Geoscientific
Model Development

Open Access

Supplement of

Development of the Community Water Model (CWatM v1.04) – a high-resolution hydrological model for global and regional assessment of integrated water resources management

Peter Burek et al.

Correspondence to: Peter Burek (burek@iiasa.ac.at)

Supplementary material

1 Description of input data

1.1 Mask map

CWATM can run globally at 0.5° (30' or ≈ 50 km by 50 km at the equator) or 5' (≈ 10 km x10 km) or a set of coordinates or a mask map can be defined to use the global dataset but running CWATM locally. Figure S1 gives an example of a mask map for the Rhine basin. Each red grid cell (here: 5') is considered for calculation. Every other grid cell is not used.
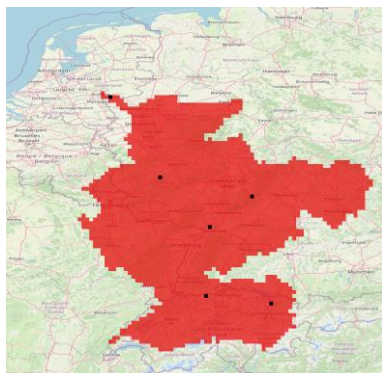

Figure S1.1: Mask map for the Rhine basin at 5' with background from © OpenStreetMap contributors 2019. Distributed under a Creative Commons BY-SA License.

1.2 Digital elevation model and river channel network

The model uses a digital elevation model and its derivate (e.g. standards deviation, slope) as variables for the snow processes and for the routing of surface runoff. The Shuttle Radar Topography Mission - SRTM (Jarvis et al. (2008)) is used for latitudes ≤ 60° N and DEM Hydro1k (USGS, 2002) is used for latitudes > 60° N.

CWATM uses a local drainage direction map which defines the dominant flow direction in one of the eight neighboring grid cells (D8 flow model). This forms a river network from the springs to the mouth of a basin. To be compliant with the ISIMIP framework (https://www.isimip.org), the 0.5° drainage direction map (DDM30) of Petra Döll et al. (2002) is used. For higher resolution e.g. 5' different sources of river network maps are available e.g. HydroSheds (B. Lehner et al., 2008) – DRT (Wu et al., 2011) and CaMa-Flood (Yamazaki et al., 2009). These approaches uses the same hydrological sound digital elevation model but differ in the upscaling methods. Zhao et al. (2017) show the importance of routing schemes and river networks in peak discharge simulation. In CWATM the DDM30 is used for 0.5° and DRT is used for 5'.

To calculate the kinematic wave in CWATM static maps of channel width, channel bankful depth, channel gradient, manning's coefficient and channel length are needed which are calculated using the river network, the flow accumulation, elevation data and average river discharge data. Figure S1.2 illustrated the channel gradient at 5' for two different regions.
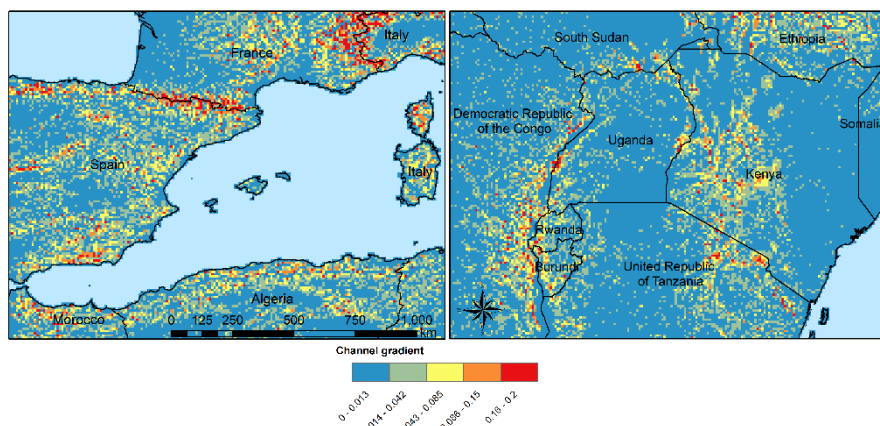

Figure S1.2: Channel gradient at 5' in percentage

1.3 Lakes and reservoirs

The HydroLakes database (Bernhard Lehner et al., 2011; Messager et al., 2016) provides 1.4 million global lakes and reservoirs with a surface area of at least 10ha. CWATM differentiate between big 'global' lakes and reservoirs which are connected to the river network and smaller 'local' lakes and reservoirs which are part of a single grid cell and part of the runoff concentration within a grid cell. Therefore the HydroLakes database is separated into "big" lakes and reservoirs with an area ≥ 100 km^2 for 0.5° (5 km^2 for 5') or upstream area ≥ 5000 km^2 0.5° (200 km^2 for 5') and into "small" lakes which represents the non-big lakes. All lakes and reservoirs are set up grid cell level but big lakes can have the expansion of several grid cells.

1.4 Soil data

Soil data comes from the Harmonized World Soil Database 1.2 (HWSD) (FAO et al., 2012)- Version 1.2 7 March, 2012 which was developed by the Land Use Change and Agriculture Program of IIASA (LUC) and the Food and Agriculture Organization of the United Nations (FAO). The HWSD is a 30 arc-second raster database with over 16000 different soil mapping units that combines existing regional and national updates of soil information worldwide – the European Soil Database (ESDB), the 1:1 million soil map of China, various regional SOTER databases (SOTWIS Database), and the Soil Map of the World – with the information contained within the 1:5000000 scale FAO-UNESCO Soil Map of the World. The resulting raster database is linked to harmonized soil property data. From the HWSD the standard soil properties like texture, porosity, soil minerals (% of sand, clay), organic matter and bulk density are used.

