UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1211

Disinformative and Uncertain Data in Global Hydrology

Challenges for Modelling and Regionalisation

ANNA KAUFFELDT

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014

ISSN 1651-6214
ISBN 978-91-554-9121-5
urn:nbn:se:uu:diva-236864

Dissertation presented at Uppsala University to be publicly examined in Hambergsalen, Villavägen 16, Uppsala, Friday, 23 January 2015 at 10:00 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor András Bárdossy (Universität Stuttgart).

Abstract

Kauffeldt, A. 2014. Disinformative and Uncertain Data in Global Hydrology. Challenges for Modelling and Regionalisation. (Desinformativa och osäkra data i global hydrologi. Utmaningar för modellering och regionalisering). *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1211. 79 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9121-5.

Water is essential for human well-being and healthy ecosystems, but population growth and changes in climate and land-use are putting increased stress on water resources in many regions. To ensure water security, knowledge about the spatiotemporal distribution of these resources is of great importance. However, estimates of global water resources are constrained by limitations in availability and quality of data. This thesis explores the quality of both observational and modelled data, gives an overview of models used for large-scale hydrological modelling, and explores the possibilities to deal with the scarcity of data by prediction of flow-duration curves.

The evaluation of the quality of observational data for large-scale hydrological modelling was based on both hydrographic data, and model forcing and evaluation data for basins worldwide. The results showed that a GIS polygon dataset outperformed all gridded hydrographic products analysed in terms of representation of basin areas. Through a screening methodology based on the long-term water-balance equation it was shown that as many as 8–43% of the basins analysed displayed inconsistencies between forcing (precipitation and potential evaporation) and evaluation (discharge) data depending on how datasets were combined. These data could prove disinformative in hydrological model inference and analysis.

The quality of key hydrological variables from a numerical weather prediction model was assessed by benchmarking against observational datasets and by analysis of the internal land-surface water budgets of several different model setups. Long-term imbalances were found between precipitation and evaporation on the global scale and between precipitation, evaporation and runoff on both cell and basin scales. These imbalances were mainly attributed to the data assimilation system in which soil moisture is used as a nudge factor to improve weather forecasts.

Regionalisation, i.e. transfer of information from data-rich areas to data-sparse areas, is a necessity in hydrology because of a lack of observed data in many areas. In this thesis, the possibility to predict flow-duration curves in ungauged basins was explored by testing several different methodologies including machine learning. The results were mixed, with some well predicted curves, but many predicted curves exhibited large biases and several methods resulted in unrealistic curves.

Keywords: Data uncertainty, Discharge, Disinformative data, Evaporation, Flow-duration curve, Global hydrology, Neural networks, Numerical weather prediction, Precipitation, Quality control, Regionalisation, Ungauged basins, Water balance

Anna Kauffeldt, Department of Earth Sciences, LUVAL, Villav. 16, Uppsala University, SE-75236 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-554-9121-5
urn:nbn:se:uu:diva-236864 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-236864)

Akademisk avhandling som för avläggande av teknologie doktorsexamen i hydrologi vid Uppsala universitet kommer att offentligen försvaras i Hambergsalen, Villavägen 16, Uppsala, fredagen 23 januari 2015, klockan 10:00. Fakultetsopponent: Professor András Bárdossy (Universität Stuttgart). Disputationen sker på engelska.

Referat

Kauffeldt, A. 2014. Desinformativa och osäkra data i global hydrologi – utmaningar för modellering och regionalisering. Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1211. 79 sid. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9121-5.

Vatten är en förutsättning för människors och ekosystems hälsa, men befolkningsökning och förändringar av klimat och markanvändning förväntas öka trycket på vattenresurserna i många regioner i världen. För att kunna säkerställa en god tillgång till vatten krävs kunskap om hur dessa resurser varierar i tid och rum. Tillförlitligheten hos skattningar av globala vattenresurser begränsas dock både av begränsad tillgänglighet av och kvalitet hos observerade data. Denna avhandling utforskar kvaliteten av såväl observations- som modellbaserade data, ger en överblick över modeller som används för storskalig hydrologisk modellering och utforskar möjligheterna att förutsäga varaktighetskurvor som ett sätt att hantera bristen på data i många områden.

