0.306	0.390	0.318	0.194	0.151	0.001
Rhone	Le_Rhone_a_Beaucaire	0.669	0.605	0.307	0.427	-0.041	0.006
Ebro	EbroEnTortosa	0.258	-0.084	0.247	0.378	0.075	0.133
Po	Lagoscuro	0.494	0.430	0.202	0.356	-0.003	0.004
Elbe	NeuDarchau	0.401	0.460	0.207	0.291	-0.155	0.105
Oder	Hohensaaten	0.584	0.447	0.001	0.050	-0.755	-0.132
RHINE	REES	0.640	0.668	0.101	0.317	-0.252	0.039
VOLGA	VOLGOGRAD POWER PLANT	-0.431	-0.595	-0.180	-0.220	-0.325	-0.207
DANUBE	CEATAL IZMAIL	0.745	0.801	0.158	0.593	0.059	0.262
THAMES	KINGSTON_ON_THAMES	0.832	0.651	0.607	0.607	0.245	0.436
Sava	S.Mitrovica	0.701	0.746	0.073	0.542	0.199	0.259
Tisza	Senta / Szeged	0.806	0.573	0.185	0.592	-0.372	0.101
MISSISSIPPI	VICKSBURG, MS	0.351	0.504	0.195	0.500	0.575	0.564
AMAZONAS	OBIDOS - LINIGRAFO	0.813	0.852	0.624	0.830	0.889	0.792
CONGO	KINSHASA	0.523	0.452	0.638	0.361	0.585	-0.696
ZAMBEZI	KATIMA MULILO (64370001)	-0.449	-0.871	-1.268	-0.992	0.168	0.551
ORANGE	VIOOLSDRIF (27811003)	-0.922	-0.265	-1.400	-1.507	-0.593	-0.189
MURRAY	BELOW WAKOOL JUNCTION	-5.733	-3.263	-4.020	-2.839	-2.715	-3.435
TOTAL KGE		1.671	3.495	-3.695	-0.185	-3.234	-2.214

4.4 ORCHIDEE

4.4.1 Model description

The model used is version #4030 of the ORCHIDEE-ROUTING branch. This is to very small details the same release as the one to be used by IPSL for CMIP6. A number of innovations have been introduced in this version compared to the one used for WRR1 in Earth2Observe.

(a) New ancillary data

The ORCHIDEE group has taken the opportunity of CMIP6 and the new regional applications to update most of the ancillary data used to set surface parameters. The objective is to take advantage of new data sets, often remote sensing based, as well as increasing the resolution of the maps used.

The latest version of the ESA CCI land cover is used. In contrast to the previously used IGBP map, this one is directly reporting fractions of plant functional types (PFT) as defined by ORCHIDEE. Thus, it is not anymore the model which transforms the Olson classification into the PFTs and thus this operation could be done with more care. This in particular decreases the fraction of bare soil for many ecosystems. Based on the ESA-CCI land cover, a probability distribution of PFTs is computed for each grid box. The model then only interpolates this distribution onto the grid used. This allows to preserve more of the original information.

For the soil texture classes the model has moved from the Zobler maps (Post and Zobler, 2000) to the USDA maps. The original maps were at 0.5° resolution and the

new ones are at 0.08° . This allowed to increase the resolution at which soil types are described and thus represent contrasts in the vertical water transports up to higher resolutions of the model.

In WRR1, ORCHIDEE used a soil colour map (Dickinson et al., 1993) to define the background albedo of soils, i.e. when the surface is not covered by vegetation. This has now been replaced by a map derived from MODIS albedo using the vegetation states predicted by ORCHIDEE. By assimilating the MODIS albedo into the model the background albedo was determined as the residual of the vegetation. It thus ensures that, given the correct vegetation albedo, the soil will have the right reflective characteristics (Peylin et al. Personal communication). This procedure also produced a new table of optimal albedo for the 13 PFTs simulated by ORCHIDEE.

(b) New formulation for surface roughness

In WRR1, ORCHIDEE had the same surface roughness for momentum and sensible and latent heat fluxes. This is a simplification which needed to be removed from the model and it was chosen to use the parametrisation proposed by Su et al. (2001). Based on the “localized near-field” Lagrangian theory the model predicts the roughness values to compute latent and sensible heat fluxes based on the state of the canopy. The implementation of this approach in ORCHIDEE was tested against a large number of FLUXNET stations and proved to reduce the systematic biases which characterized our model. In particular the overestimation of bare soil evaporation was largely reduced. It also produced a very similar effective roughness length as the 3D energy balance version of ORCHIDEE (Ryder et al., 2016). This innovation in our model reduces the overestimation of evaporation which was noted in the WRR1 simulation.

(c) New snow scheme

A major innovation in ORCHIDEE for CMIP6 and WRR2 is the introduction of a multilayer snow scheme. This scheme is documented in Wang et al. (2013). Compared to the version described in the original paper, some modification had to be applied so that the snow temperature profile can be solved implicitly and thus coupled with the energy balance of ORCHIDEE. Furthermore, the scheme has been extended to treat properly fractional snow covers in the grid boxes to make it more meaningful at the coarse resolution used in climate models. This scheme has been found to improve significantly the predicted albedo of snow covered regions and more generally is beneficial for the simulated water cycle of cold regions of the world.

(d) Soil freezing scheme

A soil freezing scheme has been introduced by (Gouttevin et al., 2012) in the multilayer soil moisture model. After a validation period this formulation is now fully operational in ORCHIDEE. Some improvement had to be made to ensure that during brief freezing events, which only affect the top soil layers, not all the rainfall or snow melt is directed toward surface runoff. So the impact of soil freezing on infiltration has been revised so as to take into account the sub-grid processes which allow incoming water to infiltrate and be stored in the soil. This avoid early soil drying in the summer in boreal regions.

(e) New routing scheme

The model has been equipped with a new version of the routing scheme which is based on the HydroSHED database (Lehner and Grill, 2013) which provides a description of the world's rivers at 1km resolution. This allows to route the runoff and drainage produced by ORCHIDEE to the oceans and endorheic lakes and thus simulate the discharge of all major rivers at their gauging stations. These developments are documented in a publication under preparation (Nguyen Quang et al. 2017). This routing scheme is an ideal tool to validate the continental water cycle simulated by the model and also offers an opportunity for data assimilation.

As the HydroSHED database is limited to the continents south of 60° North, it cannot yet be applied at the global scale. But it has been used here to validate the model and the Earth2Observe forcing data set over the Mediterranean drainage basin as shown below.

4.4.2 Data assimilation

Data assimilation has been used at CNRS for two applications with the Earth2Observe project :

- to determine the optimal background albedo and the albedo of the 13 plant functional types by merging the model output with MODIS observations, and
- to evaluate the evaporation error over the Mediterranean basin in WRR2 by assimilating into the routing the GRDC (Fekete et al., 2000) data base of observed river discharge.

Results of these studies will be reported elsewhere.

4.4.3 Evaluation

In an effort to validate the Earth2Observe forcing data for WRR2 (hereafter, E2OFD) we compare the WRR2 simulation with the WRR1 simulation, which used the WFDEI atmospheric conditions, and an independent simulation of WRF/ORCHIDEE at a resolution of 20km². The region chosen is the Mediterranean basin as it is the focus of the HyMex project and includes a number of test basins chosen for Earth2Observe. The period for the validation was chosen to be 1990-1999 as this is a golden period of the GRDC database.

Within the Mediterranean basin, two regions are presented here as they were found to show the largest precipitation difference between MSWEP (the precipitation product used in E2OFD) and WFDEI. These regions are the Caucasus and the Alps. To simplify the discussion, only the mean annual cycle over the 1990-1999 period is presented here.

(a) The Kura basin

The Kura basin takes its source in the Caucasus and thus an ideal basin to evaluate which forcing data provides the most realistic discharge and help us decide which precipitation estimation is more realistic. Figure 4-29 shows that the difference between the precipitations of E2OFD, WFDEI and WRF are quite large with the first

one clearly showing the highest values. On the other hand, the phase of the annual cycle is similar between E2OFD and WFDEI.

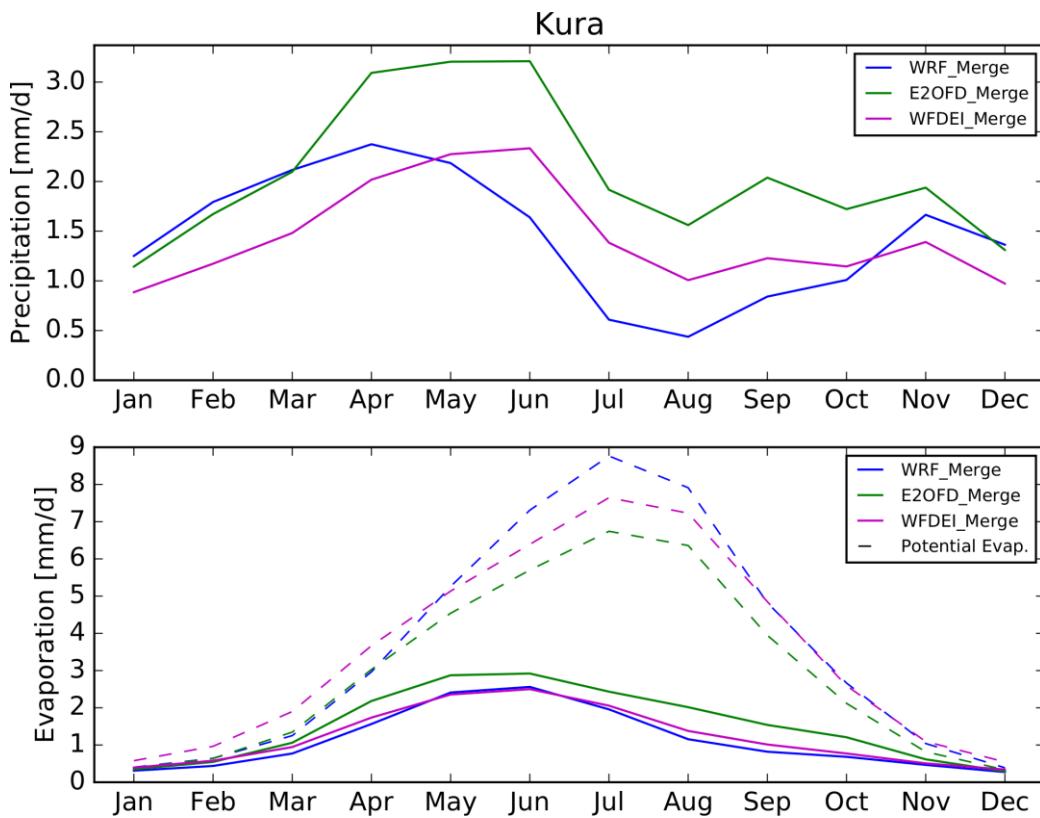


Figure 4-29: Precipitation and simulated evaporation averaged over the Kura basin. The dashed lines in the lower panel are potential evaporation estimated by ORCHIDEE with the forcing. The 3 simulations presented in these figures are : WRF, E2OFD and WFDEI forcing.

The other variable which is largely driven by the forcing is the Penman-Monteith potential evaporation, shown as dashed lines in Figure 4-29. Here we clearly see that WRF has the largest values and E2OFD the lowest. This does not necessarily lead to corresponding differences in the actual evaporation as can be seen in the same figure. The atmospheric demand can be better satisfied in E2OFD with large precipitation values and thus soil moisture stresses. This leads to a higher total evaporation in E2OFD when compared to the two other estimates.

The resulting discharge of the Kura can be compared to the climatology (i.e. a long period before the years 1990-1999) at the station Surra (see Figure 4-30). It shows that with the E2OFD and WRF atmospheric conditions the model carries too much water in the river. On the other hand for WFDEI the annual mean discharge is very close to the observations and the model only has an error in the phasing of the peak flows. From this analysis we can conclude that MSWEP has in all likelihood too large precipitation values over the Caucasus. The errors in WRF and WFDEI are difficult to attribute. These can be the result of model or forcing deficiencies.

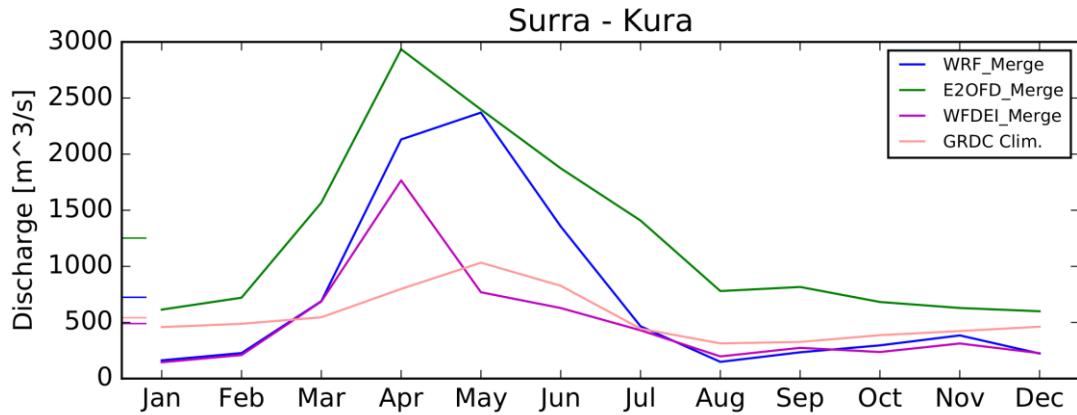


Figure 4-30: Discharge simulated by ORCHIDEE and observed for the Kura river at the Surra

(b) The Rhone and Po

The two large basins fed by Alpine precipitation, the Rhone and the Po, are also interesting. Over the Alps, the WFDEI forcing shows the largest mean annual rainfall with the largest differences mostly during the autumn. Over the Rhone basin, WRF displays larger precipitation values in winter when compared to WFDEI. Overall it is E2OFD which has the smallest precipitation values over this mountain system.

In both the Rhone and Po (Figure 4-31 and Figure 4-32) the simulation forced by E2OFD shows too low flows during the summer. The annual mean discharge is also strongly underestimated. On the other hand the WFDEI forced run displays discharge values very close to observations. It is most obvious in the Rhone that WRF has too much snowfall in winter and thus exaggerates the flow of the river during the melt period.

This short discussion focuses on two problem areas of the E2OFD forcing. This should not hide the fact that in general the behaviour of ORCHIDEE is much better when forced with E2OFD as could be evaluated with over 300 gauging stations on the rivers of the Mediterranean region. The higher resolution brings spatial details to the model's simulations which are in good agreement with WRF but absent from WFDEI. For a complete analysis of the simulated hydrological cycle of the mountainous areas within the Mediterranean the separation of precipitation in rainfall and snow fall should be analysed in detail. ORCHIDEE is quite sensitive to this distinction. A more complete analysis of these simulation will be proposed in a separate report after an in depth study of the three simulated continental water cycles.

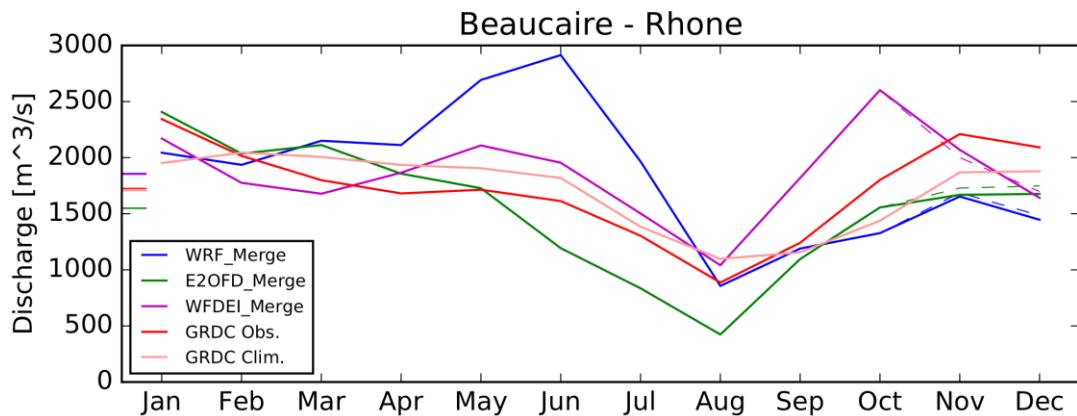


Figure 4-31: Observed and simulated discharge of the Rhone at the Beaucaire station.

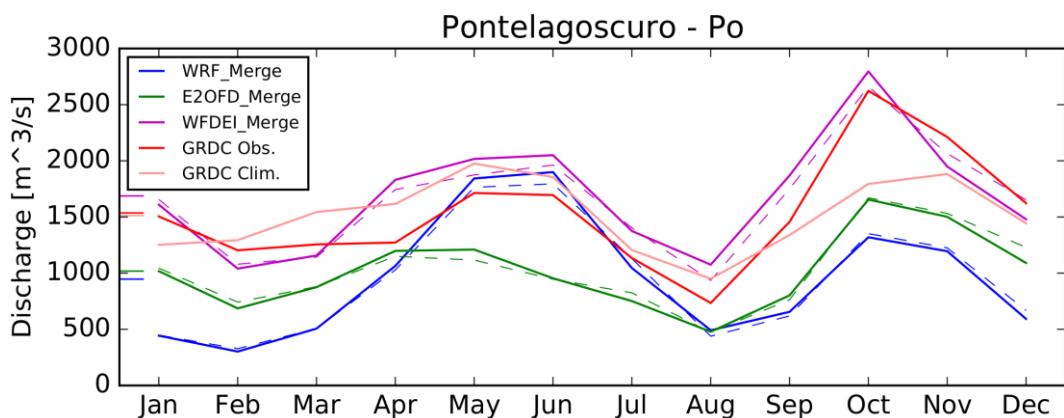


Figure 4-32: Observed and simulated discharge of the Po at the Pontelagoscuro station.

4.5 PCR-GLOBWB

4.5.1 Model description

The global hydrological model PCR-GLOBWB (PCRaster GLOBal Water Balance, van Beek et al., 2011) is a leaky bucket type of model providing a regular grid based representation of terrestrial hydrology. PCR-GLOBWB is forced by temperature, precipitation and reference potential evapotranspiration. PCR-GLOBWB responds to the prescribed atmospheric conditions through simulating the resulting fluxes and changes in the water content of its different storages (i.e. canopy interception, snow, soil, groundwater, channels and water bodies) for the land and water surface within each cell.

Over land, precipitation can fall as rainfall and snow. Snow fall, accumulation and melt are temperature dependent and modelled using a temperature-based snow module similar to the HBV model (Bergström, 1995) using 0°C as melt temperature. Any precipitation can be intercepted and evaporate, the remainder reaching the soil surface where it is partitioned into direct runoff and infiltration that enters the upper of two stacked soil layers. Through the soil layers water moves vertically and can come to delayed runoff as interflow or base flow. Base flow originates from the groundwater

store in that is situated below the second soil layer and that is replenished by net recharge (percolation minus capillary rise). Water stored in the groundwater reservoir is exempt from any direct influence of vegetation.

All drainage from the soil and groundwater layers accumulates as discharge and is routed along a river network. The routing can be described by the kinematic wave or, alternatively, by the computationally more efficient travel time approach. Channels can be intersected by lakes and reservoirs with one of the three water surface types prevailing in a cell. From all water surfaces water can be lost to evaporation.