A pedotransfer function from Zhang et al. (2017) is used to transfer the standard soil properties (soil texture, porosity, organic matter and bulk density) to the van Genuchten model parameters: maximal amount of moisture, residual amount of moisture, pore-size index, saturated conductivity of the soil (see Figure S1.3), inverse of air entry suction.
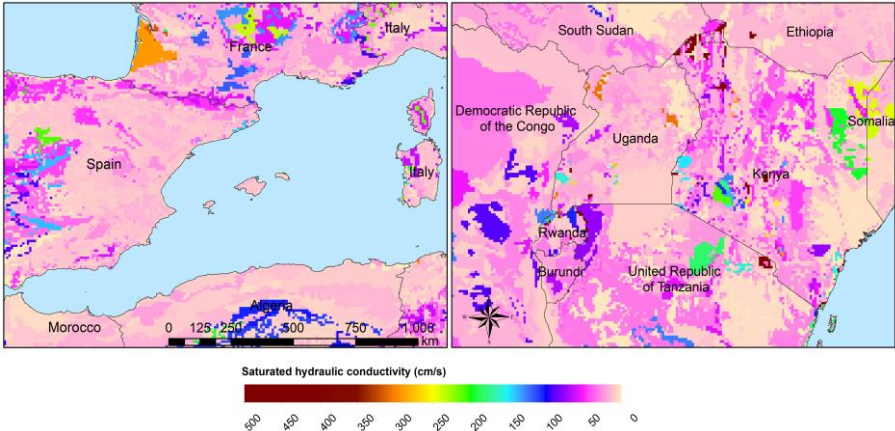

Figure S1.3: Saturated hydraulic conductivity (Ks) [cm/day] second soil layer 5-30 cm at 5'

1.5 Groundwater

For groundwater modeling maps of the recession constant of the hydraulic conductivity and the storage coefficient are needed. Gleeson et al. (2011) and the database of GLHYMPS—Global Hydrogeology Maps of permeability and porosity (Gleeson et al., 2014) are providing the necessary data. Figure S1.4 shows the global recession constant.
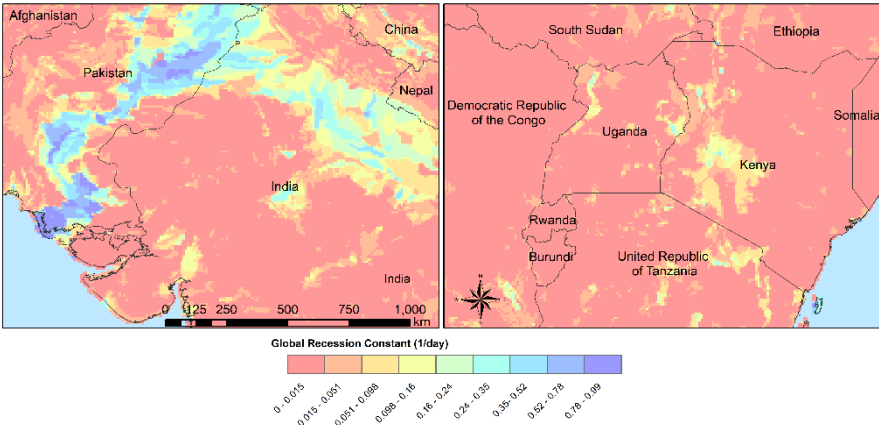

Figure S1.4: Recession constant GLIM: [1/day] at 5'

1.6 Land cover

Land cover is used to calculate fraction of water, forest, irrigated area, rice irrigated area, sealed (impermeable area) and the remaining fraction for each cell. For each fraction the soil module runs separately. The total runoff of each cell is calculated by weighting the cell according to the different fractions. Data on urban area or imperious area is based on the 1 km version of Elvidge et al. (2007), Forest land cover data are provided in high resolution by Hansen et al. (2013). Global maps of irrigated areas are used from P. Döll et al. (2002) and Siebert et al. (2005) with global data on crop coefficient based on MIRCA2000, Global data set of monthly irrigated and rainfed crop areas around the year 2000 (Portmann et al., 2010). A historical, gridded land use data set with 5' data on 15 different crop groups is used from the Hyde 3.2 database (Klein Goldewijk et al., 2017).

1.7 Global Albedo

A global dataset to calculate potential evaporation using albedo is provided by Muller et al. (2012).

1.8 Data sets of meteorological forcing

For calculating potential evaporation with Penman-Monteith some meteorological forcing data are needed:
- max, min, average temperature [K]
- humidity (relative[%] or specific[%])
- surface pressure [Pa]
- radiation (short wave and long wave downwards) [W m-2]
- wind speed [m/s]

If potential evaporation is already calculated in a previous run or from external source:
- Precipitation [Kg m-2 s-1] or [m] or [mm] (can be adjusted by a conversion factor in the settings file)
- Temperature (average) [K]
- Potential evaporation [Kg m-2 s-1] or [m] or [mm] (can be adjusted by a conversion factor in the settings file)

From observation: (ISI-MIP 2a):
- WFDEI.GPCC (Weedon et al., 2014)
    - WFD—Watch forcing data set: 0.5 3/6 hourly meteorological forcing from ECMRWF reanalysis (ERA40) bias-corrected and extrapolated by CRU TS and GPCP (rainfall) and corrections for under catch
- PGMFD v.2 - Princeton (Sheffield et al., 2006)
- GSWP3 (Kim et al., 2012)
- MSWEP (Beck et al., 2017)

From GCM (ISIMIP 2b):
- HadGem2-ES (Met Office Hadley Centre, UK)
- IPSL-CM5A-LR (Institut Pierre-Simon Laplace, France)
- GFDL-ESM2M (NOAA, USA)
- MIROC-ESM-CHEM (JAMSTEC, AORI, University of Tokyo, NIES, Japan)

For downscaling of 0.5° meteorological data to 5' or even to 30'' the dataset of WorldClim version2 (Fick et al., 2017) is used. They provide average monthly dataset for different meteorological variables for 1970-2000 for different spatial resolutions from 30'' to 10'. The dataset is used to estimate the spatial heterogeneity inside a 0.5° or 5' grid cell.

1.9 Datasets for model calibration and validation

Observed historical daily river discharge data and lake water levels were available originating from several sources. Data made available through the GRDC (Global Runoff Data Centre, 2007) were used. Several discharge station data were obtained through bilateral exchanges between IIASA and national hydrological services. As much as possible, these historical discharge data have been used for model calibration and verification.

1.10 Socio economic datasets

Data on population and gross domestic product (GDP) are based on the SSP Database of IIASA (Riahi et al., 2017). Methods and data for spatial disaggregation of country based data to gridded 5' data are taken from Jones et al. (2016); Gao (2017); Kummu et al. (2018) and Gidden et al. (2018).