Utvärderingen av observationsbaserade datas kvalitet baserades på hydrografiska data och driv- och utvärderingsdata för storskaliga hydrologiska modeller. Resultaten visade att en uppsättning data över hydrografin baserad på GIS-polygoner representerade avrinningsområdesareorna bättre än alla de som byggde på rutor. En metod baserad på långtidsvattenbalansen identifierade att kombinationen av drivdata (nederbörd och potentiell avdunstning) och utvärderingsdata (vattenföring) var fysiskt orimlig för så många som 8–43 % av de analyserade avrinningsområdena beroende på hur olika datauppsättningar kombinerades. Sådana data kan vara desinformativa för slutsatser som dras av resultat från hydrologiska modeller och analyser.

Kvaliteten hos hydrologiskt viktiga variabler från en numerisk väderprognosmodell utvärderades dels genom jämförelser med observationsdata och dels genom analys av landytans vattenbudget för ett flertal olika modellvarianter. Resultaten visade obalanser mellan långtidsvärden av nederbörd och avdunstning i global skala och mellan långtidsvärden av nederbörd, avdunstning och avrinning i både modellrute- och avrinningsområdesskala. Dessa obalanser skulle till stor del kunna förklaras av den data assimilering som görs, i vilken markvattenlagret används som en justeringsfaktor för att förbättra väderprognoserna.

Regionalisering, som innebär en överföring av information från områden med god tillgång på mätdata till områden med otillräcklig tillgång, är i många fall nödvändig för hydrologisk analys på grund av att mätdata saknas i många områden. I denna avhandling utforskades möjligheten att förutsäga varaktighetskurvor för avrinningsområden utan vattenföringsdata genom flera metoder inklusive maskininlärning. Resultaten var blandade med en del kurvor som förutsas väl, och andra kurvor som visade stora systematiska avvikelser. Flera metoder resulterade i orealistiska kurvor (ickemonotona eller med negativa värden).

Nyckelord: Avdunstning, avrinningsområden utan vattenföringsdata, dataosäkerhet, desinformativa data, global hydrologi, kvalitetskontroll, nederbörd, neurala nät, numerisk vädermodell, regionalisering, varaktighetskurva, vattenbalans, vattenföring

Anna Kauffeldt, Institutionen för geovetenskaper, LUVAL, Villav. 16, Uppsala universitet, 752 36 Uppsala.

ISSN 1651-6214
ISBN 978-91-554-9121-5
urn:nbn:se:uu:diva-236864 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-236864)

“Let me begin, as all science begins,
with the subject of data.“

(Kirchner, 2006)

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I	Juston, J., Kauffeldt, A., Quesada Montano, B., Seibert, J., Beven, K., Westerberg, I. (2013). Smiling in the rain: Seven reasons to be positive about uncertainty in hydrological modelling. *Hydrological Processes*, 27(7): 1117–1122 © 2012 John Wiley & Sons, Ltd, reprinted with permission.
II	Kauffeldt, A., Halldin, S., Rodhe, A., Xu, C.-Y., Westerberg, I. (2013). Disinformative data in large-scale hydrological modelling. *Hydrology and Earth System Sciences*, 17(7): 2845–2857 © Authors 2013. CC Attribution 3.0 License.
III	Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P. and Thielen, J. (2014). Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling & Software. In review.*
IV	Kauffeldt, A., Halldin, S., Pappenberger, F., Wetterhall, F., Xu, C.-Y., Cloke, H.L. (2014). Imbalanced land-surface water budgets in a numerical weather prediction system. *Geophysical Research Letters. In review.*
V	Kauffeldt, A., Halldin, S., Rodhe, A., Xu, C.-Y. (2014). Regionalisation of flow-duration curves on global scale – assessment of methods. *Manuscript.*

In paper **I**, I drafted the sections “We can engender trust by recognising and communicating uncertainties” and “We deepen academic understanding” and contributed in the writing and the discussions from which the commentary originated. In paper **II**, I performed the quality controls and all the analyses and had the overall responsibility for writing the paper. In paper **III**, I was responsible for the literature review and communication with model developers and had the main responsibility for writing of the paper. In paper **IV**, F. Pappenberger retrieved and interpolated all model outputs from the European Centre for Medium-Range Forecasts (ECMWF) and I was responsible for gathering and preparing the remaining data, performing the analysis and writing the paper. In paper **V**, I was responsible for developing the experimental setup, all data collection and analyses and writing the paper. All co-authors have contributed with ideas and by giving advice and feedback on the analyses and the papers. Reprints were made with permission from the respective publishers.