In addition to the subdivision between land and open water, PCR-GLOBWB includes sub-grid variability with regards to land cover, soil conditions and lithology. Land cover is subdivided on the basis of the GLCC dataset (USGS EROS Data Center, 2002) for natural vegetation and this can be complemented with agricultural land cover types (e.g., paddy and non-paddy irrigated areas and rain-fed agriculture). For each cover type, vegetation phenology and crop calendars are prescribed by repeating a specific annual course. Sub-grid variations in soil and hydro-lithological properties are derived from global data sources (FAO DSW, Gleeson et al., 2014) and assigned as effective value per land cover type. A noticeable exception is the distribution of the water holding capacity that is parameterized by its mean and minimum value and a shape factor, representing the combined effect of soil and vegetation properties which is used to generate the direct runoff using the Improved Arno Scheme (Todini, 1996; Hagemann and Gates, 2003). The distribution of the surface elevation within each cell, as used to compute inundation and the capillary rise, is described using a cumulative function of the 1×1 km Hydro1k data set (<https://lta.cr.usgs.gov/HYDRO1K>).

Further anthropogenic interference with the global water cycle can be included optionally by considering dynamic withdrawal, allocation and consumptive use of ground- and surface water resources -including irrigation- and return flows of unconsumed water to surface water and groundwater resources (Wada et al., 2014; Sutanudjaja et al., 2014). Allocation of the gross demand and subsequent return flows is dynamic and responds over time to the changing availability of surface water and groundwater (de Graaf et al., 2014).

Reservoirs can be included, and more than 6000 reservoirs can be included in the parameterization on the basis of the GRanD (Lehner et al., 2011) database, and alter the river flow regime as a function of storage, inflow and downstream demand.

Since WRR1, several improvements in the underlying model code have been carried out. For Tier 1 of Earth2Observe, PCR-GLOBWB simulations were carried out to provide a baseline reference run and only “natural” (non-humanly modified) components of the water cycle were considered. For Tier2, these human influences are included in full:

- Paddy and non-paddy irrigated areas have been included in the land cover parameterization in addition to the natural short and natural tall vegetation types considered in WRR1 and the prescribed areas represent the development over time;
- Irrigation water demand is computed for these areas using the FAO guidelines (Allen et al., 1998);
- For three other sectors -livestock, municipalities and industry- water demand is prescribed at a monthly resolution and reflecting the historical development

over the period 1961-2014. Gross demand prescribes the potential withdrawal and net demand the eventual return flows;

- Desalination is included as an external supply (Wada et al., 2011);
- Reservoirs, available in the GRaND dataset, have been included.

In WRR2, the temporal resolution of all simulations was daily. The spatial resolution of the simulations was either 0.5° or 0.08333° (30 and 5 arc minutes respectively) globally. The degraded spatial resolution of 0.5° was preferred to increase the computational speed for the more demanding simulations and to facilitate comparison with the outcomes of WRR1. For computational expedience, also the routing was performed with the less demanding travel time option. Meteorological input for WRR2 was available at 0.25° and was used without downscaling in the higher-resolution simulations in order to assess the effects of model resolution more directly. Precipitation and air temperature provided at the daily resolution and used accordingly. Potential reference evapotranspiration was computed from the provided non-precipitation data according to the FAO guidelines (Allen et al., 1998). For all simulations, a substantially long spin-up period was used, typically using the meteorological forcing twice back-to-back before the eventual simulation was carried out.

4.5.2 Data assimilation

Pending the completion and analysis of this data assimilation, a brute force simulation has been carried out including 24 members, including the global pre-factors for the parameters of PCR-GLOBWB shown in Table 4-11. This dataset is evaluated against GRDC discharge over the available period. Eventually, as part of the data assimilation within WRR2, the global hydrological model PCR-GLOBWB is embedded in an Ensemble Kalman Filter framework (EnKF). Hydrological states and fluxes are constrained on the basis of the available in-situ river discharge measurements of the Global Runoff Data Centre (GRDC) and soil moisture measurements (ESA-CCI, Dorigo et al., 2015). The assimilated information will lead to a calibrated version of PCR-GLOBWB in which parameters are related to the snow module, rainfall-runoff generation, unsaturated zone process and base flow generation. The calibration period covers 1979-2014 –the later years affected by the absence of discharge measurements, the earlier ones by the absence of soil moisture- and the performance is evaluated using the long-term GRDC dataset and evaporation from satellites and FLUXNET observations (Miralles et al., 2011; Jung et al., 2009). Calibration is performed at the 0.5° spatial resolution, including all available human influences and starting from 1961 as this is the earliest year with demand information, replicating the climate forcing of WRR2 up to the start of the actual simulation period in 1979. A similar setup has also been adopted for the intermediate brute force simulation.

Table 4-11: Parameter combinations for the brute force calibration. 24 parameter combinations are used ($3 \times 4 \times 2$). The default parameter setting (Run 10) has been omitted in favour of Run 24.

Saturated hydraulic conductivity	Recession coefficient	Storage capacity	Saturated hydraulic conductivity	Recession coefficient	Storage capacity
----------------------------------	-----------------------	------------------	----------------------------------	-----------------------	------------------

Ru n	(log-scale)	(log-scale)	(linear scale)	Run	(log-scale)	(log-scale)	(linear scale)
0	-0.25	-0.50	1.00	13	0.00	0.50	1.25
1	-0.25	-0.50	1.25	14	0.00	1.00	1.00
2	-0.25	0.00	1.00	15	0.00	1.00	1.25
3	-0.25	0.00	1.25	16	0.25	-0.50	1.00
4	-0.25	0.50	1.00	17	0.25	-0.50	1.25
5	-0.25	0.50	1.25	18	0.25	0.00	1.00
6	-0.25	1.00	1.00	19	0.25	0.00	1.25
7	-0.25	1.00	1.25	20	0.25	0.50	1.00
8	0.00	-0.50	1.00	21	0.25	0.50	1.25
9	0.00	-0.50	1.25	22	0.25	1.00	1.00
11	0.00	0.00	1.25	23	0.25	1.00	1.25
12	0.00	0.50	1.00	24	0.00	0.00	0.80

4.5.3 Evaluation

The influence of the differences in meteorological forcing and model changes between WRR1 and WRR2 were evaluated progressively by the model setups listed in Table 4-12.

Table 4-12: Model setups evaluated in WRR2. All model runs cover the period 1979-2014 with the exception of WRR1_Nat30 which covers the period of the WRR1 forcing 1979-2012

Model name	Forcing	Spatial resolution	Details
WRR1_Nat30	WRR1	30"	Natural only: two land cover types; no water use or reservoirs, WRR1 forcing
WRR2_Nat30	WRR2	30"	As above but WRR2 forcing
WRR2_Hum30	WRR2	30"	Including full human interaction using 4 land cover types (incl. irrigation), water use and reservoir operations using WRR2 forcing
WRR2_Hum05	WRR2	5"	As above but with increased spatial resolution from 30 to 5"
WRR2_Hum30_BF	WRR2	30"	As above at 30" with consideration of the parameter combination of Table 1 (brute force calibration) pending the results of the full data assimilation
WRR2_Hum30_DA	WRR2	30"	As above but with EnKF, to replace the information of WRR2_Hum30_BF

For the evaluation of the model runs, the following aspects are considered:

- The components of the global water balance and their average seasonal variation over the common period (1979-2012);
- The agreement between monthly discharge as simulated and observed at selected stations of the Global Runoff Data Centre (GRDC);
- Total water demand and the provenance of water withdrawals from surface water and groundwater;
- The range in parameter values and performance of the calibrated model (by brute force or data assimilation) in terms of observed discharge and the storage of soil moisture and rate of actual evapotranspiration inferred from indirect measurements.

4.6 SURFEX-TRIP

4.6.1 Model description

The SURFEX-TRIP modelling system of the CNRM is based on the coupling of the ISBA Land Surface Model and the CNRM version of the TRIP river routing model. In the WRR1 version, ISBA computes the soil/snow/vegetation energy and water budgets and provides TRIP with runoff and drainage which are then routed within the river network.

Two main improvements have been implemented in the WRR2 version. First, a simple groundwater scheme has been added in the TRIP model to simulate the low-frequency variations of groundwater and the groundwater-river interactions (Vergnes and Decharme, 2012). The two-dimensional groundwater flow equation is solved for the uppermost unconfined aquifer and groundwater-river interactions are computed from the water level in both pools. The new parameters of the groundwater scheme are based on lithological and hydrogeological information available at the global scale. Namely, the World-wide Hydrogeological Mapping and Assessment Programme (WHYMAP, <http://www.whymap.org>) was used as primary information to delineate major regional groundwater basins concerned by diffusive groundwater movements. Moreover, the new groundwater scheme is coupled with ISBA to simulate upward capillary fluxes from shallow groundwater into the unsaturated soil column (Vergnes et al., 2014). The simulated water table depth is used as the lower boundary condition for the soil moisture diffusive equation. Compared to a free drain experiment, upward capillary fluxes at the bottom of the soil tend to decrease the simulated recharge from ISBA to the aquifers. Globally the impact on the river discharge remains low while the mean annual evapotranspiration slightly increases over the aquifer domains.

The second main improvement is the implementation of a flood model which describes the flood dynamics through the daily coupling between ISBA and TRIP including a prognostic flood reservoir (Decharme et al., 2012). This reservoir fills when the river height exceeds the critical river bank full height and vice versa. The flood interacts with the soil hydrology through infiltration and with the overlying atmosphere through precipitation interception and free water surface evaporation. Generally, the flood model significantly improves the discharge estimation first by increasing the evapotranspiration that limits the annual discharge overestimation, second by smoothing river peak flows when the floodplain storage is significant.

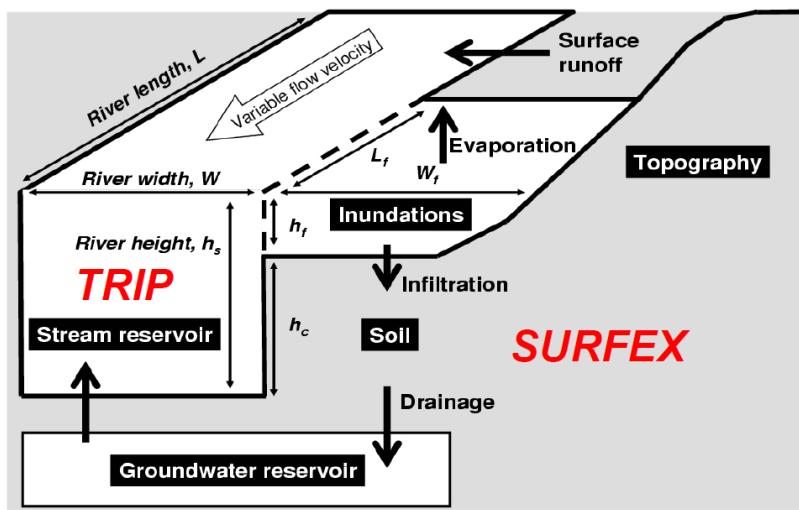


Figure 4-33: Scheme of the WRR2 model version, including the representation of floodplains and groundwater reservoir.

In WRR2, the ISBA simulations were carried out at the $0.25^\circ \times 0.25^\circ$ resolution. On the other hand, since no river routing network was available at this resolution, the resolution for the TRIP model was $0.5^\circ \times 0.5^\circ$ as in WRR1. Hence, output variables from TRIP (RivOut, SurfStor, GroundMoist) are given at this coarser resolution. Besides, a 50 years spin-up was carried out in order to get the groundwater storage at its equilibrium point. As in WRR1, in both ISBA and TRIP the default set of parameters was used, with no particular calibration performed for the simulations. The model state of ISBA and TRIP on 01-01-2000 from the baseline simulation was used as a restart for the ensemble simulations.

4.6.2 Data assimilation

In addition to the four experiments of WRR2 ensemble simulations based on different precipitation forcing fields, a data assimilation experiment has been conducted with the SURFEX-TRIP model. This experiment is an extension of the existing SURFEX Land Data Assimilation System to the global scale (LDAS-Monde) that was recently implemented over France (Barbu et al., 2014).

Here, two global operational datasets derived from satellite observations are assimilated simultaneously using a multivariate Extended Kalman Filter (EKF): (i) Surface Soil Moisture (SSM) from the ESA Climate Change Initiative and (ii) Leaf Area Index (LAI) from the Copernicus Global Land Service project. The SSM observations consist in daily maps at the resolution of 0.25° covering the period 1978-2015, while the LAI observations are given every 10 days at the resolution of 1 km over the period 1999-2015. The SSM data were bias corrected with respect to the model climatology by using a season-based CDF (Cumulative Distribution Function) matching technique. Each model grid cell is divided into a number of land covers, each having its own set of prognostic variables. The filter algorithm is designed to provide a distinct analysis for each land cover while using one observation per grid cell. The updated values are aggregated by computing a weighted average.

The SURFEX-TRIP model configuration and meteorological forcings are the same as for the WRR2 baseline simulation. Nevertheless, in order to limit the computation

time, the simulations were conducted at the global scale at a 1 degree spatial resolution over the period 2000-2013. In the next section, the results of the LDAS-Monde simulation are compared to a reference run in which no assimilation is done. The reference run is equivalent to the WRR2 baseline simulation, but with a 1 degree resolution.

4.6.3 Evaluation

In this section, the performances of the WRR2 experiment are analysed. The analysis criterion is based on the Nash-Sutcliffe Efficiency computed for a large number of world-wide in situ gage stations where discharge observations were available. Discharge measurements were collected from different sources, including the Global Runoff Data Centre (GRDC), the United States Geological Survey (USGS), the HYBAM network (HYdrologie et Biogéochimie du Bassin Amazonien) and French hydrological network (Banque HYDRO).

(a) Impact of the new forcing

To evaluate the impact of the forcing used in the WRR2 setup, an extra simulation has been carried out using the model setup of WRR1 and forcings from WRR2 (WRR1_F2). Results of this simulation are compared to results from WRR1 (WRR1_F1), so that only the forcing fields change from one experiment to the other. Figure 4-34 compares the Nash-Sutcliffe Efficiency over the available gauge stations. It is shown that the WRR2 forcing has neutral to positive impact on simulation of the discharge almost all over the globe, even though a large number of stations still show quite poor performances, such as in North America. As shown in Figure 4-35, where accumulated distribution of NSE is plotted for several regions of the globe, the improvement is quite significant, except over the Eastern part of USA and over the Mediterranean region. Finally, Figure 4-36 shows the seasonal cycle of the runoff for 14 major basins for both experiments and from in situ observations. The new forcing has a clear impact of discharge simulations. In some basins, such as the Amazon or the Mekong basins, it clearly improves the discharge, while the impact is negative for the Danube basin.

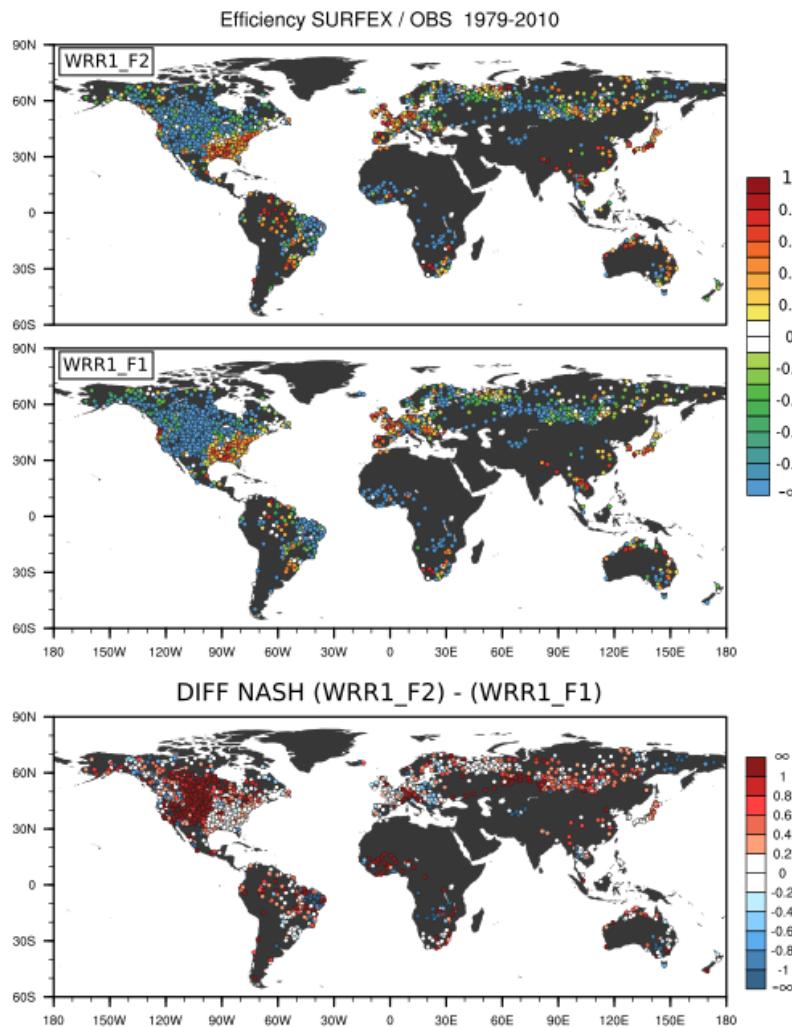


Figure 4-34: Comparison of Nash-Sutcliffe Efficiency for experiments WRR1_F1 (model WRR1, forcings WRR1) and WRR1_F2 (model WRR1, forcings WRR2).