The gridded livestock densities of cattle, buffalo, sheep, goats, pigs and poultry in the year 2000 and their drinking water requirements are obtained from FAO (2007) and Steinfeld et al. (2006). The number of each livestock type per country from 1979-2010 are provided by FAO (2012) and is downscaled to 5' or 30' using the distribution of livestock density in 2000. Gridded industrial water data for 2000 is using the data from Shiklomanov (1997). The approach of Shen et al. (2008) and Wada et al. (2011) is including GDP, electricity production, energy consumption and household consumption. Domestic water demand needs population data and the rate of domestic water withdrawal per capita from FAO (2007) and Gleick et al. (2009).
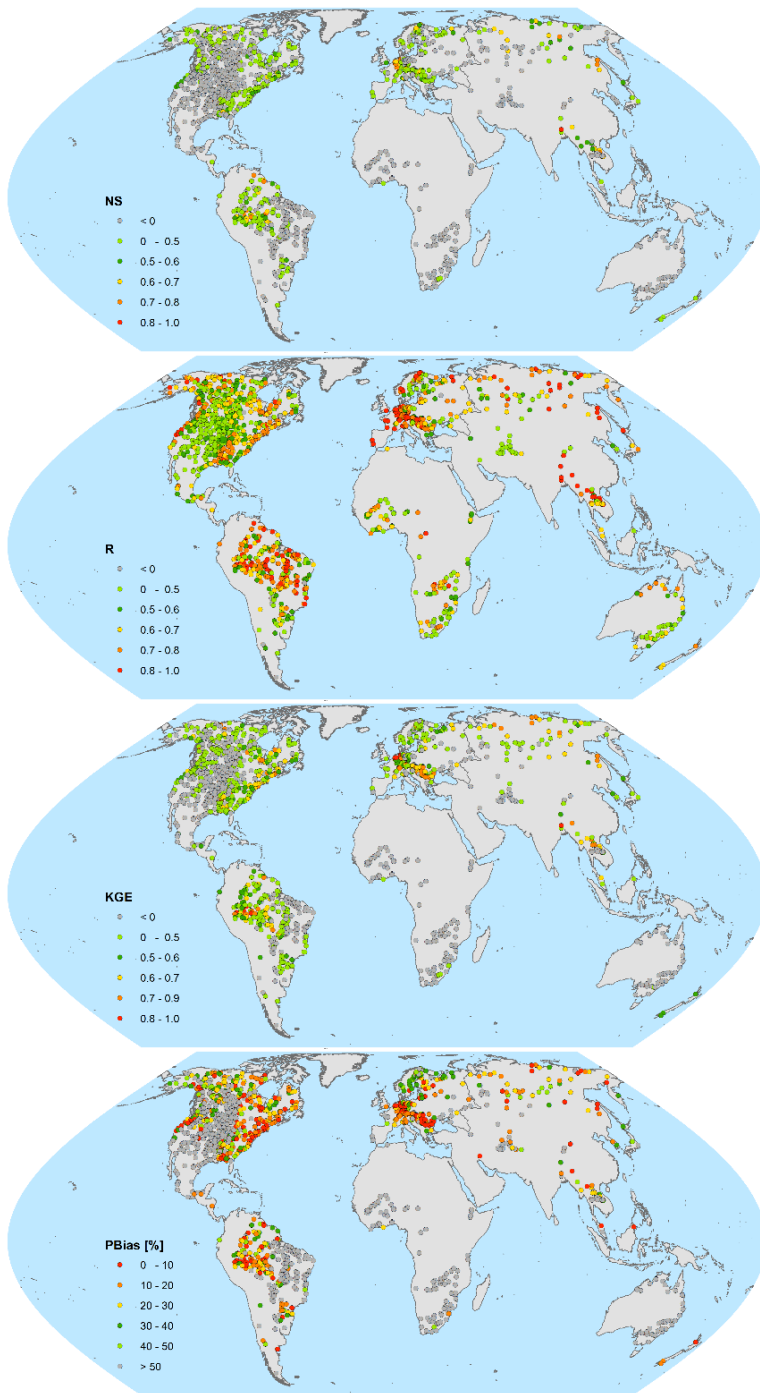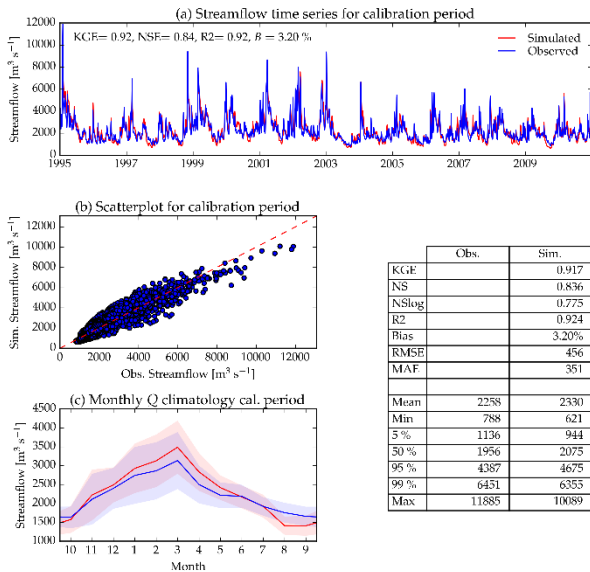
2 Global model validation

Figure S2.1: Global maps of Kling-Gupta efficiency, Nash-Sutcliffe efficiency, Pearson's correlation and percent bias based on 1366 GRDC stations

3 Calibration results

Calibration results for Rhin basin 0.5° resolution

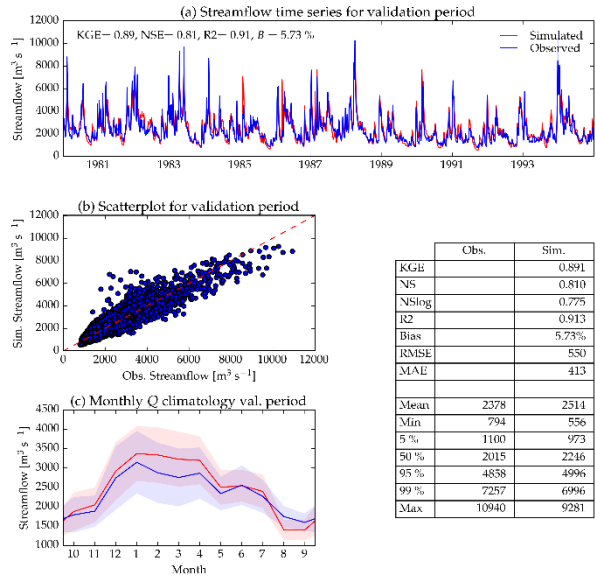### Calibration: 1995-2010

Validation: 1980-1994

Figure S3.1: Calibration result Rhine

a) Observed (blue) and simulated (red) daily or monthly streamflow time series during variable calibration periods (depending on the availability of observed data)
(b) Scatter plot showing the relationship between observed and simulated streamflow.
(c) Monthly climatology of observed (blue) and simulated (red) streamflow for the water year (from October to September) during the calibration period. Shaded area shows the standard deviation for each month for observed and simulated streamflow