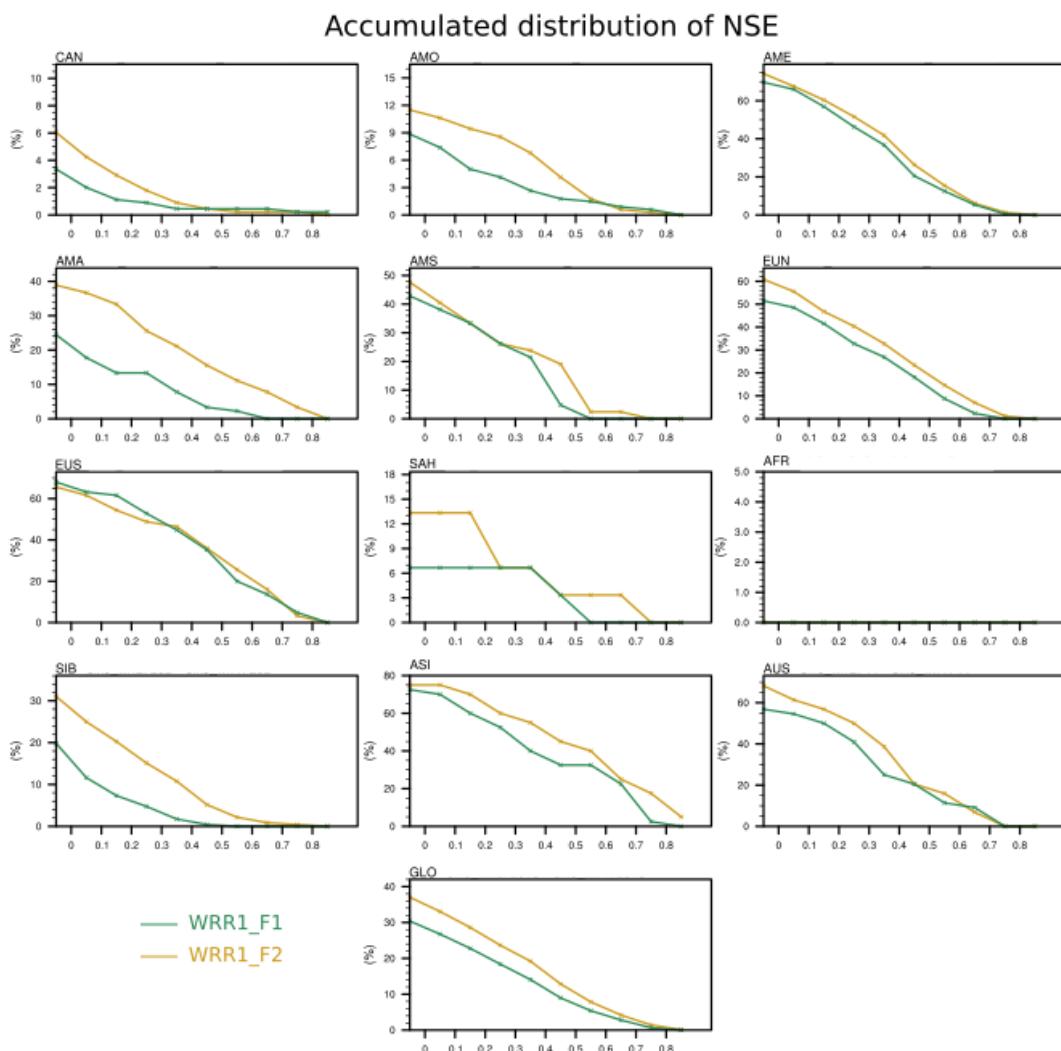
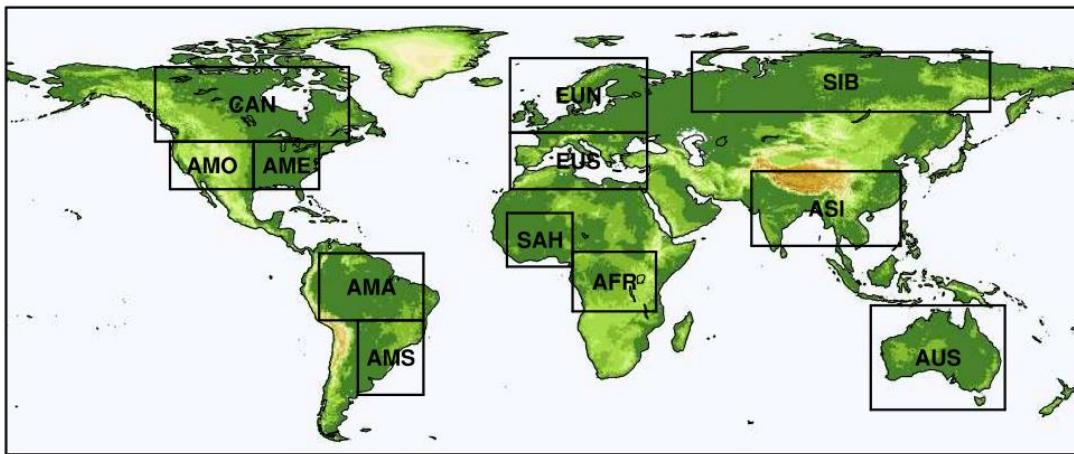


Figure 4-35: Comparison of the accumulated distribution of NSE over different regions of the globe, for both experiments WRR1_F1 and WRR1_F2.

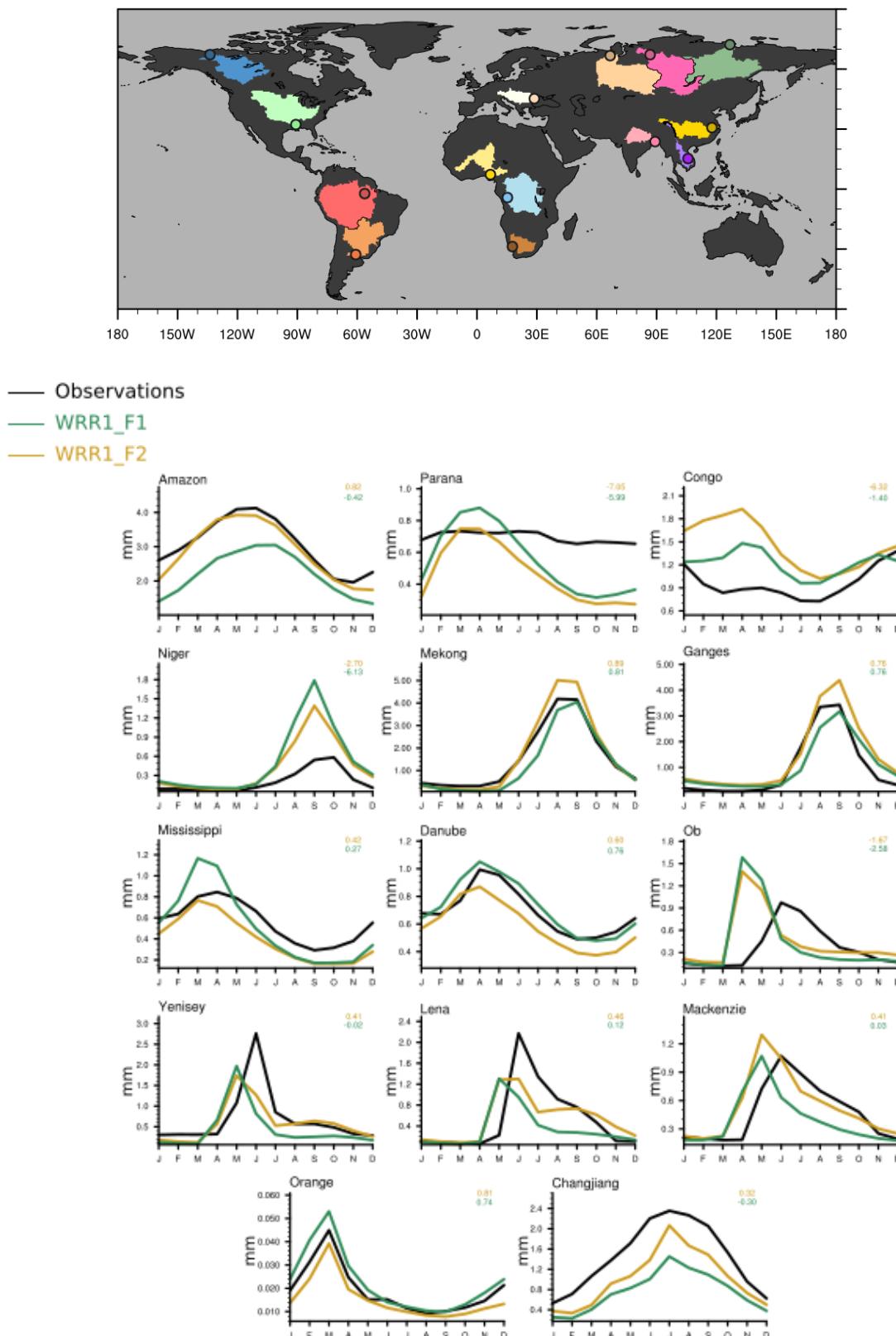


Figure 4-36: Seasonal cycle of the runoff for both WRR1_F1 and WRR1_F2 experiments and comparison with in situ observations.

(b) Impact of the model improvements

The same analysis is performed to analyze the impact of the model changes. The previous WRR1_F2 experiment is compared to the WRR2_F2 experiment which has been run with the WRR2 version of the model and the new forcings. Hence, in the following, only the impact of the model improvements are analyzed. From Figure 4-37, Figure 4-38, and Figure 4-39, it is clear that the model changes have much less impact than the forcing. Figure 4-37 and Figure 4-38 show that the impact is neutral or positive for every station. Most of the improvements are located in Siberia, Africa and in the Amazon basin. Figure 4-39 shows that most of the major basins are slightly impacted by the model changes. Nevertheless, the representation of the floodplain in the Niger River Basin significantly decreases the peak discharge, namely because of the direct evaporation from the Niger Inner Delta which is better represented. The effect is quite similar for the Ob river, even though the peak discharge occurs one month earlier than expected, which is probably due to a too fast melting of the snow layer.

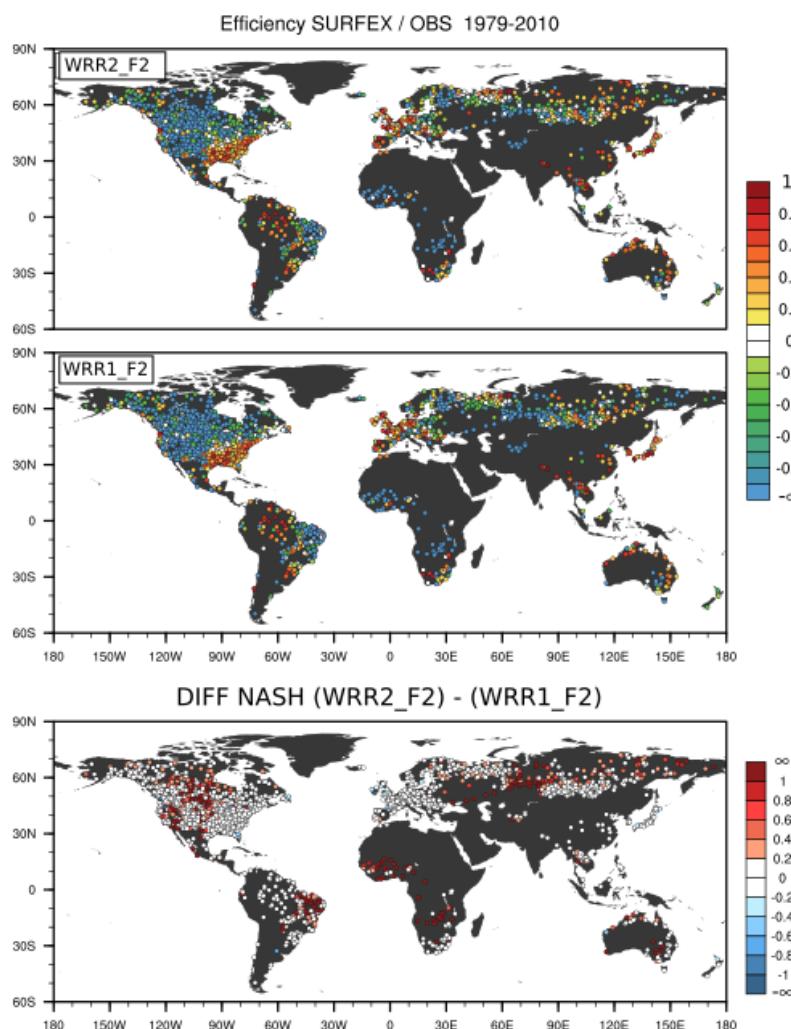


Figure 4-37: Comparison of Nash-Sutcliffe Efficiency for experiments WRR1_F2 (model WRR1, forcings WRR2) and WRR2_F2 (model WRR2, forcings WRR2).

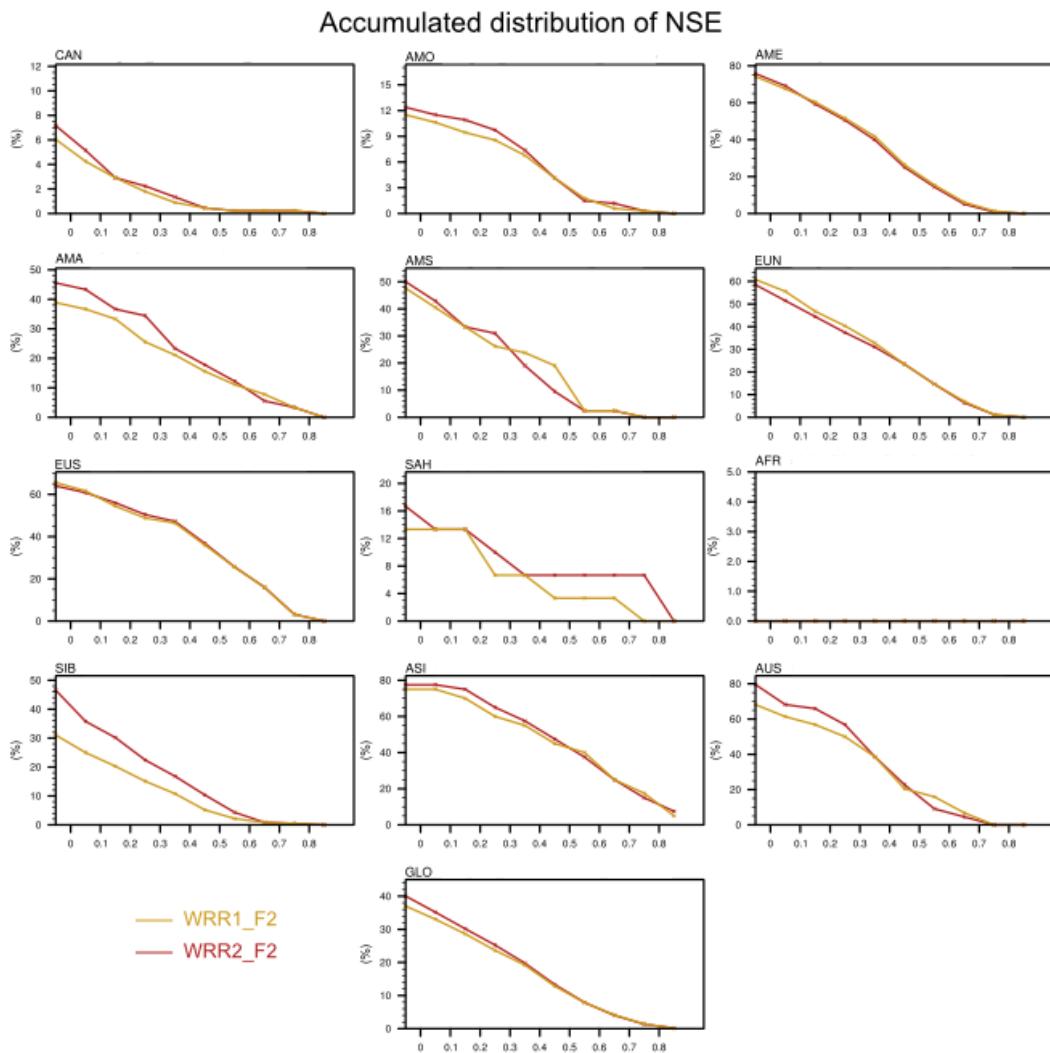


Figure 4-38: Accumulated distribution of NSE over different regions of the globe, for both experiments WRR1_F2 and WRR2_F2.

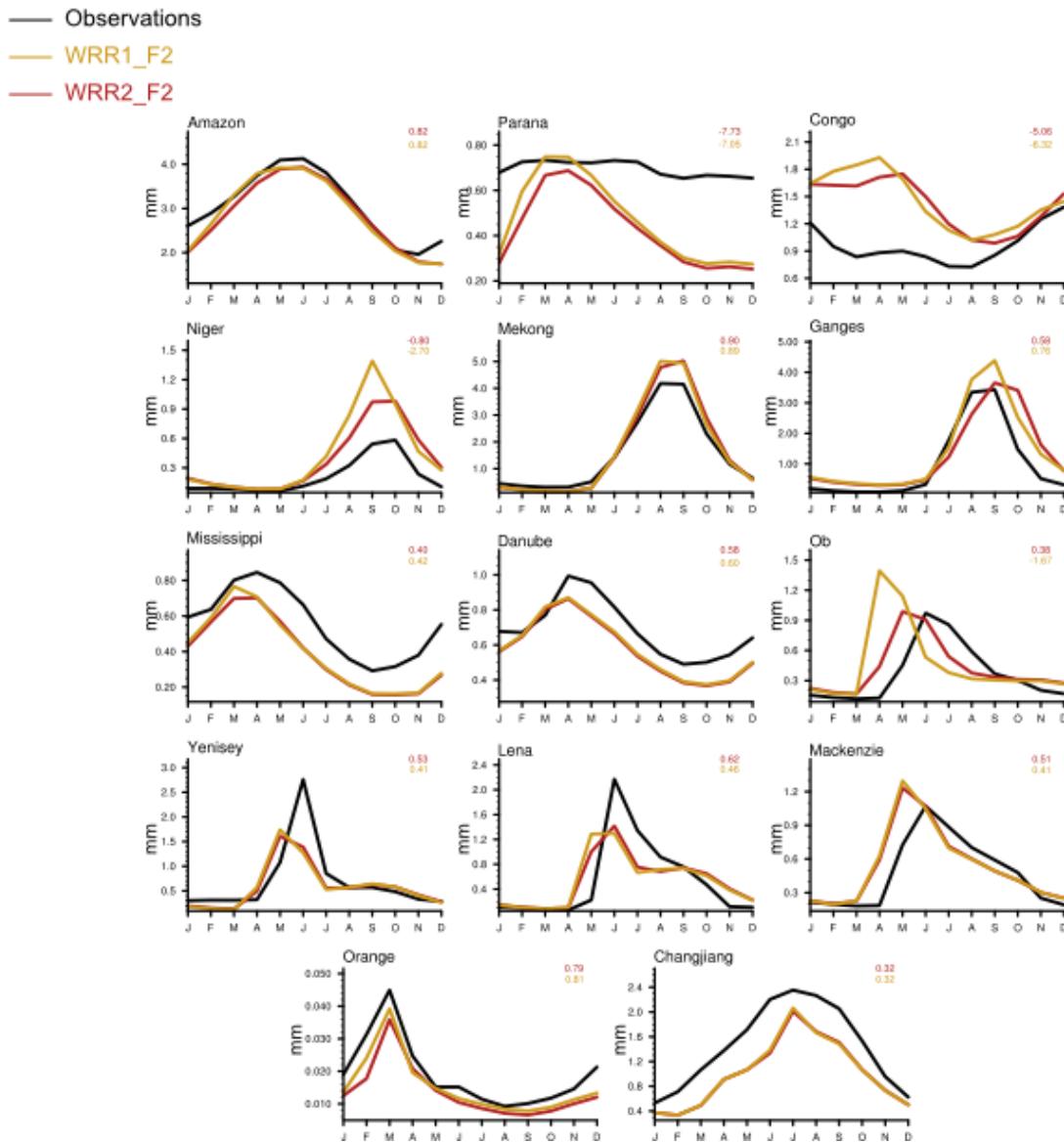


Figure 4-39: Seasonal cycle of the runoff for both WRR1_F2 and WRR2_F2 experiments and comparison with in situ observations.

(c) Analysis of the SURFEX LDAS-Monde

This section presents the main results of the SURFEX LDAS-Monde. Performances of the analysis are evaluated against an open-loop simulation in which no satellite data are assimilated. The open-loop simulation is referred as ‘Model’ in the results.

Figure 4-40 and Figure 4-41 show the improvement due to the assimilation, in terms of correlation and Root Mean Square Difference (RMSD) with respect to the observations, on both assimilated variables LAI and SSM. The column on the left and the one in the middle correspond to the open-loop and to the analysis, respectively, while the column on the right represents their difference. The assimilation has a clear impact on the LAI all over the globe, with resulting correlations close to 1 and very low RMSD compared to the open-loop. Namely, in the tropical zone and at high latitudes, the model is not able to correctly represent the vegetation dynamics, which is

efficiently corrected by the assimilation. The impact on the SSM is much less important but still positive.