The Kling-Gupta Efficiency (KGE) is used as the objective function. Other model performance metrics are summarized in the table, including NS (Nash-Sutcliffe coefficient of efficiency) and its log form (NSlog), R^2 (coefficient of determination), Bias, RMSE (root mean squared error), and MAE (mean absolute error). Observed and simulated streamflow statistics are shown, including mean, minimum, maximum, and different quantiles (5%, 50%, 95%, 99%).
Some basins do not show a validation period. Here we used all available observation data for calibration.
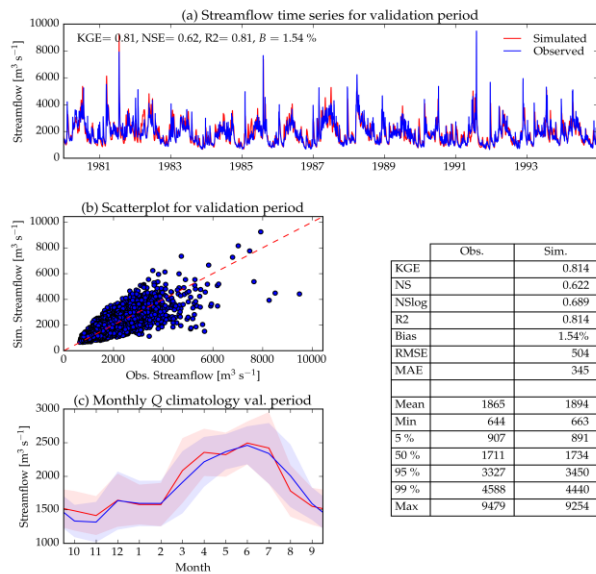
Calibration results for Danube basin

Calibration: 1995-2010

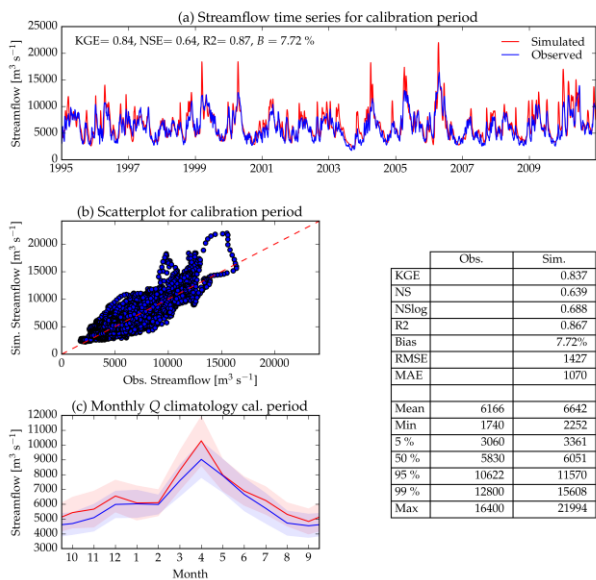### River: Danube Station: Kienstock, AT
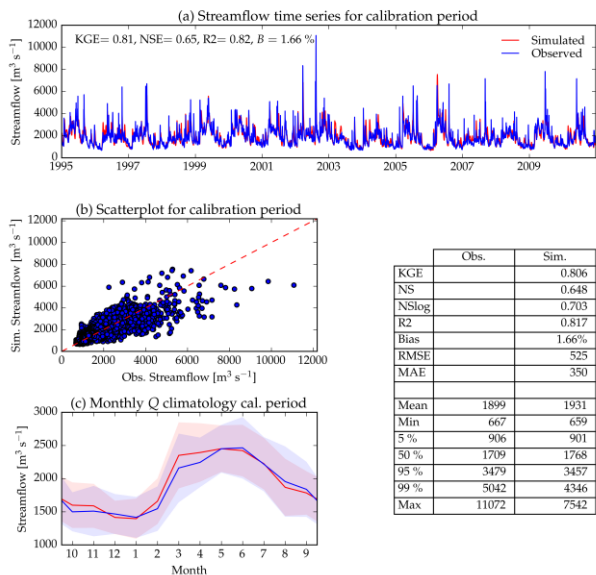
Validation: 1980-1994

River: Danube Station: Kienstock, AT

Calibration: 1995-2010

River: Danube Station: Zimnicea, RO

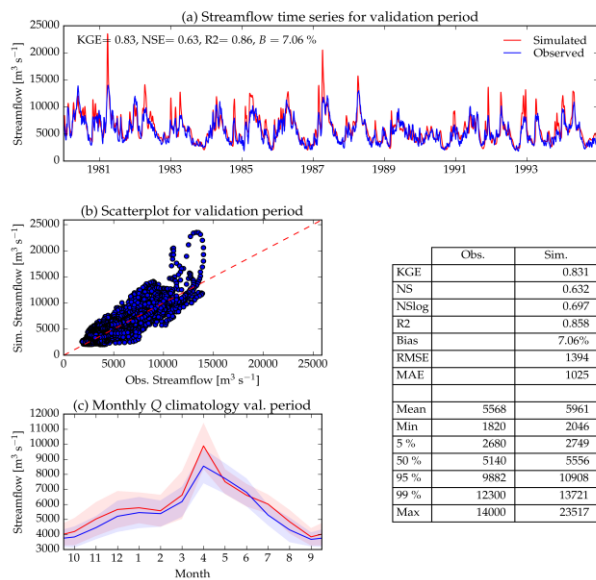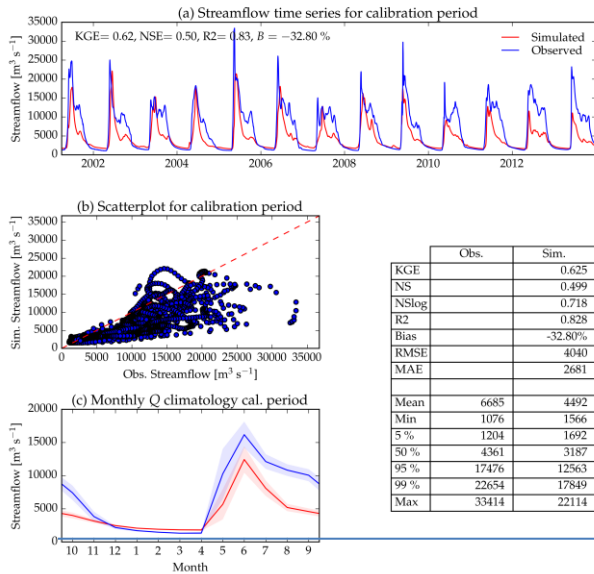## Validation: 1980-1994

River: Danube Station: Zimnicea, RO

Figure S3.2: Calibration result Danube

Calibration results for Yukon / USA basin
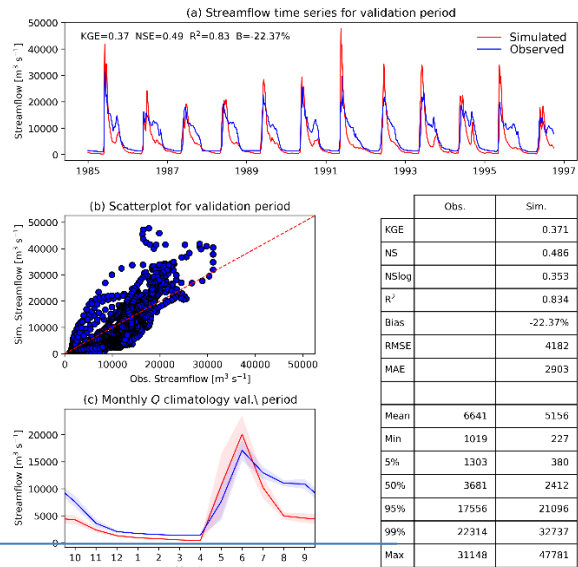
Calibration: Calibration:2001-2013

Validation: 1985-1997

Calibration results for Amazon basin, Brazil

Calibration: 1985-1997
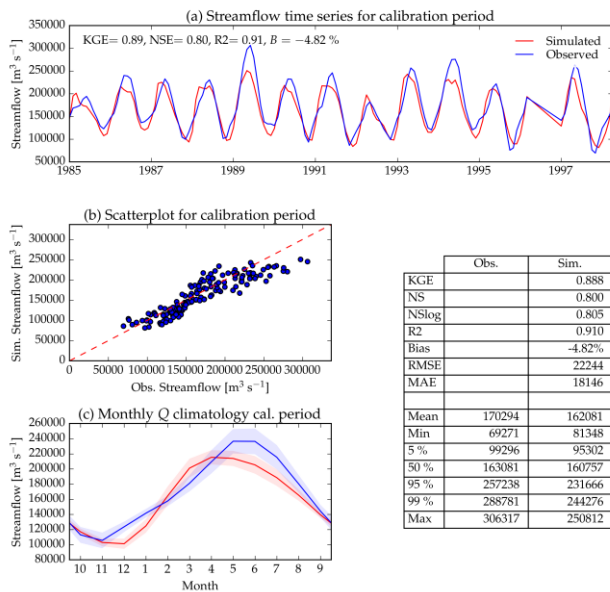
Validation: 1970-1984
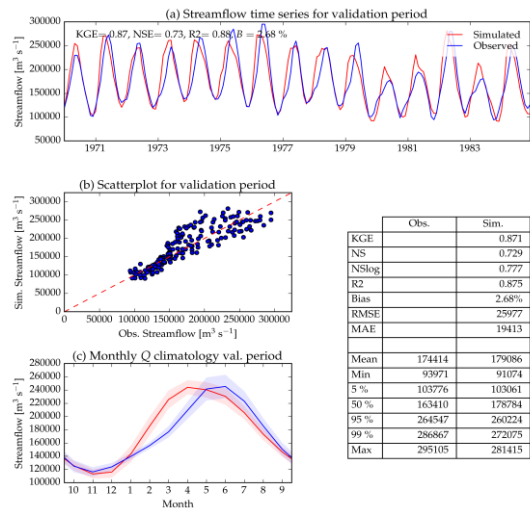
Figure S3.3: Calibration result Yukon and Amazon

Calibration results for Murray River, Australia

Calibration: 2000-2010

Validation: 1990-1999

River: Murray River station: Wakool Junction

KGE= 0.70, NSE= 0.32, R2= 0.74, B = 7.93 %
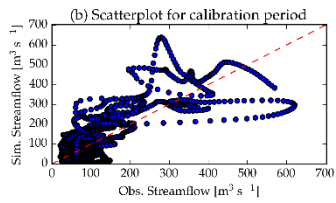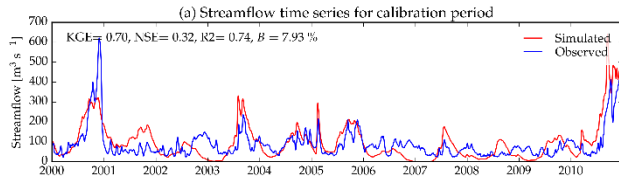
	Obs.	Sim.
KGE		0.704
NS		0.323
NSlog		-10.575
R2		0.735
Bias		7.93%
RMSE		66
MAE		50
Mean	89	96
Min	21	0
5 %	29	4
50 %	69	70
95 %	251	298
99 %	456	474
Max	621	639