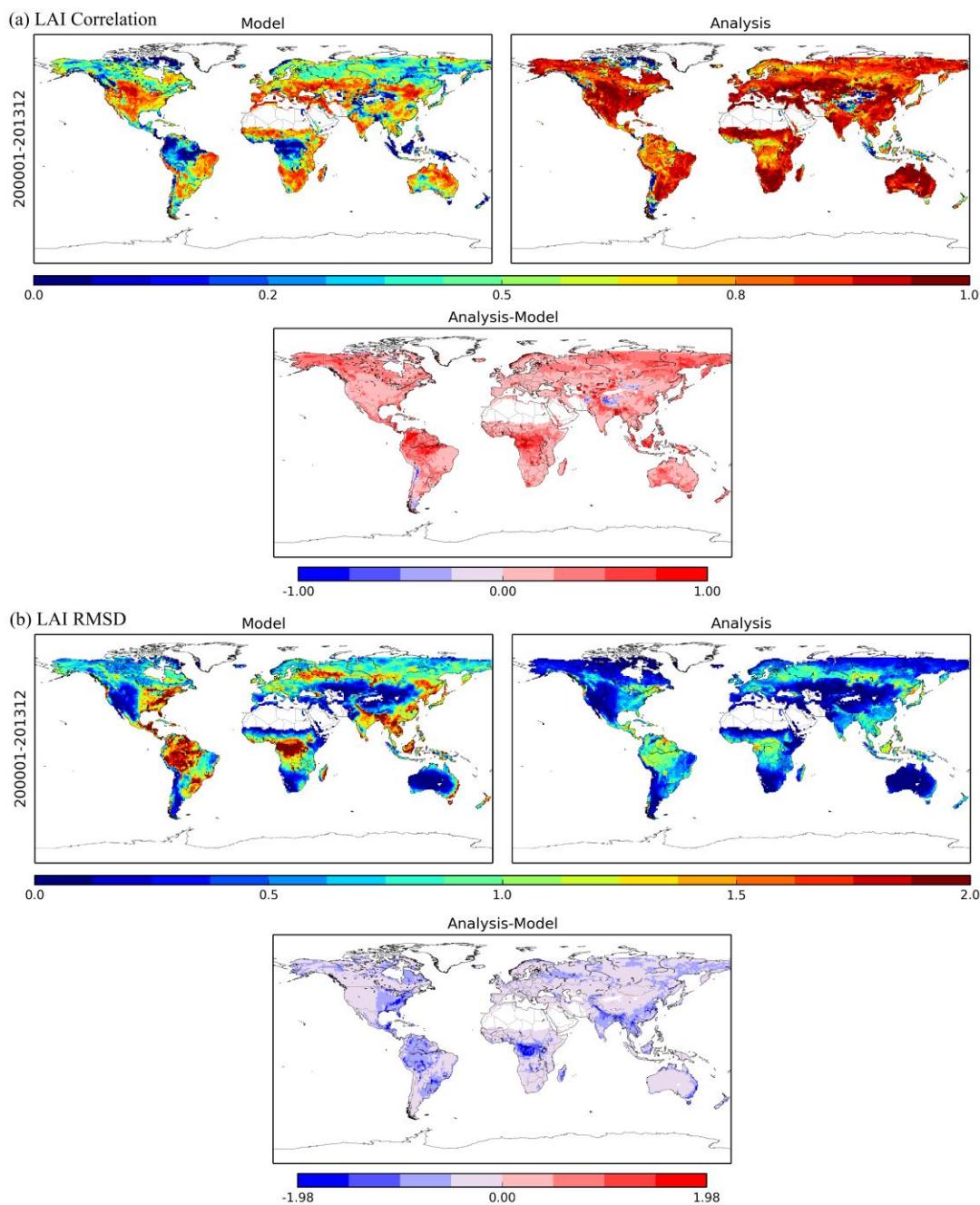


Figure 4-40: Correlation and Root Mean Square Difference (RMSD, in m^2/m^2) of LAI for experiments without data assimilation (Model) and with data assimilation (Analysis).

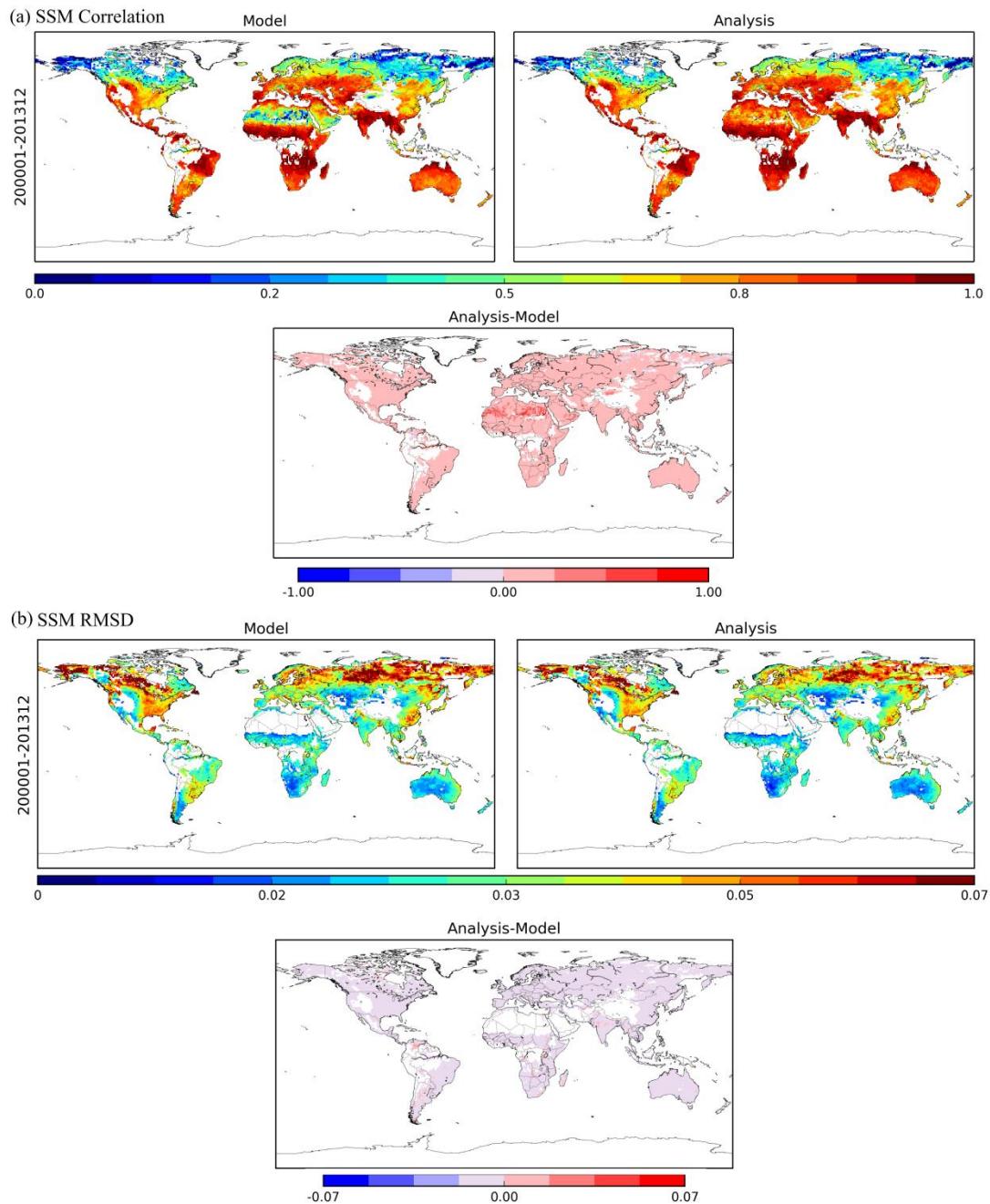


Figure 4-41: Correlation and Root Mean Square Difference (RMSD, in m^3/m^3) of SSM for experiments without data assimilation (Model) and with data assimilation (Analysis).

In Figure 4-42, the modelled LAI averaged over Europe is compared to observations for both the open-loop and the analysis. This figure shows that most of the corrections provided by the assimilation is done during the second part of the year. Indeed, the growing phase of plants is quite well simulated by the ISBA model. The simulated vegetation peak is a little delayed, as well as all the senescence phase, which is responsible for the limited performances of the open-loop. The assimilation is able to partly correct this delay. Similar conclusions can be drawn from Figure 4-43 which shows four metrics describing the seasonal performances of both the open-loop and the analysis: the standard deviation (SDD), the bias, the RMSD and the correlation. All

the metrics are greatly improved by the assimilation. Finally, the analysis increments averaged over Europe are plotted in Figure 4-44. The increments represent the values added or subtracted to the control variables (LAI and WG2, the latter representing the surface soil moisture) during the assimilation. It is shown that, for both variables, the corrections are quite seasonal and occur mostly during summer and autumn, which is consistent with previous remarks. Besides, the assimilation tends to remove some vegetation (negative LAI increments) and to add water in the soil (positive WG2 increments). It has to be noticed that the average of WG2 increments over Europe increased during the 14 years of simulation. The discontinuity of the satellite missions integrated into the ESA-CCI SSM product is potentially responsible for this increase. Namely, the number of available observations drastically increased in 2003 with the integration of AMSR-E observations, and since 2010 with the integration of SMOS (Dorigo et al., 2017, under review).

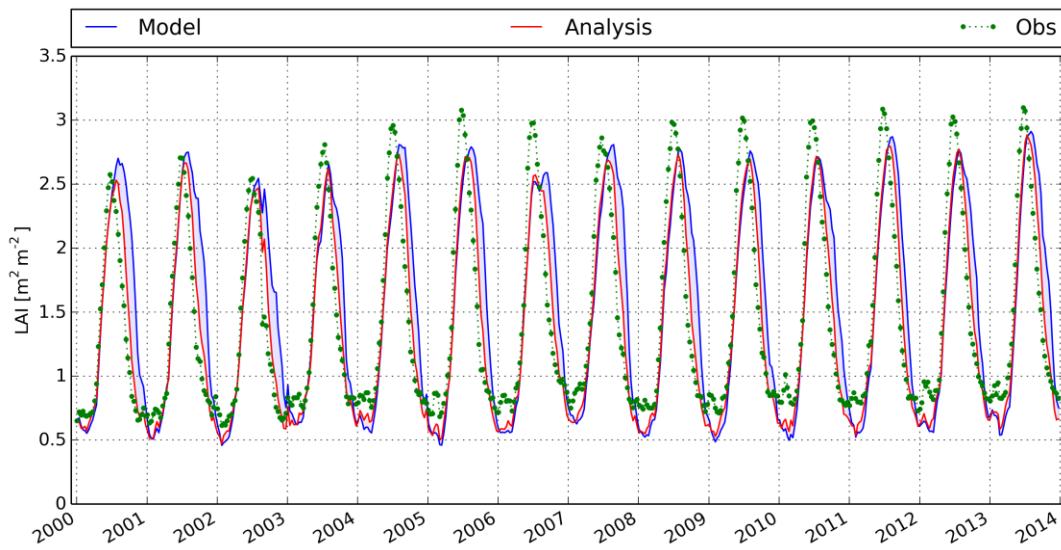


Figure 4-42: Comparison of LAI from open-loop, analysis and observation over Europe.

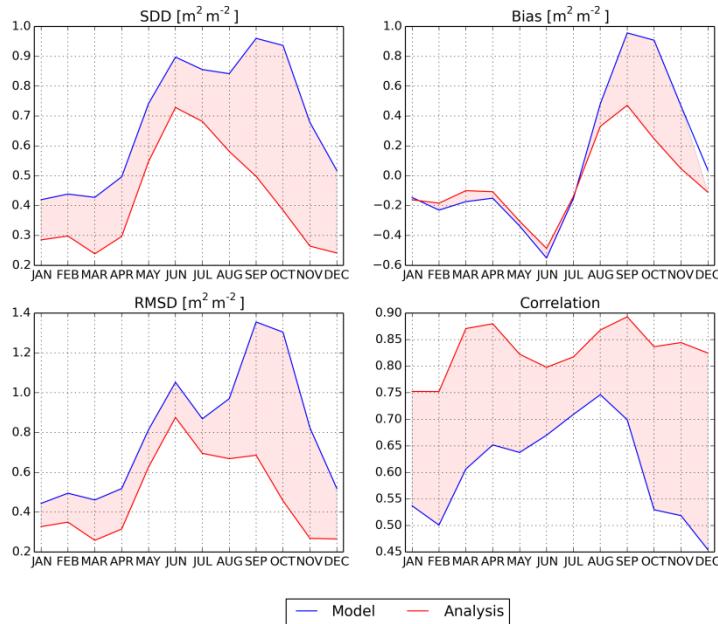


Figure 4-43: LAI seasonal scores of the open-loop and analysis over Europe: Standard Deviation (SDD), Bias, Root Mean Square Difference (RMSD) and correlation.

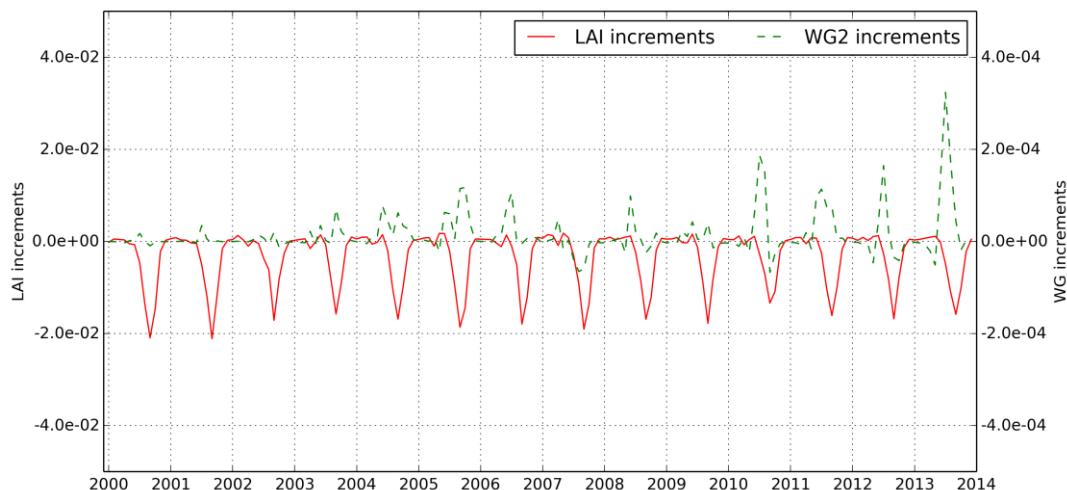


Figure 4-44: Time series of analysis increments averaged over Europe for both control variables. WG2 is the water content in the second soil layer (1cm-4cm). Units are m^2/m^2 for LAI increments and m^3/m^3 for WG2 increments.

Lastly, Figure 4-45 shows the impact of the assimilation on variables of the water budget: runoff, drainage and evapotranspiration. Contrarily to LAI and SSM, this is an indirect impact since these variables are not assimilated. Even though the impact of the assimilation is much lower than for LAI and SSM, it is not negligible and can reach 10 % or more in some regions. Globally, the assimilation increases the runoff and drainage and decreases the evapotranspiration, especially in the tropical regions.

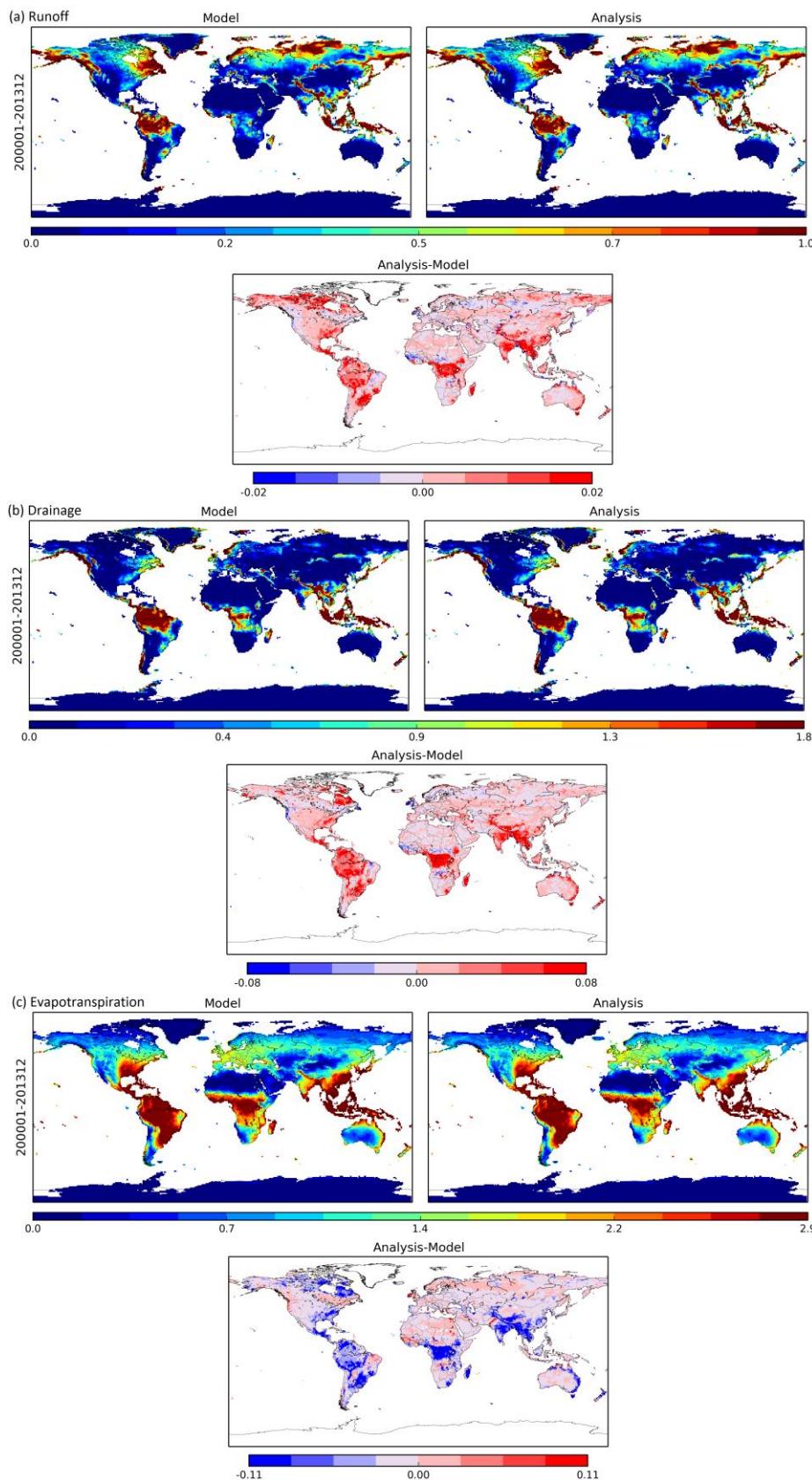


Figure 4-45: Average over the whole period of Runoff, Drainage and Evapotranspiration for the open-loop and analysis experiments, and impact of the assimilation. Units are kg/m²/day.

4.7 W3

4.7.1 Model description

The World-Wide Water (W3) model (formerly W3RA) is a grid-based global hydrological model based on the AWRA-L series of models used operationally by the Australian Bureau of Meteorology for national-scale water assessments, water accounts and the daily-updated Australian Landscape Water Balance web site (<http://www.bom.gov.au/water/landscape/>). W3 version 2.0 builds on a global implementation of AWRA-L v5.0. The original AWRA-L model (version 0.5) is comprehensively described in a technical report (Van Dijk, 2010), while the changes in version 5.0 are described in (Viney et al., 2015). The conceptual model structure is illustrated in Figure 4-46.