River: Murray River station: Wakool Junction

KGE=0.51 NSE=0.48 R²=0.74 B=-20.01%
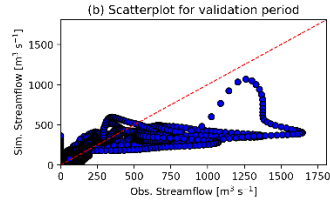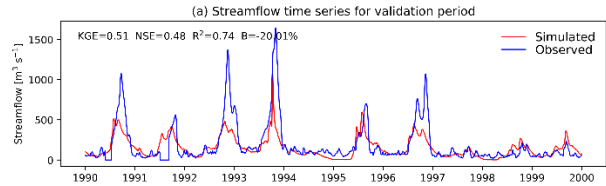
	Obs.	Sim.
KGE		0.511
NS		0.475
NSlog		nan
R²		0.744
Bias		-20.01%
RMSE		184
MAE		96
Mean	188	150
Min	0	2
5%	29	7
50%	89	110
95%	712	410
99%	1322	513
Max	1644	1068
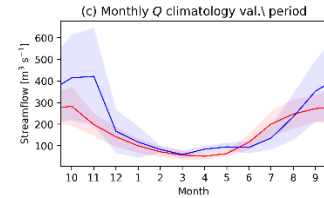
Figure S3.4: Calibration result Murray River

Calibration results for White Nile basin / Uganda

Calibration: 1996-2005

River: Nile Station: Jinja, Uganda

	Obs.	Sim.
KGE		0.941
NS		0.900
NSlog		0.860
R2		0.949
Bias		-0.00%
RMSE		70
MAE		55
Mean	992	992
Min	474	563
5 %	642	668
50 %	971	951
95 %	1394	1392
99 %	1588	1589
Max	1661	1635
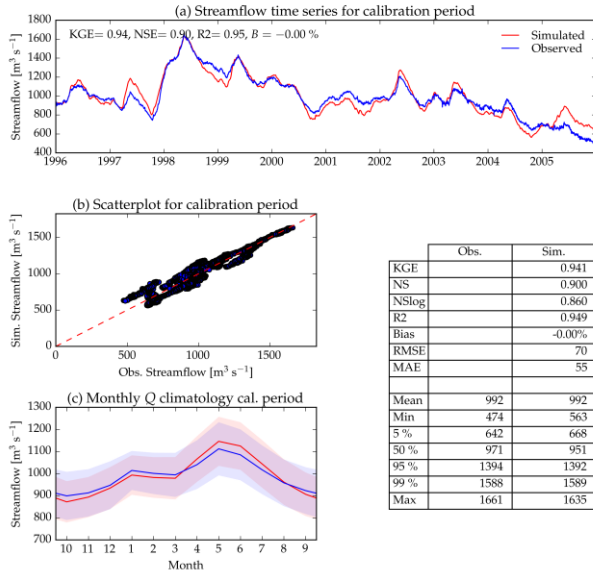
Calibration: 1996-2005

River: Nile Station: Laropi

	Obs.	Sim.
KGE		0.172
NS		-2.213
NSlog		-1.700
R2		0.213
Bias		15.13%
RMSE		346
MAE		306
Mean	1188	1367
Min	826	751
5 %	874	857
50 %	1199	1368
95 %	1456	1835
99 %	1539	1939
Max	1646	1978
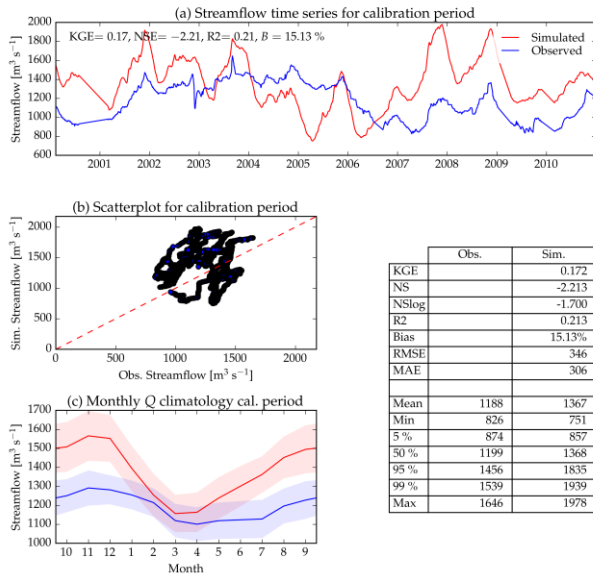
Figure S3.5: Calibration result White Nile

Calibration results for Zambezi basin

Calibration: 1979-1990

Station: Lukulu / Zambezi

	Obs.	Sim.
KGE		0.874
NS		0.785
NSlog		0.883
R2		0.887
Bias		-1.17%
RMSE		212
MAE		133
Mean	706	698
Min	196	201
5 %	236	242
50 %	515	576
95 %	1652	1459
99 %	1980	1897
Max	2084	2270
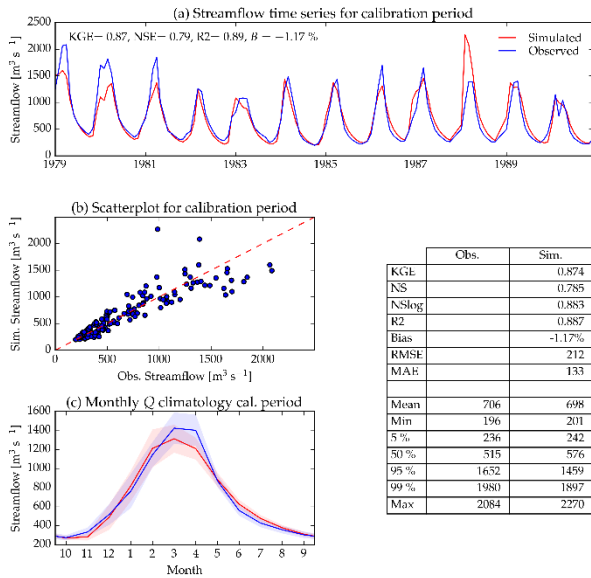
Station: Matundo-Cais / Zambezi