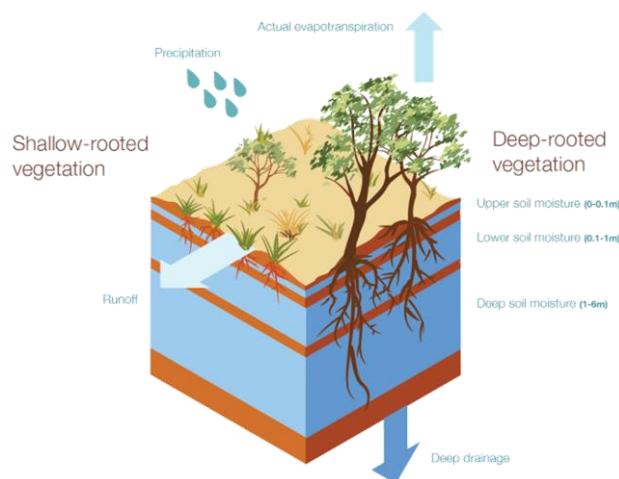


Figure 4-46: Simplified conception structure of W3 v2.0

For implementation at global scale, a number of modifications were made.

- A snow model was coupled to AWRA-L v5.0 based on HBV. This was also done in the W3RA version contributed to ensemble WRR1. However, for WRR2 parameters for the HBV model were obtained for HBV-SIMREG for WRR1 (Beck et al., 2016).
- The model was run at 0.05° resolution. As the data to be supplied to WRR2 was requested at not more than 0.25° , a distinction between subgrid vegetation response units (i.e., deep- or shallow-rooted vegetation) was considered unnecessary. Instead, each grid cell was assigned to either class based on available vegetation mapping.
- The WRR2 forcing data were downscaled to 0.05° for precipitation and temperature. This was done using a monthly climatology of average precipitation gain (a multiplier) and temperature offset (in $^\circ\text{C}$), based on the WorldClim (Hijmans et al., 2005) data set.
- Compared to the WRR1 contribution (which was based on AWRA-L v0.5), W3 version 2.0 has considerable changes in the description of soil and groundwater hydraulics (Viney et al., 2015).

- A global high-resolution map of Height Above Nearest Drainage (HAND) was used for sub-grid topography parameterisation, used in the simulation of saturated area and vegetation root access to groundwater. The map was derived primarily from the SRTM DEM completed with GTOPO30 at higher latitudes (G. Donchyts, Deltares, pers. comm).
- The radiation and energy balance description were slightly modified, which improved simulations particularly in high latitudes.
- Land surface temperature was simulated as a diagnostic, which allowed the assimilation of MODIS land surface temperature to estimate water use from sources other than rainfall (e.g. irrigation, wetlands and water bodies).

Other changes related mainly to the input fields, which included:

- ocean mask (0.01° NOAA);
- slope (SRTM/GTOPO30);
- aquifer hydraulic properties (GLHYMPS, Gleeson et al., 2014);
- groundwater recession coefficient (GSCD, Beck et al., 2016);
- soil hydraulic properties (GSDE , Shangguan et al., 2014);
- permanent water, permanent ice, impervious areas (GlobCover);
- broad vegetation type (MOD12C1);
- vegetation height (ICESat-GLAS, Simard);

4.7.2 Data assimilation

For the data assimilation (online) run, the following MODIS-derived 0.05° resolution data products were assimilated:

- white- and black-sky albedo (MCD43C3.005);
- BRDF reflectance (MCD43C4.005);
- inundation fraction (Van Dijk et al., 2016);
- leaf area index (GLASS);
- snow cover (MCD43C3.005 / MCD43C4.005);
- land surface temperature (MOD11C1.006);

The data assimilation approach was an efficient as well as somewhat crude updating using either nudging or direct insertion. Dynamic updating was done for 9 parameters and 3 (non-mass) states. Because water store were not directly updated, water balance is maintained even in the data assimilation run.

The evaporation of water not derived from rainfall is estimated through the assimilation of LST and inundation observations (Figure 4-47). These estimates are not reconciled within the model water budget and should be seen as separate and in addition to rainfall driven ET. They are however considered in estimating net river discharge, where they are routed as a loss that goes to the detriment of runoff gains.

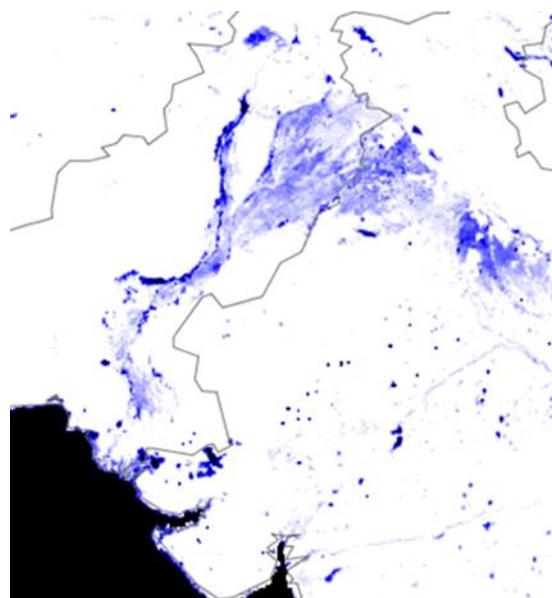


Figure 4-47: Example of additional evaporation from non-rainfall sources estimated for 03 July 2012 over Pakistan and NW India. Values range from 0 (white) to 5 mm/d (blue).

4.7.3 Evaluation

Model simulations were evaluated against streamflow from around 9,000 stations worldwide (cf. Van Dijk et al., 2016), as well as observations the in situ global soil moisture data base and daytime-average estimates of radiation budget components as well as latent and sensible heat flux from the FLUXNET network, made available through the La Thuile data set. These showed satisfactory agreement between model and observations. It is noted that the additional evaporation estimates derived from LST assimilation have not been evaluated against independent data sources and therefore should be considered experimental. A more comprehensive evaluation will be documented in forthcoming publications.

4.8 WaterGAP3

4.8.1 Model description

The global water model WaterGAP3 (Water – Global Assessment and Prognosis) is a grid-based, integrative assessment tool to examine the state of global freshwater resources. The model framework consists of a spatially-distributed rainfall-runoff model, five sectorial water use models, and a large-scale water quality model.

The global hydrological model simulates the terrestrial part of the global hydrological cycle by a sequence of storage equations for the most relevant continental storage compartments: canopy, snowpack, soil, renewable groundwater, and surface water bodies. The model requires daily fields of precipitation, near-surface air temperature, downwelling shortwave and longwave radiation as external meteorological forcing. Potential evapotranspiration is estimated according to the Priestley-Taylor approach

with surface net radiation calculated on the basis of land-cover dependent albedo and emissivity values.

The canopy storage is conceptualized as a single layer that intercepts precipitation until the maximum storage capacity is exceeded, and intercepted water evaporates at potential rate. Maximum canopy storage depends on daily LAI which is modelled as a function of land use dependent maximum LAI, fraction of deciduous plants and climate.

Snow accumulation and melt are simulated on a 1 arc minute sub-grid following a degree-day approach based on land-cover specific melting rates. Surface air temperature is disaggregated from 5 to 1 arc minute according to the elevation difference between each 5 arc minute cell and its corresponding sub-grid cells.

The soil column is represented as a single layer whose storage capacity is determined as a function of the soil texture dependent available water capacity (derived from soil maps) and the land-cover specific rooting depth. Total runoff from land is a function of effective precipitation, soil saturation, a non-linearity parameter and the fraction of urban area. The remaining part (effective precipitation minus runoff from land) is passed as infiltration to the soil storage. Groundwater recharge is calculated as a function of slope, soil texture, aquifer type, and the occurrence of permafrost/glaciers. Renewable groundwater resources are represented as a linear storage whose outflow is passed as subsurface runoff to the routing scheme.

Surface and subsurface runoff generated in each grid cell and inflow from upstream cells is transported through a series of linear and nonlinear retention storages representing lakes, reservoirs, and wetlands before contributing to streamflow. Flow velocity in the river segment is calculated as a function of river bed roughness, river bed slope and hydraulic radius of the channel according to the Manning-Strickler equation (Verzano et al., 2012). Lateral flow, i.e. between grid cells, is assumed to occur as streamflow only.

WaterGAP3 is calibrated in a basin-specific manner against long-term mean discharge by adjusting a runoff-nonlinearity parameter in each basin. In basins where adjusting this parameter does not provide an acceptable runoff estimate, i.e. the deviation from observed discharge remains larger than $\pm 1\%$, an additional runoff correction factor is assigned to each cell within the basin (same value for each cell). This correction factor effects all water balance components; if simulated discharge is too low, runoff will be increased and evaporation will be reduced and vice versa. This procedure is aimed to ensure a closed water balance, hence mass conservation, within the individual basin.

WaterGAP3 is explicitly designed to account for human interference on the natural streamflow regime through water abstraction and flow regulation by large dams and reservoirs. Spatially explicit time series of water withdrawal and water consumption are provided by the WaterGAP3 water use models (aus der Beek et al., 2010; Flörke et al., 2013) for five sectors: Domestic use (households and small businesses), manufacturing industries, thermal power plant cooling, irrigated agriculture and livestock farming. Sectorial water demands can be abstracted from surface water (rivers, reservoirs and lakes) and groundwater resources. Water demands for thermoelectric power plant cooling and livestock farming are assumed to be exclusively abstracted from surface water. For the remaining sectors, water withdrawals are allocated to groundwater and surface water abstractions according

to sector- and cell-specific temporally constant groundwater use fractions derived from national and sub-national statistics (Döll et al., 2012).

The WaterGAP3 simulations were carried out at 5 by 5 arc minute horizontal spacing. Storage compartments were initialized by re-running the first year of available meteorological forcing (1979) ten times. In a final step, simulated daily states and fluxes were scaled from the native 5 arc minute model resolution to 0.25° WRR-tier 2 resolution through taking the maximum (river discharge) or thorough area-weighted arithmetic averaging and summation, respectively.

4.8.2 Model improvements with respect to WRR1

Discharges at the major rivers are regulated to use the available river waters for hydropower generation, irrigation, water supply, flood control, and other aims. The extensive surface water usage affects discharges of the global major rivers. Hydrological models fail to reproduce the hydrographs of large rivers if they don't consider reservoir management. By upgrading WaterGAP3.1 (used in WRR1) to WaterGAP3.2 (used in WRR2) the human impact on river discharges are now much better resolved: an improved prospective management scheme defines the target storage in a monthly interval based on an objective function for one of the following reservoir purposes (for simplicity each reservoir is assumed to have only a single purpose):

- Flood control: flood damages shall be minimised.
- Hydropower: hydropower generation shall be maximised.
- Navigation: deviations from constant flow shall be minimised.

There are constraints for the management scheme: the reservoir capacity cannot be exceeded. Reservoir content at the end of each time interval is below the sum of the storage at the beginning of each time step plus reservoir inflow. Sufficient storage should be kept to maintain a minimum flow for 30 days. Sufficient storage capacity should be reserved to absorb larger floods. Finally, for flood protection, the release should be smaller than bankfull flow.

The daily release is calculated based on the difference between target and actual storage, the predicted inflow, and the number of remaining days. The new management scheme was tested and evaluated with different EO data (see next section).

4.8.3 Evaluation

The evaluation of the WaterGAP's improved WRR calculations is performed in four aspects: (i) The global average water balance components of WRR1 and WRR2 are compared and improvements are evaluated; (ii) the change in spatial distribution of different water balance components is discussed; (iii) the new implemented reservoir management representation in WaterGAP3.2 is evaluated with respect to different EO data products at four large reservoirs; (iv) WaterGAP's improved world water resources recalculation (especially river discharge) is discussed with respect to changed model representation and the new meteorological forcing.

(a) The global average water balance

The global distribution of major water balance components is primarily affected by the new spatial resolution and the new WRR2 forcing. The model improvements focused on the reservoir management scheme. Thus the model improvements affect primarily the water stored in the reservoirs (see next section) and not the surface runoff, evapotranspiration, and other surface fluxes. Hence, for the evaluation of the improvement of the water cycle representation between WRR1 and WRR2 Table 4-13 provides some basic data: the main difference between the calculated water balances in WRR1 and WRR2 are caused by the higher global average precipitation in the new MSWEP global gridded precipitation of about 896 mm/yr in WRR2 (863 mm/yr in WRR1). This increase of +33 mm is primarily transferred into a higher Runoff. In a global view this indicates that the average global evaporation, which increases by only 7 mm, is more sensitive to atmospheric conditions (air temperature and radiation) than to local water stress and precipitation.

The total increase of Runoff + Evaporation (+36 mm) corresponds to the increase of Precipitation and the water balance is closed. The change in soil and other storages decreased from +8 to +5mm per year between WRR1 and WRR2 and during the period 1979-2014.

Table 4-13: Global average water balance calculated by WaterGAP3.1 (WRR1) and WaterGAP3.2 (WRR2)

	P [mm/yr]	R [mm/yr]	ET [mm/yr]	ΔS [mm/yr]
WaterGAP3.1 WRR1	863	306	549	+8
WaterGAP3.2 WRR2	896	335	556	+5

(b) The global spatial distribution of major water balance components

Precipitation: The major change between the WRR1 and WRR2 forcing is the new precipitation data (see section 3.1.1, the mean annual MSWEP precipitation on the WRR2 grid is shown in Figure 4-48a). Beside several changes, the most distinct difference to the WRR1 precipitation is in areas of high precipitation such as the Amazonas basin (Figure 4-48b). E.g. a WRR1 precipitation hotspot in the Kongo area is now levelled to its surrounding area.

Evapotranspiration: The WRR2 WaterGAP3.2 calculations show that most water is evaporated in the tropic zone (Figure 4-49a). The results confirm features such as a relative high evapotranspiration in the eastern part of the U.S. relative to other areas at the same latitudes. The differences/improvements with respect to the WRR1, besides the better spatial resolution, go along with the changes in precipitation (Figure 4-49b). E.g., higher precipitation in some areas of the amazon cause altered evapotranspiration in those areas.

Runoff: As it was shown for Evapotranspiration, the changes in Runoff between WRR1 and WRR2 are highly correlated to the changes in precipitation (Figure 4-50). The pattern of changes is very similar.

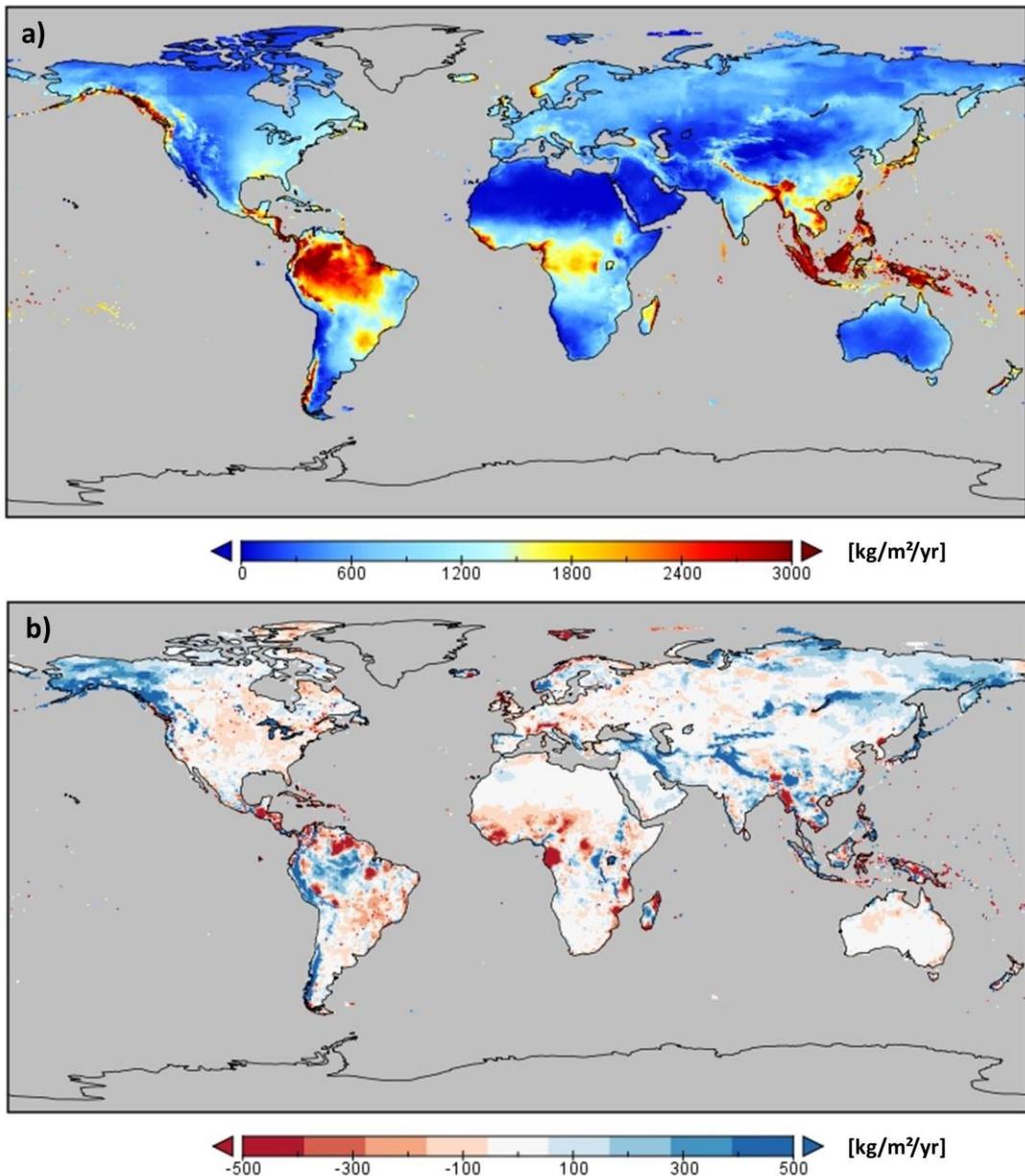


Figure 4-48: Global distribution of mean annual precipitation: a) WRR2 1979-2014 b) difference to WRR1 (WRR2-WRR1) 1979-2012.