	Obs.	Sim.
KGE		0.451
NS		-0.431
NSlog		-0.330
R2		0.492
Bias		20.88%
RMSE		1416
MAE		1015
Mean	2188	2645
Min	666	549
5 %	876	878
50 %	1789	2586
95 %	4669	5284
99 %	6160	7453
Max	7160	10484
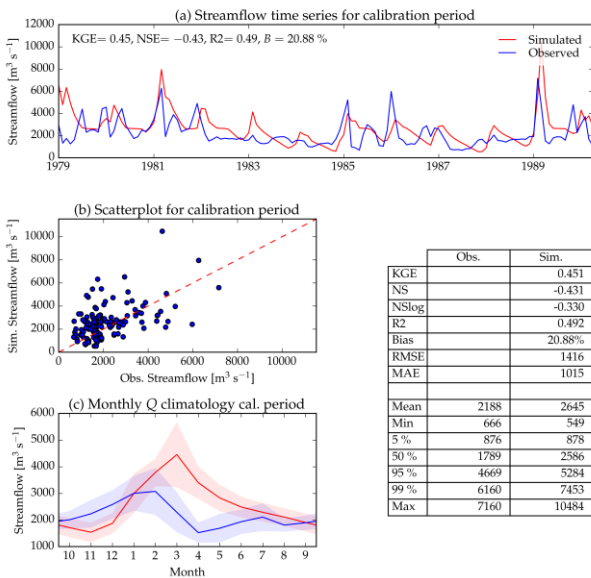
Figure S3.6: Calibration result Zambezi

4 CWatM modular structure

Figure S4.1 shows the modules of CWatM and the interlinkage of the modules. The model starts with *cwatm3* and initiate the class *dynamic model*, which has two children cwatm_initial and cwatm_dynamic. Cwatm_initial role is to initiate all hydrological modules while cwatm_dynamic runs the model through the time steps. Each hydrological process group e.g. soil has an own module. Support modules for data handling, output generating or error interception etc are separate modules and triggered by the hydrological module e.g. for reading input data.

Figure S4.1: Modular structure of CWatM

References

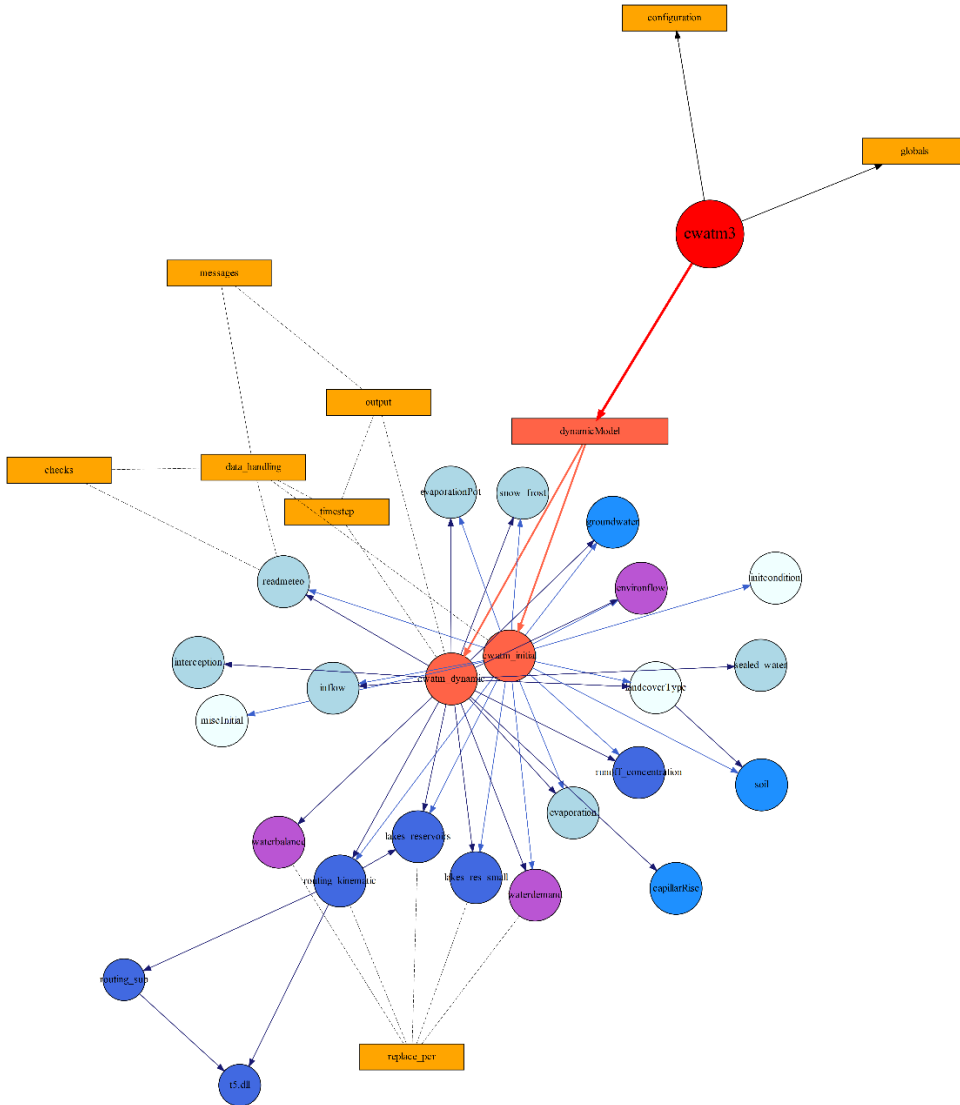Beck, H. E., Van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & De Roo, A. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences, 21*(1), 589-615. 10.5194/hess-21-589-2017: 10.5194/hess-21-589-2017.

Döll, P., & Lehner, B. (2002). Validation of a new global 30-min drainage direction map. *Journal of Hydrology, 258*(1), 214-231. 10.1016/S0022-1694(01)00565-0: 10.1016/S0022-1694(01)00565-0.

Döll, P., & Siebert, S. (2002). Global modeling of irrigation water requirements. *Water Resources Research, 38*(4), 81-811.

Elvidge, C. D., Tuttle, B. T., Sutton, P. S., Baugh, K. E., Howard, A. T., Milesi, C., . . . Nemani, R. (2007). Global distribution and density of constructed impervious surfaces. *Sensors, 7*(9), 1962-1979. 10.3390/s7091962: 10.3390/s7091962.

FAO. (2007). *Gridded livestock of the world*. Food and Agriculture Organization of the United Nations Rome.

FAO. (2012). *FAOSTAT online database*.

FAO, IIASA, ISRIC, ISSCAS, & JRC. (2012). *Harmonized World Soil Database (version 1.2)*.

Fick, S. E., & Hijmans, R. J. (2017). Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*.