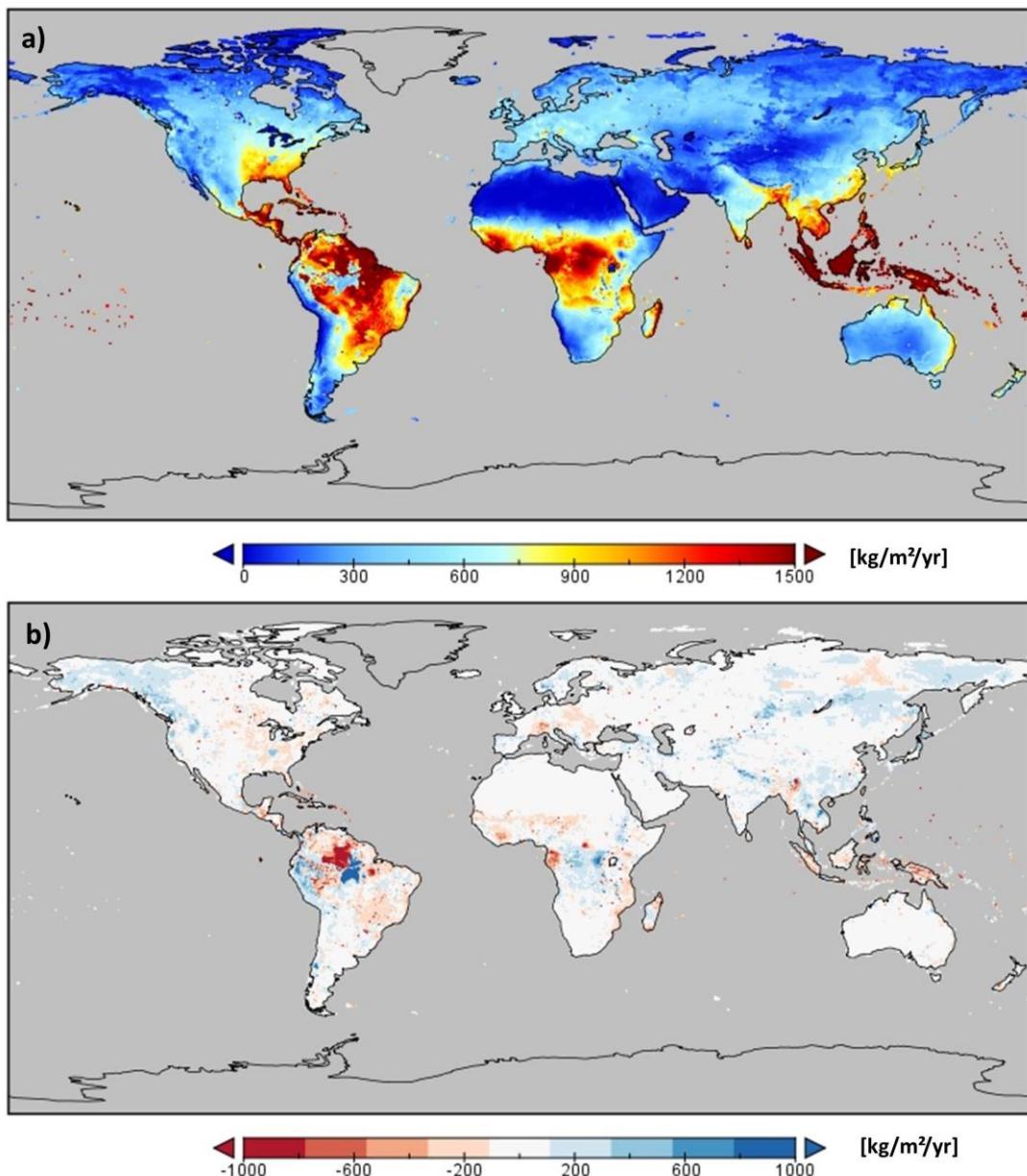


Figure 4-49: Global distribution of mean annual evapotranspiration: a) WRR2 1979-2014 b) difference to WRR1 (WRR2-WRR1) 1979-2012.

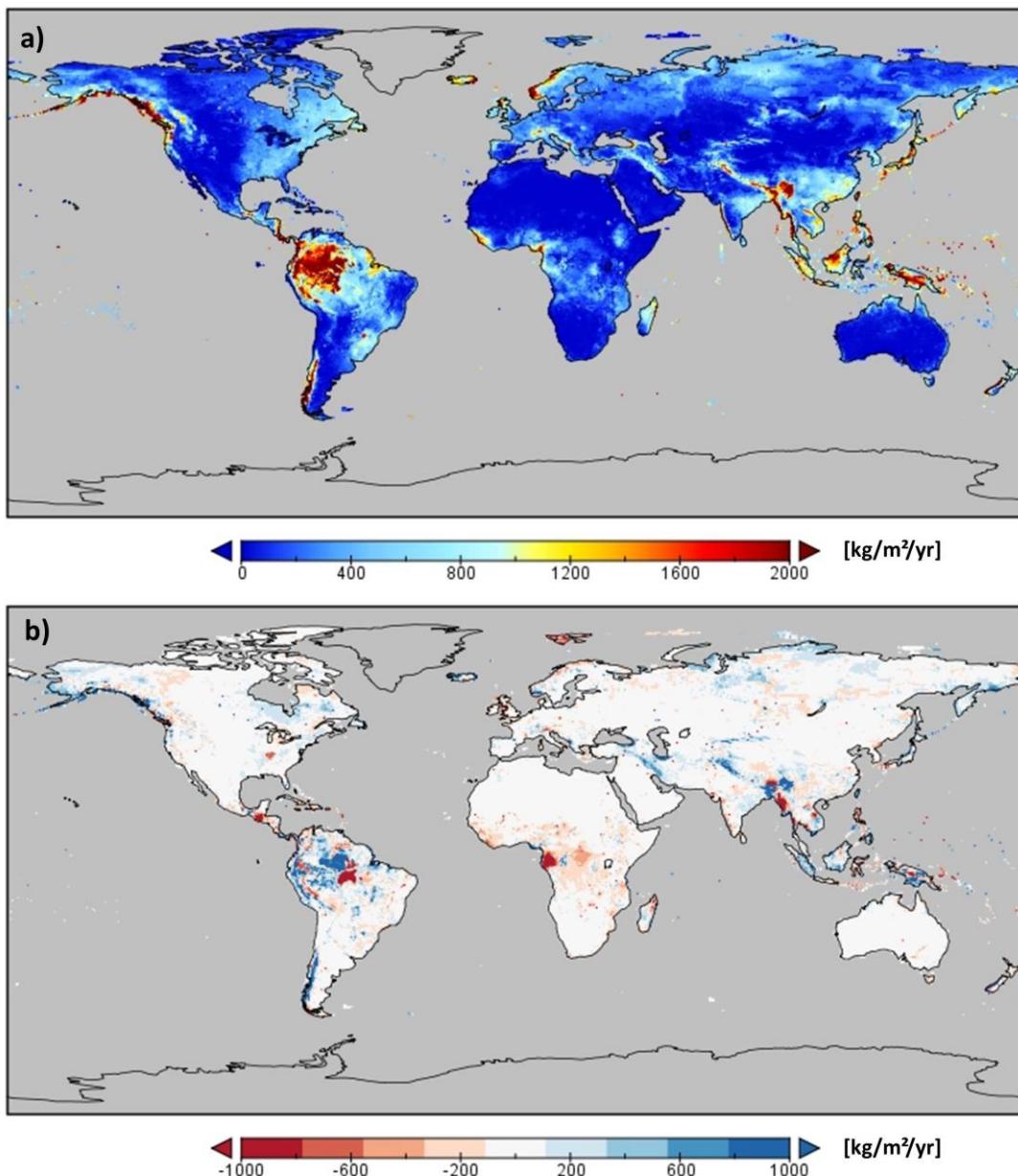


Figure 4-50: Global distribution of mean annual runoff: a) WRR2 1979-2014 b) difference to WRR1 (WRR2-WRR1) 1979-2012.

(c) Testing of WaterGAP 3.2 reservoir management scheme with EO data

The major improvement beside the better meteorological forcing and the higher spatial resolution is the new implementation of a reservoir management scheme (see section 4.8.2). EO data of reservoir surface water level were used to evaluate the better representation of the impact of reservoir management on the global water cycle. However, on the one hand, the satellites provide water surface levels and no reservoir volumes, on the other hand WaterGAP3 provides reservoir water volume information but no surface water level. Both variables are physically linked by the reservoir's bathymetry. Hence, the evaluation of the models performance in terms of the temporal behavior can be performed by (i) visually compare the time series and, more informative, by (ii) calculating the correlation coefficient.

We used 4 different EO products (reservoir water level time series) to test and validate the modelled reservoir management:

- DAHITI data (Schwatke et al., 2015): multi-mission satellite altimetry
- TP satellite data (prepared and provided by project partner I-MAGE)
- Envisat EO data (prepared and provided by project partner I-MAGE)
- Jason2 EO data (prepared and provided by project partner I-MAGE)

These data were compared with WaterGAP3.2 calculations at four large reservoirs:

Cahora Bassa

The African Cahora Bassa Reservoir in Mozambique is a reservoir in the Zambezi river system ($15^{\circ}35'09''S$ $32^{\circ}42'17''E$). The surface area is about 2739 km^2 with a maximum water depth of 157m, and a total capacity of 55.8 km^3 .

Figure 4-51 compares the satellite measurements of the Cahora Bassa reservoir surface water level and the WaterGAP3.2 estimated reservoir volumes (monthly means). A correlation coefficient of about 0.71 indicates a high linear dependency between both data. The modeled annual cycle corresponds to the measured one. Especially the low water levels/volumes in 2005 are reproduced by the model very well.

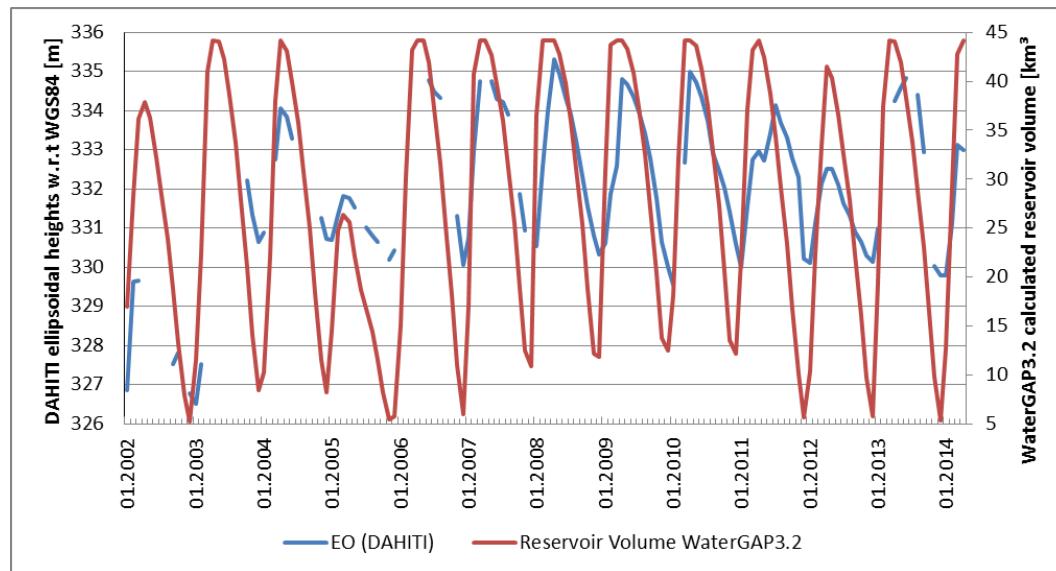


Figure 4-51: Calculated Water volumes and observed water levels of the Chora Bassa Reservoir, Mozambique, for the period 2002 - 2014.

Kariba Reservoir

The Kariba Reservoir in Zimbabwe and Zambia is another big reservoir in the Zambezi river system ($16^{\circ}55'S$ $28^{\circ}00'E$). The surface area is about 5580 km^2 , the maximum water depth is 97 m, and the total maximum capacity is 180 km^3 . The main purpose is hydropower generation.

Figure 4-52 compares the satellite measurements of the Kariba reservoir surface water level and the WaterGAP3.2 estimated reservoir volumes (monthly means). A correlation coefficient of about 0.85 indicates a high linear dependency between both

data. The modeled annual cycle corresponds to the measured one. Especially the low water levels/volumes in the period 1992-1997 and the following increase until 2000 are reproduced by the model very well.

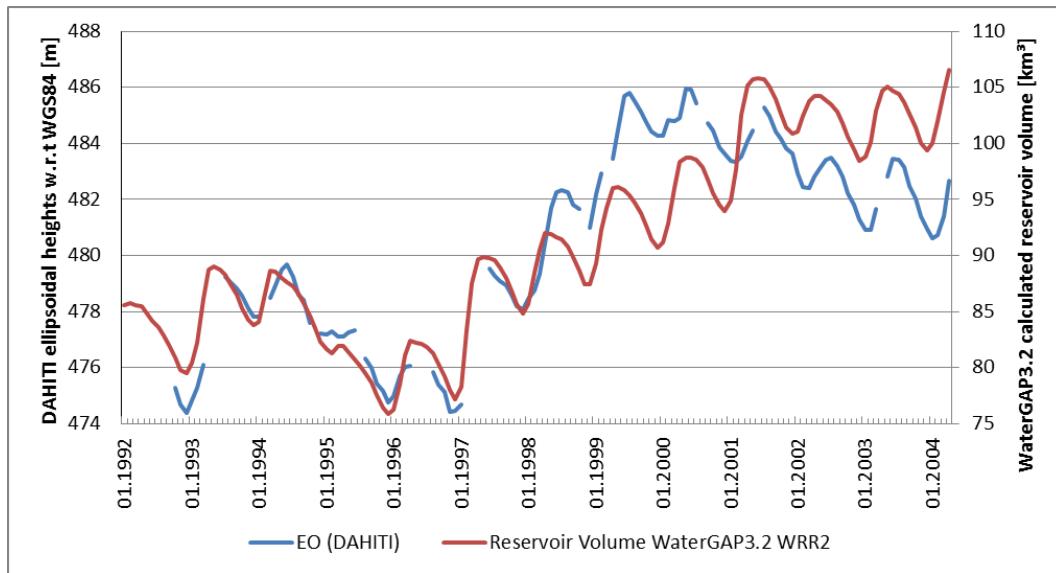


Figure 4-52: Calculated Water volumes and observed water levels of the Kariba Reservoir for the period 1992-2004.

Lake Mead (U.S.)

Lake Mead, formed by the Hoover Dam, is a large reservoir in the Colorado River catchment (36.25°N 114.39°W). The surface area is about 640 km^2 , the maximum water depth is 162 m.

Figure 4-53 compares the satellite measurements of the Lake Mead surface water level and the WaterGAP3.2 estimated reservoir volumes (monthly means). A correlation coefficient of about -0.54 indicates an inverse linear dependency between both data sources. The modeled annual cycle corresponds to the measured one. However, the long term trends are not calculated by the model. Only the decreasing water levels/volumes since 2013 are captured.

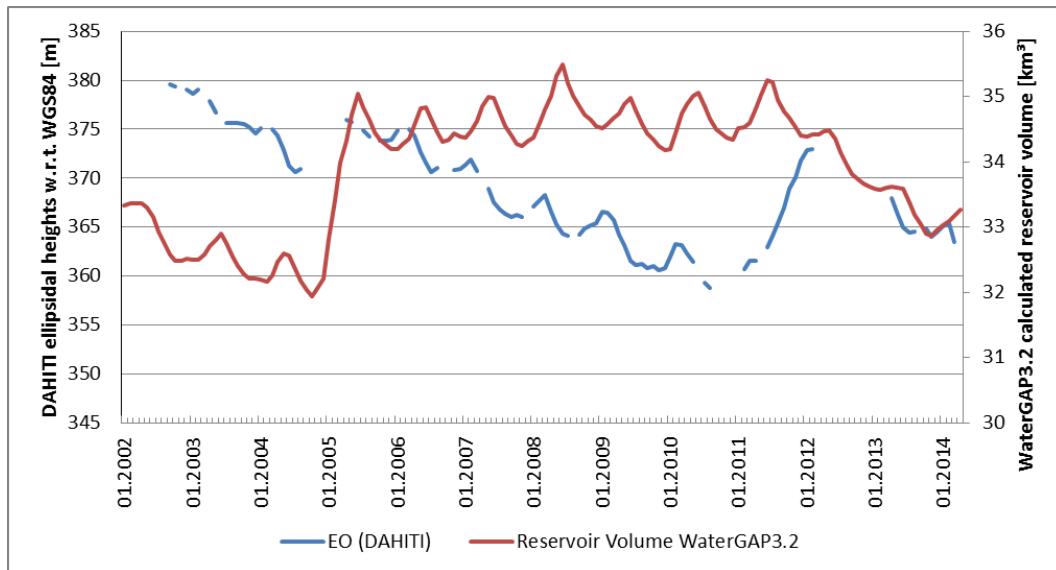


Figure 4-53: Calculated Water volumes and observed water levels of Lake Mead for the period 2002 to 2014.

Lake Victoria

This is Africa's largest lake by area (68800 km^2). It is a formally natural lake situated in the Nile basin. The water level is regulated since decades, thus it is an additional lake that can be used to test the reservoir management scheme.

Figure 4-54 compares the satellite measurements (three different products) of the Lake Victoria surface water level and the WaterGAP3.2 estimated reservoir volumes (monthly means). A correlation coefficient of about 0.59 between the TP satellite measurements and the model calculations indicates that the model captures the real reservoir management. The Envisat data and Jason data are shifted with respect to the TP satellite data and are not reliable due to the difference to DAHITI data (not shown in Figure 4-54). The fluctuations of the lake are different to the smaller reservoirs, e.g. there is no distinct annual cycle. The modeled temporal fluctuations correspond to the measured one.

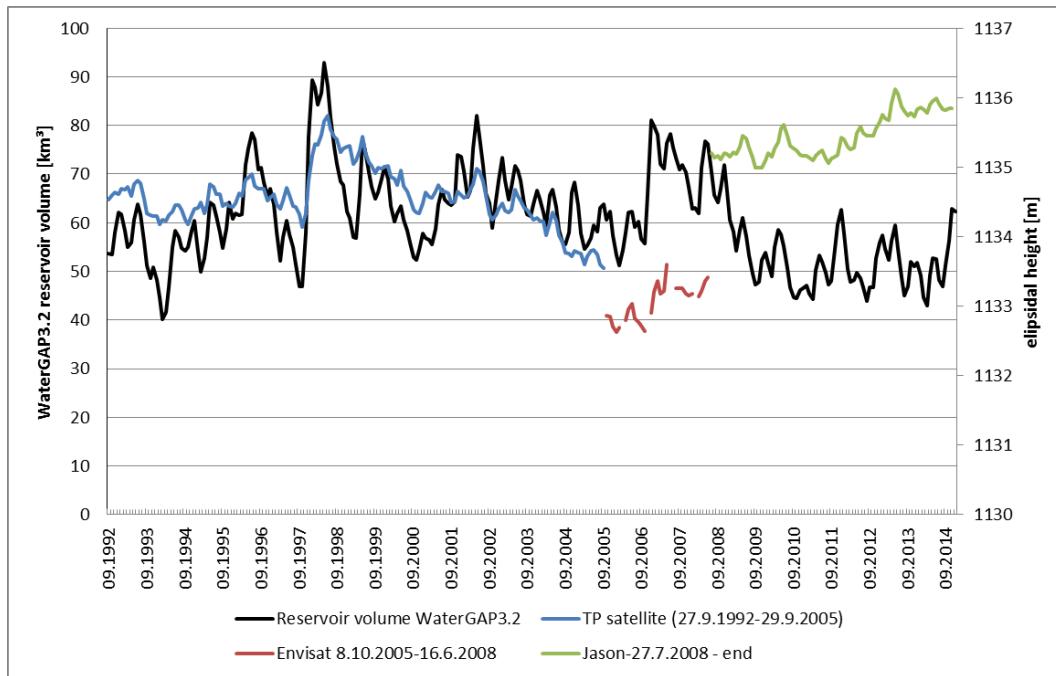


Figure 4-54: Calculated Water volumes and observed water levels of Lake Victoria for the period 1992 to 2014.

(d) Effects on the improvement on water availability and discharge

Because the model improvements affect primarily the reservoir algorithm, the Nile catchment was selected to highlight these effects on the modelled river discharges. There are numerous reservoirs in this catchment (e.g. Lake Victoria) which makes this basin an appropriate choice for such an evaluation. Figure 4-55 compares the calculated Nile River hydrographs at Al Akshas (GRDC station 1362100), situated approximately 30km upstream of Kairo. Some basic statistics are listed in Table 4-14 for WRR1, WRR2 and the WRR1_F2 setup. Here, WRR1_F2 is the old model forced with the meteorology of WWR2.

There is a clear difference between WRR1 and WRR2: on average, the Nile discharge is higher than in the WRR1 calculations and it is less variable. The latter indicates a higher impact of the management schemes on the discharge ("flattening" of the hydrograph) in the WRR2 calculations. However, this is not caused by the better representation of the reservoir management. But it is an effect of higher water availability: the WRR1 model, forced with the new meteorology, deviates only slightly from the WRR2 setup (compare green and red line in Figure 4-55). To conclude, there is less fluctuation on a higher discharge level in WRR2. This is primarily because of the new forcing. The effects of the new precipitation data overtop the effects of the better representation of reservoir management.

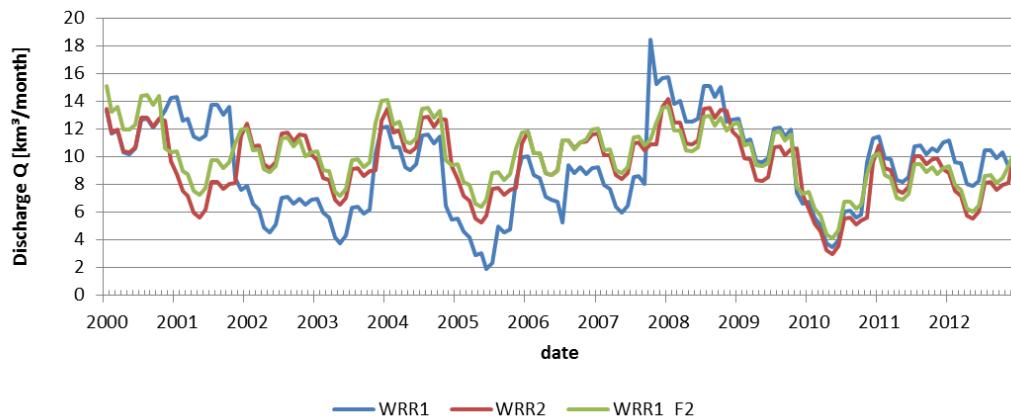


Figure 4-55: Modelled Nile River Discharges at Al Akhsas gauging station (GRDC ID: 1362100) for the period 2000-2010.

Table 4-14: The WaterGAP3 modelled Nile discharges statistics (2000-2012) of different WRR settings.

	WRR1	WRR2	WRR1_F2
Average monthly discharge [km ³ /month]	9.2	9.5	10.0
Variance of the monthly discharge [(km ³ /month) ²]	10.7	5.8	5.0

4.9 WaterWorld

The WWR2 data used by the WaterWorld model and the model changes being reported here include (a) the use of downscaled WFDEI (WATCH Forcing Data methodology applied to ERA-Interim data) data in the Estonia Case Study supplied at 1Km spatial resolution by our Partners at Deltares, (b) the use of downscaled TRMM data in the Magdalena, Colombia Case Study, (c) the improvements made to the WaterWorld (WW) model to allow better Policy decision making by users of the EO products uploaded to WW and (d) adapting of WW from 1Km scale to 10Km scale for global application.

The WaterWorld model (WW) is parameterised with 140 inputs from an extensive global gridded database of more than 600 variables, developed from many sources, since 1998, the so-called simTerra database. All of these data are available globally at 1km resolution and some variables (terrain, land cover, urban areas, water bodies) are available globally at 1ha resolution. WaterWorld requires some 140 maps to run including monthly climate data, land cover and use, terrain, population, water bodies and more.

Whilst publically-available terrain and land cover data are now well developed for local scale application, climate and land use data remain a challenge. The global climate data products made available by WRR2 remain at 0.25 degree resolution or worse, which is rather coarse for local application. Our options for new and improved WRR2 datasets to integrate into WW are limited to the 1Km spatial data required by the WW model. We have used WFDEI downscaled to 1km resolution to provide an alternative

input to WaterWorld's preloaded climate data. Downscaling was performed by Deltares.

We have then examined the impacts of the new data on key hydrological outputs of water balance and water quality for the entire catchment of Estonia at 1Km resolution. Estonia is a Case Study area in Earth2Observe. In D5.3 we discuss the remaining challenges of using datasets like these for local scale application and decision making.

In addition to WFDEI data, the WW Model has been parameterised with other EO datasets. As reported in D6.2, WorldClim rainfall data was compared to an alternative high resolution global rainfall database: TRMM 2b31 (Tropical Rainfall Measurement Mission, mean of all rainfall observations to March 2006). The data was originally 5km resolution, but has been downscaled to 1km for upload into WW. Approximately 50,000 satellite swaths (0.5 TB of data) were processed and the results are shown below.

Rainfall is broadly similar between the two products, though the TRMM data has a higher catchment total rainfall. Upon comparing WaterWorld results , TRMM data produced a higher water balance, but a lower erosion rate. The latter is due to spatial pattern of TRMM rainfall in relation to the distribution of bare ground and slope in the catchment. The TRMM data has lower rainfall than WorldClim in areas dominated bare areas and steep slopes on mountain peaks (red) which may be erosion hotspots (see Figure 4-56 and Figure 4-57).

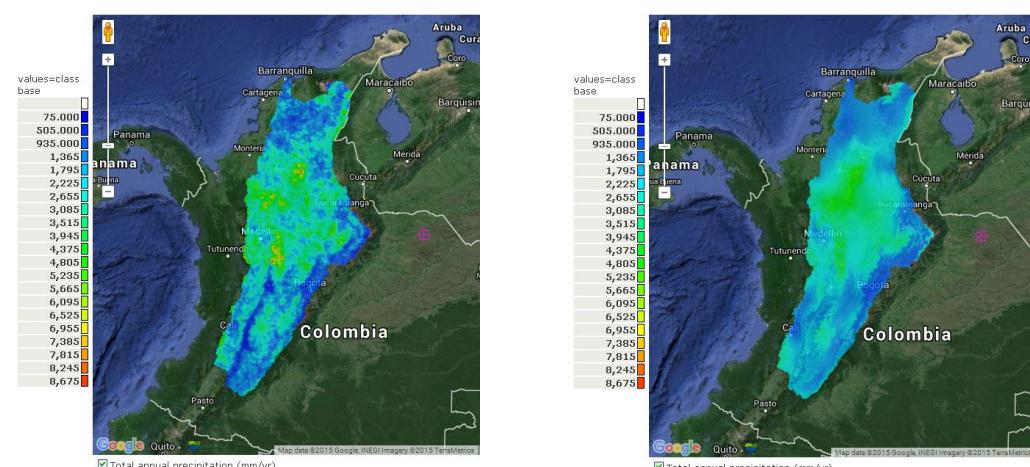


Figure 4-56: WorldClim (right) and TRMM (left) rainfall data in the Magdalena catchment.

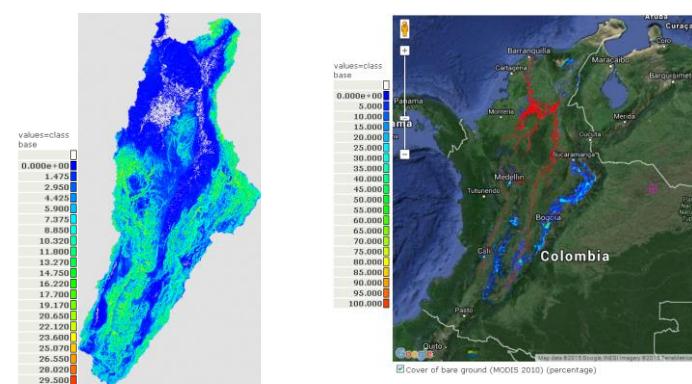


Figure 4-57: Slope angle (from Hydrosheds data) and areas of bare ground (from MODIS data) in the Magdalena catchment.

The WW policy support system itself has been extended in a number of ways. Users can now upload their own land use and climate data so that the impact of uploaded new land use climate scenarios can be analysed. WW users therefore have more options for testing the impact of their decisions. The new 2016 land use scenario options include:

- Business as Usual (eg same deforestation rate)
- Accelerated deforestation
- Percentage change of individual land classes
- Option to allocate new land use by agricultural suitability
- Include planned infrastructure / transport routes
- Vary Management Effectiveness Index - for Protected Areas for example

These can then be used in the new ‘Stacked Scenario’ model function where users can combine a number of plausible scenarios for land use, climate, population, and assess the impacts of these multiple changes as well as each individual scenario on water resources. For example, users can assess the impact of a combination of climate change, deforestation, population growth and rural-urban migration, mining and oil and gas exploitation.

Finally, in order to incorporate coarse scale global datasets, work is in progress to adapt WW to 10Km resolution simulations globally. This is still in testing phase (v2.92). Figure 4-58 shows WaterWorld's index of water stress. Water Stress is the % of water demand that is unavailable or contaminated. The data are averaged by watershed.

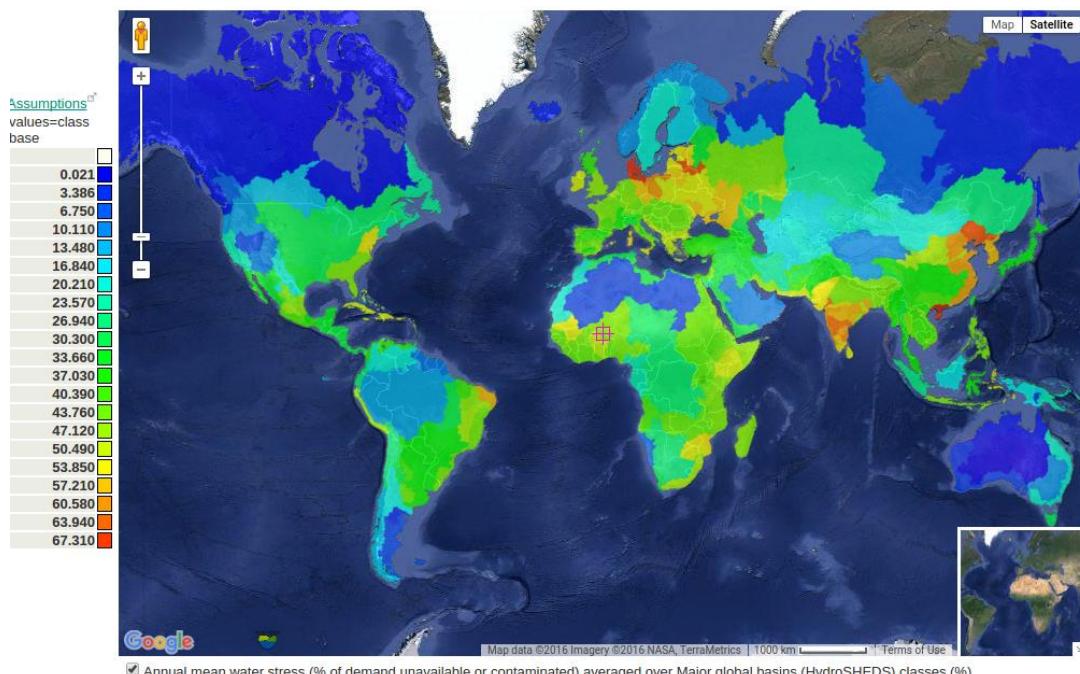


Figure 4-58: Water stress by global basin modelled at 10Km resolution.

5 Data

5.1 Data access

5.1.1 Meteorological forcing

The meteorological forcing used in WRR2 baseline is available on the WCI³ data server and the direct links listed in Table 5-1. The Rainfall products used in WRR2 ensemble are available on the ftp server and their links also listed in table Table 5-1

Table 5-1: Forcing data

Forcing	Location
WRR1	http://wci.earth2observe.eu/thredds/catalog/ecmwf/met_forcing_v0/catalog.html
WRR2 except Rainfall & Snowfall	http://wci.earth2observe.eu/thredds/catalog/ecmwf/met_forcing_v1/catalog.html
WRR2 MSWEP Rainfall & Snowfall	http://wci.earth2observe.eu/thredds/catalog/jrc/MSWEP/3hourly_e2o_netcdf_convention/catalog.html http://wci.earth2observe.eu/thredds/catalog/jrc/MSWEP/daily_e2o_netcdf_convention/catalog.html
WRR2 TRMM Rainfall	ftp://wci.earth2observe.eu/exchange/e2o_ecmwf/forcing_wrr2ens/rainf_trmm/
WRR2 TRMMRT Rainfall	ftp://wci.earth2observe.eu/exchange/e2o_ecmwf/forcing_wrr2ens/rainf_trmmrt/
WRR2 GSMAP Rainfall	ftp://wci.earth2observe.eu/exchange/e2o_ecmwf/forcing_wrr2ens/rainf_gsmap/
WRR2 CMORPH	ftp://wci.earth2observe.eu/exchange/e2o_ecmwf/forcing_wrr2ens/rainf_cmorph/

5.1.2 Model output convention

All models used the same grid for the post-processing (equal to the meteorological forcing), for both the baseline and ensemble simulations with the exception of the data assimilations experiments and river discharge of SURFEX-TRIP. The simulations are available on the server with a daily frequency, and their location is shown in Table 5-5. For the models with a sub-daily time-step, the daily means are computed from all the model time steps. The monthly means are computed from the daily means.

³ <https://wci.earth2observe.eu/>

As in WRR1, the file format is netcdf4⁴ containing a single variable following the convention:

e2o_[ID]_[VER]_[DOMAIN]_[FREQ]_[VAR]_[YEAR_START-YEAR_END].nc

1. **ID**: Institution identifier (see Table 5-2);
2. **VER**: experiment names (see Table 5-3);
3. **DOMAIN**: domain identification (see Table 5-3);
4. **FREQ**: output frequency: **day** (daily means) **mon** (monthly means);
5. **VAR**: variable name contained in the file (see Table 5-4);
6. **YEART_START-YEAR_END**: start and end year denoting the temporal extension (see Table 5-3).

Table 5-2: Institution identifiers in the model output

Institution ID	Institution (mode)
ecmwf	ECMWF (HTESSEL, see 4.1)
univu	University of Utrecht (PCR-GLOBWB, see 4.5)
metfr	Meteo-France (SURFEX-TRIP, see 4.6)
nerc	NERC, CEH (JULES, see 4.2)
jrc	JRC (LISFLOOD, 4.3)
cnrs	CNRS (ORCHIDEE, see 4.4)
univk	University of Kassel (WaterGAP3, see 4.8)
Anu	Australian National University (W3, see 4.7) (csiro in wrr1)
ambio	Ambiotek (WaterWorld, see 4.9)

Table 5-3: Experiment version names, associated start and end year and domain.

Experiment name	YEAR_START - YEAR_END	Domain	Description
wrr1	1979-2012	glob30	WRR1 simulations
wrr2	1980-2014 (monthly files) 1980-1989 (daily 1 st decade) 1990-1999 (daily 2 nd decade) 2000-2014(daily 3 rd decade)	glob15 ^a	WRR2 simulations
wrr2gsmap	2000-2013	glob15 ^a	WRR2 ensemble forced with GSMAP rainfall

⁴ <https://www.unidata.ucar.edu/software/netcdf/>

Experiment name	YEAR_START – YEAR_END	Domain	Description
wrr2trmm	2000-2013	glob15 ^a	WRR2 ensemble forced with TRMM rainfall
wrr2cmorph	2000-2013	glob15 ^a	WRR2 ensemble forced with CMORPH rainfall
wrr2trmmrt	2000-2013	glob15 ^a	WRR2 ensemble forced with TRMMRT rainfall
wrr2da	2000-2013	glob15 ^{a,b}	WRR2 with data assimilation

^a – glob30 for the river routing related variables for metfr simulations

^b – glob60 for the metfr simulations

Table 5-4: List of output variables and conventions

Name	long_name attribute	standard_name attribute	units attribute	Definition	Positive direction
Precip ^a	total precipitation	precipitation_flux	kg m ⁻² s ⁻¹	Average of total precipitation (Rainf+Snowf)	downwards
Evap ^a	total evapotranspiration	water_evaporation_flux	kg m ⁻² s ⁻¹	Sum of all evaporation sources, averaged over a grid cell	downwards
Runoff ^a	Total runoff	runoff_flux	kg m ⁻² s ⁻¹	Average total liquid water draining from land	into grid cell
Rainf	rainfall	rainfall_flux	kg m ⁻² s ⁻¹	Average of the total rainfall (liquid phase)	downwards
Qsa ^a	surface runoff	surface_runoff_flux	kg m ⁻² s ⁻¹	Runoff from the land surface and/or subsurface stormflow	into gridcell
Qsb ^a	Subsurface runoff	subsurface_runoff_flux	kg m ⁻² s ⁻¹	Gravity drainage and/or slow response lateral flow. Ground water recharge will have the opposite sign.	into gridcell
Qrec	Recharge	N.A	kg m ⁻² s ⁻¹	Recharge from river to the flood plain and from groundwater to soil column	into of gridcell
Qsm	snowmelt	surface_snow_melt_flux	kg m ⁻² s ⁻¹	Average liquid water generated from solid to liquid phase change in the snow	solid to liquid

Name	long_name attribute	standard_name attribute	units attribute	Definition	Positive direction
PotEvap	potential evapotranspiration	water_potential_evaporation_flux	kg m ⁻² s ⁻¹	The flux as computed for evapotranspiration but will all resistances set to zero, except the aerodynamic resistance.	downwards
ECanop	interception evaporation	water_evaporation_flux_from_canopy	kg m ⁻² s ⁻¹	Evaporation from canopy interception, averaged over all vegetation types within a grid cell.	downwards
TVeg	vegetation transpiration	transpiration_flux	kg m ⁻² s ⁻¹	Vegetation transpiration, averaged over all vegetation types within a grid cell.	downwards
ESoil	bare soil evaporation	water_evaporation_flux_from_soil	kg m ⁻² s ⁻¹	Evaporation from bare soil.	downwards
EWater	Open water evaporation	N.A	kg m ⁻² s ⁻¹	Evaporation from surface water storage (lakes, river Chanel, floodplains, etc.)	downwards
RivOut	river discharge	N.A	m ³ s ⁻¹	Water volume leaving the cell	downstream
Dis	point river discharge	N.A	m ³ s ⁻¹	Water volume leaving the cell	downstream
SWnet	Net shortwave radiation	surface_net_downward_shortwave_flux	W m ⁻²	Incoming solar radiation less the simulated outgoing shortwave radiation, averaged over a grid cell	downward
LWnet	Net longwave radiation	surface_net_downward_longwave_flux	W m ⁻²	Incident longwave radiation less the simulated outgoing longwave radiation, averaged over a grid cell	downward
Qle	Latent heat flux	surface_downward_latent_heat_flux	W m ⁻²	Energy of evaporation, averaged over a grid cell	downward
Qh	Sensible heat flux	surface_downward_sensible_heat_flux	W m ⁻²	Sensible energy, averaged over a grid cell	downward
AvgSurfT	Average surface temperature	surface_temperature	K	Average of all vegetation, bare soil and snow skin temperatures	-

Name	long_name attribute	standard_name attribute	units attribute	Definition	Positive direction
Albedo	Surface Albedo	surface_albedo	-	Grid cell average albedo for all wavelengths.	-
LAI	Surface Albedo	leaf_area_index	-	Grid cell average leaf area index	-
SWE ^a	Snow water equivalent	liquid_water_content_of_surface_snow	Kg m ⁻²	Total water mass of the snowpack (liquid or frozen), averaged over a grid cell (including SWEVeg)	-
CanopInt	Total canopy water storage	N.A	Kg m ⁻²	Total canopy interception, averaged over all vegetation types within a grid cell (included both solid and liquid)	-
SWEVeg	SWE intercepted by the vegetation	N.A	Kg m ⁻²	Total water mass of the snowpack (liquid or frozen), averaged over a grid cell and intercepted by the canopy	-
SurfStor	Surface Water Storage	N.A	Kg m ⁻²	Total liquid water storage, other than soil, snow or interception storage (i.e. lakes, river channel or depression storage).	-
WaterTableD	Water table Depth	N.A	M	Depth of the water table (distance from surface)	-
SnowFrac	Snow covered fraction	surface_snow_area_fraction	-	Grid cell snow covered fraction	-
SnowDepth	Depth of snow layer	surface_snow_thickness	M	total snow depth	-
SurfMoist ^a	Surface soil moisture	N.A	Kg m ⁻²	best of 5 cm depth or first model layer	-
RootMoist ^a	Root zone soil moisture	N.A	Kg m ⁻²	Total soil moisture available for evapotranspiration (or up to 1 meter depth if not defined)	-
TotMoist	Total soil moisture	N.A	Kg m ⁻²	Vertically integrated total soil moisture	-

Name	long_name attribute	standard_name attribute	units attribute	Definition	Positive direction
GroundMoi st	ground water	N.A	Kg m ⁻²	ground water not directly available for evapotranspiration	-
lsm	Land sea mask	land_area_fraction	-	fraction of land in each cell	
SurfSoilSal	Surface soil saturation	N.A	Kg m ⁻²	saturation soil moisture for the layer reported in SurfMoist	
RootSoilSat	Root soil saturation	N.A.	Kg m ⁻²	saturation soil moisture for the layer reported in RootMoist	
TotSoilSat	Total soil saturation	N.A	Kg m ⁻²	saturation soil moisture for the layer reported in TotMoist	

^a Default variables available daily on the WRR2 ensemble.