Gao, J. (2017). *Downscaling Global Spatial Population Projections from 1/8-degree to 1-km Grid Cells*. National Center for Atmospheric Research, Boulder, CO, USA

Gidden, M. J., Fujimori, S., van den Berg, M., Klein, D., Smith, S. J., van Vuuren, D. P., & Riahi, K. (2018). A methodology and implementation of automated emissions harmonization for use in Integrated Assessment Models. *Environmental Modelling and Software, 105*, 187-200. 10.1016/j.envsoft.2018.04.002: 10.1016/j.envsoft.2018.04.002.

Gleeson, T., Moosdorf, N., Hartmann, J., & Van Beek, L. P. H. (2014). A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity. *Geophysical Research Letters, 41*(11), 3891-3898. 10.1002/2014GL059856: 10.1002/2014GL059856.

Gleeson, T., Smith, L., Moosdorf, N., Hartmann, J., Dürr, H. H., Manning, A. H., . . . Jellinek, A. M. (2011). Mapping permeability over the surface of the Earth. *Geophysical Research Letters, 38*(2). 10.1029/2010GL045565: 10.1029/2010GL045565.

Gleick, P. H., Cooley, H., Cohen, M. J., Morikawa, M., Morrison, J., & Palaniappan, M. (2009). *The Worlds Water 2008–2009, The Biennial Report on Freshwater Resources*. DC, USA.

GRDC (2007): Major River Basins of the World / Global Runoff Data Centre, GRDC. Koblenz, Germany: Federal Institute of Hydrology (BfG).

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., . . . Townshend, J. R. G. (2013). High-resolution global maps of 21st-century forest cover change. *Science, 342*(6160), 850-853. 10.1126/science.1244693: 10.1126/science.1244693.

Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). *Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database (http://srtm.csi.cgiar.org)*.

Jones, B., & O'Neill, B. C. (2016). Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environmental Research Letters, 11*(8), 084003.

Kim, H., Watanabe, S., Chang, E.-C., Yoshimura, K., Hirabayashi, Y., Famiglietti, J., & Oki, T. (2012). *Century long observation constrained global dynamic downscaling and hydrologic implication*. Paper presented at the American Geophysical Union, Fall Meeting 2012.

Klein Goldewijk, K. K., Beusen, A., Doelman, J., & Stehfest, E. (2017). Anthropogenic land use estimates for the Holocene - HYDE 3.2. *Earth System Science Data, 9*(2), 927-953. 10.5194/essd-9-927-2017: 10.5194/essd-9-927-2017.

Kummu, M., Taka, M., & Guillaume, J. H. A. (2018). Gridded global datasets for Gross Domestic Product and Human Development Index over 1990-2015. *Scientific Data, 5*. 10.1038/sdata.2018.4: 10.1038/sdata.2018.4.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., . . . Wisser, D. (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment, 9*(9), 494-502. 10.1890/100125: 10.1890/100125.

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, 89*(10), 93-94. 10.1029/2008EO100001: 10.1029/2008EO100001.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *7*, 13603. 10.1038/ncomms13603: 10.1038/ncomms13603.

Muller, P. J., Lewis, P., Fischer, J., North, P., & Framer, U. (2012). *The ESA GlobAlbedo Project for mapping the Earth's land surface albedo for 15 Years from European Sensors., paper presented at IEEE Geoscience and Remote Sensing Symposium (IGARSS)* Paper presented at the IEEE Geoscience and Remote Sensing Symposium (IGARSS) 2012, Munich, Germany.

Portmann, F. T., Siebert, S., & Döll, P. (2010). MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles, 24*(1), n/a-n/a. 10.1029/2008GB003435: 10.1029/2008GB003435.

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., . . . Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change, 42*, 153-168. 10.1016/j.gloenvcha.2016.05.009: 10.1016/j.gloenvcha.2016.05.009.

Sheffield, J., Goteti, G., & Wood, E. F. (2006). Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate, 19*(13), 3088-3111. 10.1175/JCLI3790.1: 10.1175/JCLI3790.1.

Shen, Y., Oki, T., Utsumi, N., Kanae, S., & Hanasaki, N. (2008). Projection of future world water resources under SRES scenarios: Water withdrawal. *Hydrological Sciences Journal, 53*(1), 11-33. 10.1623/hysj.53.1.11: 10.1623/hysj.53.1.11.

Shiklomanov, I. A. (1997). Assessment of Water Resources and Water Availability in the World. *Comprehensive Assessment of the Freshwater Re-sources of the World*.

Siebert, S., Döll, P., Hoogeveen, J., Faures, J. M., Frenken, K., & Feick, S. (2005). Development and validation of the global map of irrigation areas. *Hydrology and Earth System Sciences, 9*(5), 535-547.

Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M., & de Haan, C. (2006). Livestock's long shadow: Environmental issues and options. *Renewable Resources Journal, 24*(4), 15-17.

USGS, U. G. S. C. f. E. R. O. a. S. (2002). *Hydro1k*. Retrieved from: https://lta.cr.usgs.gov/HYDRO1K

Wada, Y., Van Beek, L., & Bierkens, M. F. (2011). Modelling global water stress of the recent past: on the relative importance of trends in water demand and climate variability. *Hydrology and Earth System Sciences, 15*(12), 3785-3808.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing data methodology applied to ERA-Interim reanalysis data. *Water Resources Research, 50*(9), 7505-7514. 10.1002/2014WR015638: 10.1002/2014WR015638.

Wu, H., Kimball, J. S., Mantua, N., & Stanford, J. (2011). Automated upscaling of river networks for macroscale hydrological modeling. *Water Resources Research, 47*(3). 10.1029/2009WR008871: 10.1029/2009WR008871.

Yamazaki, D., Oki, T., & Kanae, S. (2009). Deriving a global river network map and its sub-grid topographic characteristics from a fine-resolution flow direction map. *Hydrology and Earth System Sciences, 13*(11), 2241-2251.

Zhang, Y., & Schaap, M. G. (2017). Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). *Journal of Hydrology, 547*, 39-53. 10.1016/j.jhydrol.2017.01.004: 10.1016/j.jhydrol.2017.01.004.

Zhao, F., Veldkamp, T. I. E., Frieler, K., Schewe, J., Ostberg, S., Willner, S., . . . Yamazaki, D. (2017). The critical role of the routing scheme in simulating peak river discharge in global hydrological models. *Environmental Research Letters, 12*(7), 075003.