Table 5-5: Model output locations

Model	Direct link		Comments
	WRR1 / WRR2	WRR2-ENS	
HTESEL	wrr1 wrr2	wrr2trmm wrr2trmmrt wrr2gsmap wrr2cmorph wrr2da	
JULES	wrr1 wrr2	wrr2trmm wrr2trmmrt wrr2gsmap wrr2cmorph	
LISFLOOD	wrr1 wrr2	wrr2trmm wrr2trmmrt wrr2gsmap wrr2cmorph	Currently the wrr2-ens only contains the River discharge data. Remaining data will be made available.
ORCHIDEE	wrr1 wrr2	wrr2trmm wrr2trmmrt wrr2gsmap wrr2cmorph	wrr2-ens simulation are not available yet due to technical problems, and will be made available as soon as possible

Model	Direct link		Comments
	WRR1 / WRR2	WRR2-ENS	
PCR-GLOBWB	wrr1 wrr2	wrr2trmm wrr2trmmrt wrr2gsmap wrr2cmorph wrr2da	wrr2 data to be made available soon. Wrr2-ens data not available due to technical problems
SURFEX	wrr1 wrr2	wrr2trmm wrr2trmmrt wrr2gsmap wrr2cmorph wrr2da	wrr2da is available at 1°x1° resolution (glob60) for all point variables and at 0.5°x0.5° for the river routing variables (glob30)
WaterGAP3	wrr1 wrr2	wrr2trmm wrr2trmmrt wrr2gsmap wrr2cmorph	
W3	wrr1	wrr2da	In wrr1 the model institution id was "csiro" and in wrr2da is "anu". wrr2da is available from 2000 to 2014.
SWBM	wrr1		Did not participated in WRR2

6 Summary

This report describes the tier-2 water resources reanalysis (WRR2) development, implementation and data structure and storage. This work was built on the initial WRR1 dataset by exploring several key sources of improvement. The main achievements include:

- A new high resolution 0.25x0.25 global meteorological forcing with a 3-hourly frequency covering 1979 to 2014. It is based on ERA-Interim with a topographic temperature adjustment and a new precipitation dataset that integrates atmospheric reanalysis, satellite products and in-situ observations;
- A variety of modelling system enhancements including process description (e.g. snowpack, runoff generation, groundwater representation, etc..), calibration and water use (withdrawal, consumption, management).
- Different data assimilation methodologies and use of Earth observations (e.g. snow depth, soil moisture, river discharge, vegetation state, etc...)
- Ensemble of simulation using four different satellite precipitation products.

This report focused on the technical details of the simulations and the description of the modelling system. Each modelling system presented an evaluation of their developments along with the impact of the meteorological forcing. Due to the different nature of each modelling and data assimilation system changes, a consistent evaluation among models was not pursued in this report. On the contrary, each modelling system was evaluated independently, and this report serves as a detailed description of the dataset and not an integrated quality assessment of its quality. Such analysis will be presented in following project reports. However, among the several evaluation examples there is a clear signal of the improvements due to the new meteorological forcing used in WRR2. This is mainly attributed to the new precipitation datasets (MSWEP) which leads to clear improvements of the models simulations. Each modelling system change also shows improvements, but these tend to be restricted to the main process that was addressed. Overall, it is expected that WRR2 will bring a significant improvement in terms of the land water cycle representation over land when compared with WRR1. This new dataset will have an important impact on the community and users due to its higher resolution and state-of-the art modelling systems and data assimilation.

7 References

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration, FAO Irrig. Drain. Pap. 56, Food and Agric. Organ., Rome, 1998.

aus der Beek, T., Flörke, M., Lapola, D. M., Schaldach, R., Voß, F., and Teichert, E.: Modelling historical and current irrigation water demand on the continental scale: Europe, Adv. Geosci., 27, 79-85, 10.5194/adgeo-27-79-2010, 2010.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrol. Earth Syst. Sci., 19, 389-407, 10.5194/hess-19-389-2015, 2015.

Barbu, A. L., Calvet, J. C., Mahfouf, J. F., and Lafont, S.: Integrating ASCAT surface soil moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation application over France, Hydrol. Earth Syst. Sci., 18, 173-192, 10.5194/hess-18-173-2014, 2014.

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Miralles, D. G., McVicar, T. R., Schellekens, J., and Bruijnzeel, L. A.: Global-scale regionalization of hydrologic model parameters, Water Resources Research, 52, 3599-3622, 10.1002/2015WR018247, 2016.

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A.: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data, Hydrol. Earth Syst. Sci., 21, 589-615, 10.5194/hess-21-589-2017, 2017.

Bergström, S.: The HBV model in: Computer models of watershed hydrology, edited by: Singh, V. P., Water Resources Publications, Highlands Ranch, Colorado, U.S.A, 1995.

Brasnett, B.: A Global Analysis of Snow Depth for Numerical Weather Prediction, Journal of Applied Meteorology, 38, 726-740, 10.1175/1520-0450(1999)038<0726:agaosd>2.0.co;2, 1999.

Clark, D. B., and Gedney, N.: Representing the effects of subgrid variability of soil moisture on runoff generation in a land surface model, J. Geophys. Res. Atmos., 113, n/a-n/a, 10.1029/2007JD008940, 2008.

de Graaf, I. E. M., van Beek, L. P. H., Wada, Y., and Bierkens, M. F. P.: Dynamic attribution of global water demand to surface water and groundwater resources: Effects of abstractions and return flows on river discharges, Advances in Water Resources, 64, 21-33, <http://doi.org/10.1016/j.advwatres.2013.12.002>, 2014.

de Rosnay, P., Balsamo, G., Albergel, C., Muñoz-Sabater, J., and Isaksen, L.: Initialisation of Land Surface Variables for Numerical Weather Prediction, Surv Geophys, 35, 607-621, 10.1007/s10712-012-9207-x, 2014.

Decharme, B., Alkama, R., Papa, F., Faroux, S., Douville, H., and Prigent, C.: Global off-line evaluation of the ISBA-TRIP flood model, Clim. Dyn., 38, 1389-1412, 10.1007/s00382-011-1054-9, 2012.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Källberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553-597, 10.1002/qj.828, 2011.

Dickinson, R. E., Henderson-Sellers, A., and Kennedy, P.: Biosphere-atmosphere transfer scheme (BATS) version 1e as coupled to the NCAR community climate model, NCAR Boulder Colorado, 1993.

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., and Hanasaki, N.: GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface, Bulletin of the American Meteorological Society, 87, 1381-1397, 10.1175/bams-87-10-1381, 2006.

Döll, P., Hoffmann-Dobrev, H., Portmann, F. T., Siebert, S., Eicker, A., Rodell, M., Strassberg, G., and Scanlon, B. R.: Impact of water withdrawals from groundwater and surface water on continental water storage variations, Journal of Geodynamics, 59–60, 143-156, <http://dx.doi.org/10.1016/j.jog.2011.05.001>, 2012.

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

Dutra, E., Viterbo, P., Miranda, P. M. A., and Balsamo, G.: Complexity of Snow Schemes in a Climate Model and Its Impact on Surface Energy and Hydrology, J. Hydrometeorol., 13, 521-538, 10.1175/jhm-d-11-072.1, 2012.

Fekete, B., Vorosmarty, C., and Wolfgang, G.: Global, Composite Runoff Fields Based on Observed River Discharge and Simulated Water Balances, Global Runoff Data Centre, Koblenz, Germany: UNH/GRDC, 2000.

Fekete, B. M., Vörösmarty, C. J., and Grabs, W.: High-resolution fields of global runoff combining observed river discharge and simulated water balances, Global Biogeochemical Cycles, 16, 15-11-15-10, 10.1029/1999GB001254, 2002.

Flörke, M., Kynast, E., Bärlund, I., Eisner, S., Wimmer, F., and Alcamo, J.: Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: A global simulation study, *Global Environ. Change*, 23, 144-156, <http://dx.doi.org/10.1016/j.gloenvcha.2012.10.018>, 2013.

Funk, C., Verdin, A., Michaelsen, J., Peterson, P., Pedreros, D., and Husak, G.: A global satellite-assisted precipitation climatology, *Earth Syst. Sci. Data*, 7, 275-287, 10.5194/essd-7-275-2015, 2015.

Gao, L., Bernhardt, M., and Schulz, K.: Elevation correction of ERA-Interim temperature data in complex terrain, *Hydrol. Earth Syst. Sci.*, 16, 4661-4673, 10.5194/hess-16-4661-2012, 2012.

Getirana, A. C. V., Dutra, E., Guimberteau, M., Kam, J., Li, H.-Y., Decharme, B., Zhang, Z., Ducharne, A., Boone, A., Balsamo, G., Rodell, M., Toure, A. M., Xue, Y., Peters-Lidard, C. D., Kumar, S. V., Arsenault, K., Drapeau, G., Ruby Leung, L., Ronchail, J., and Sheffield, J.: Water Balance in the Amazon Basin from a Land Surface Model Ensemble, *J. Hydrometeorol.*, 15, 2586-2614, 10.1175/JHM-D-14-0068.1, 2014.

Gleeson, T., Moosdorf, N., Hartmann, J., and van Beek, L. P. H.: A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity, *Geophys. Res. Lett.*, 41, 3891-3898, 10.1002/2014GL059856, 2014.

Gouttevin, I., Krinner, G., Ciais, P., Polcher, J., and Legout, C.: Multi-scale validation of a new soil freezing scheme for a land-surface model with physically-based hydrology, *The Cryosphere*, 6, 407-430, 10.5194/tc-6-407-2012, 2012.

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koitala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P., and Yeh, P.: Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results, *J. Hydrometeorol.*, 12, 869-884, 10.1175/2011jhm1324.1, 2011.

Hagemann, S., and Gates, L. D.: Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations, *Clim. Dyn.*, 21, 349-359, 10.1007/s00382-003-0349-x, 2003.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, *International Journal of Climatology*, 25, 1965-1978, 10.1002/joc.1276, 2005.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, *J. Hydrometeorol.*, 8, 38-55, 10.1175/jhm560.1, 2007.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P.: CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution, *J. Hydrometeorol.*, 5, 487-503, 10.1175/1525-7541(2004)005<0487:camtpg>2.0.co;2, 2004.

Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, *Biogeosciences*, 6, 2001-2013, 10.5194/bg-6-2001-2009, 2009.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., and Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, *Front. Ecol. Environ.*, 9, 494-502, 10.1890/100125, 2011.

Lehner, B., and Grill, G.: Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, *Hydrological Processes*, 27, 2171-2186, 10.1002/hyp.9740, 2013.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., and Merchant, J. W.: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, *Int. J. Remote Sens.*, 21, 1303-1330, 2000.

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

Moore, R. J.: The probability-distributed principle and runoff production at point and basin scales *Hydrological Sciences Journal*, 30, 273-297, 1985.

Post, W. M., and Zobler, L.: Global Soil Types, 0.5-Degree Grid (Modified Zobler), Available on-line [<http://www.daac.ornl.gov>] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A., 2000.

Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Liu, Q., Mahanama, S. P. P., and Touré, A.: Assessment and Enhancement of MERRA Land Surface Hydrology Estimates, *J. Clim.*, 24, 6322-6338, 10.1175/JCLI-D-10-05033.1, 2011.

Reichle, R. H., Draper, C. S., Liu, Q., Girotto, M., Mahanama, S. P. P., Koster, R. D., and De Lannoy, G. J. M.: Assessment of MERRA-2 Land Surface Hydrology Estimates, *J. Clim.*, 30, 2937-2960, 10.1175/JCLI-D-16-0720.1, 2017.

Robinson, E. L., Blyth, E., Clark, D. B., Comyn-Platt, E., Finch, J., and Rudd, A. C.: Climate hydrology and ecology research support system meteorology dataset for Great

Britain (1961-2015) [CHESS-met] v1.2, in, NERC Environmental Information Data Centre, 2017.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin*, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bulletin of the American Meteorological Society, 85, 381-394, 10.1175/bams-85-3-381, 2004.

Ryder, J., Polcher, J., Peylin, P., Ottlé, C., Chen, Y., van Gorsel, E., Haverd, V., McGrath, M. J., Naudts, K., Otto, J., Valade, A., and Luyssaert, S.: A multi-layer land surface energy budget model for implicit coupling with global atmospheric simulations, Geosci. Model Dev., 9, 223-245, 10.5194/gmd-9-223-2016, 2016.

Schellekens, J., Dutra, E., Martínez-de la Torre, A., Balsamo, G., van Dijk, A., Sperna Weiland, F., Minvielle, M., Calvet, J. C., Decharme, B., Eisner, S., Fink, G., Flörke, M., Peßenteiner, S., van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W., and Weedon, G. P.: A global water resources ensemble of hydrological models: the earthH2Observe Tier-1 dataset, Earth Syst. Sci. Data Discuss., 2016, 1-35, 10.5194/essd-2016-55, 2016.

Schwatke, C., Dettmering, D., Bosch, W., and Seitz, F.: DAHITI – an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry, Hydrol. Earth Syst. Sci., 19, 4345-4364, 10.5194/hess-19-4345-2015, 2015.

Shangguan, W., Dai, Y., Duan, Q., Liu, B., and Yuan, H.: A global soil data set for earth system modeling, Journal of Advances in Modeling Earth Systems, 6, 249-263, 10.1002/2013MS000293, 2014.

Su, Z., Schmugge, T., Kustas, W. P., and Massman, W. J.: An Evaluation of Two Models for Estimation of the Roughness Height for Heat Transfer between the Land Surface and the Atmosphere, Journal of Applied Meteorology, 40, 1933-1951, 10.1175/1520-0450(2001)040<1933:aeotmf>2.0.co;2, 2001.

Sutanudjaja, E. H., Van Beek, L. P. H., Wada, Y., Wisser, D., Graaf, I. E. M., Straatsma, M. W., and Bierkens, M. F. P.: Development and validation of PCR-GLOBWB 2.): a 5 arc min resolution global hydrology and water resources model., EGU, Vienna, Austria, 2014,

Todini, E.: The ARNO rainfall—runoff model, Journal of Hydrology, 175, 339-382, [http://dx.doi.org/10.1016/S0022-1694\(96\)80016-3](http://dx.doi.org/10.1016/S0022-1694(96)80016-3), 1996.

USGS EROS Data Center: Global land cover characteristics data base version 2.0, http://edc2.usgs.gov/glcc/globe_int.php, 2002.

Ushio, T., Sasashige, K., Kubota, T., Shige, S., Okamoto, K. i., Aonashi, K., Inoue, T., Takahashi, N., Iguchi, T., Kachi, M., Oki, R., Morimoto, T., and Kawasaki, Z.-I.: A Kalman Filter Approach to the Global Satellite Mapping of Precipitation (GSMap) from Combined Passive Microwave and Infrared Radiometric Data, Journal of the Meteorological Society of Japan. Ser. II, 87A, 137-151, 10.2151/jmsj.87A.137, 2009.

van Beek, L. P. H., Wada, Y., and Bierkens, M. F. P.: Global monthly water stress: 1. Water balance and water availability, Water Resources Research, 47, W07517, 10.1029/2010wr009791, 2011.

Van Dijk, A. I. J. M.: AWRA Technical Report 3, Landscape Model (Version 0.5) Technical Description, WIRADA/CSIRO Water for a Healthy Country Flagship, Canberra, [Available at <http://www.clw.csiro.au/publications/waterforahealthycountry/2010/wfhc-aus-water-resources-assessment-system.pdf>.] 2010.

van Dijk, A. I. J. M., Renzullo, L. J., Wada, Y., and Tregoning, P.: A global water cycle reanalysis (2003–2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble, Hydrol. Earth Syst. Sci., 18, 2955-2973, 10.5194/hess-18-2955-2014, 2014.

Van Dijk, A. I. J. M., Brakenridge, G. R., Kettner, A. J., Beck, H. E., De Groot, T., and Schellekens, J.: River gauging at global scale using optical and passive microwave remote sensing, Water Resources Research, 52, 6404-6418, 10.1002/2015WR018545, 2016.

Vergnes, J. P., and Decharme, B.: A simple groundwater scheme in the TRIP river routing model: global off-line evaluation against GRACE terrestrial water storage estimates and observed river discharges, Hydrol. Earth Syst. Sci., 16, 3889-3908, 10.5194/hess-16-3889-2012, 2012.

Vergnes, J. P., Decharme, B., and Habets, F.: Introduction of groundwater capillary rises using subgrid spatial variability of topography into the ISBA land surface model, J. Geophys. Res. Atmos., 119, 2014JD021573, 10.1002/2014JD021573, 2014.

Verzano, K., Bärlund, I., Flörke, M., Lehner, B., Kynast, E., Voß, F., and Alcamo, J.: Modeling variable river flow velocity on continental scale: Current situation and climate change impacts in Europe, Journal of Hydrology, 424–425, 238-251, <http://dx.doi.org/10.1016/j.jhydrol.2012.01.005>, 2012.

Viney, N., Vaze, J., Crosbie, R. S., Wang, B., Dawes, W., and Frost, A.: AWRA-L v5.0: technical description of model algorithms and inputs., CSIRO Land and Water, Canberra, Australia., 2015.

Wada, Y., van Beek, L. P. H., Vivenzoli, D., Dürr, H. H., Weingartner, R., and Bierkens, M. F. P.: Global monthly water stress: 2. Water demand and severity of water stress, Water Resources Research, 47, n/a-n/a, 10.1029/2010WR009792, 2011.

Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, *Earth Syst. Dynam.*, 5, 15-40, 10.5194/esd-5-15-2014, 2014.

Wang, T., Ottlé, C., Boone, A., Ciais, P., Brun, E., Morin, S., Krinner, G., Piao, S., and Peng, S.: Evaluation of an improved intermediate complexity snow scheme in the ORCHIDEE land surface model, *J. Geophys. Res. Atmos.*, 118, 6064-6079, 10.1002/jgrd.50395, 2013.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resources Research*, 50, 7505-7514, 10.1002/2014wr015638, 2014.