Contents

Abbreviations

ANN	Artificial Neural Network
CRU	Climate Research Unit
DGVM	Dynamic Global Vegetation Model
DRT	Hydrographic dataset based on the Dominant River Tracing technique
E_A	(Inferred) Actual Evaporation
ECMWF	European Centre for Medium-Range Weather Forecasts
E_{LFE}	Diagnostic evaporation data product from the LandFlux-EVAL multi-data set synthesis
EOF	Empirical Orthogonal Function
E_P	Potential Evaporation
ERA-40	ECMWF reanalysis
ERA-CM	ECMWF century model integration
ERA-Interim	ECMWF reanalysis
FAO	Food and Agricultural Organization
FDC	Flow-Duration Curve
GHM	Global Hydrological Model
GloFAS	Global Flood Awareness System
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GRanD	Global Reservoir and Dam Database
GRDC	Global Runoff Data Centre
GSWP	Global Soil Wetness Project
H-TESSEL	Hydrology-Tiled ECMWF Scheme for Surface Exchange over Land
IC	Index of concavity
ISLSCP	International Satellite Land-Surface Climatology Project
ISRIC-WISE	International Soil Reference and Information Centre - World Inventory of Soil Emission Potentials
O_{H1d}	High-resolution operational forecast (1 day from analysis)
O_{L10d}	Low-resolution operational forecast (10 days from analysis)
IAHS	International Association of Hydrological Sciences
LSM	Land-Surface Model
$Land_C$	Offline run of the land surface model H-TESSEL with corrected precipitation forcing

$Land_U$	Offline run of the land surface model H-TESSEL with uncorrected precipitation forcing
Macro-PDM	Macro-scale Probability-Distributed-Moisture model
NN	Nearest Neighbour
NRMSE	Normalised Root Mean Square Error
NWP	Numerical Weather Prediction
P	Precipitation
PILPS	Project for the Intercomparison of Land-surface Parameterization Schemes
PUB	Predictions in Ungauged Basins
R	Runoff
RC	Runoff Coefficient
RE	Relative Error
SR	Stepwise Linear Regression
VIC	Variable Infiltration Capacity model
WASMOD-M	Water And Snow balance MODelling system - Macro-scale
WATCH	WATer and global CHange project
WaterGAP	Water - Global Analysis and Prognosis model
WaterMIP	Water Model Intercomparison Project
WBM	Water Balance Model and Water Transport Model

Introduction

Water is essential for all known forms of life and for our socio-economical systems. The 2014 UN-water report on the status of freshwater resources states that 768 million people lack access to clean water and 2.5 billion lack access to adequate sanitation[1] (WWAP, 2014). Global water resources are under continually increasing demand due to population increase and growing economies (Oki and Kanae, 2006; Alcamo et al., 2007; Gleick and Palaniappan, 2010). In addition, climate and land-use change is expected to put further strain on the finite water resources (Vörösmarty et al., 2000a; Jiménez Cisneros et al., 2014).

Water security, which relates to sufficient quantity and quality of water not only for human needs, but also for ecosystems, is one of the most important challenges facing humanity today. It is closely interlinked to energy and food security and together they are of fundamental importance for the alleviation of poverty (WWAP, 2014). Due to the complex interactions between water, food and energy, there is a growing recognition that a *nexus*[2] approach is needed to tackle these challenges (Finley and Seiber, 2014). However, this requires knowledge about the spatio-temporal distribution of these resources.

Considering their great importance, it is astonishing how poorly we understand global water resources. For instance, in a model intercomparison project where 11 global models were forced with the same climate data, Haddeland et al. (2011) reported a range of values from 42,000 up to 66,000 $km^3\ yr^{-1}$ for global runoff. Widén-Nilsson (2007) reported differences of around 30% in global runoff estimates in the literature and for individual continents of as much as 70%.

Two key areas of concern in estimating global water resources are the limited availability, and limited quality, of observational data worldwide (Döll and Siebert, 2002; Fekete et al., 2004; Decharme and Douville, 2006; Güntner, 2008; Hunger and Döll, 2008; Widén-Nilsson et al., 2009; Peel et al., 2010). Early global water estimates were made by extrapolation of a limited set of discharge data (e.g. Baumgartner and Reichel, 1975; Korzun et

[1] Clean water refers to water supplies that are protected from outside contamination, and adequate sanitation refers to sanitation facilities that hygienically separate human excreta from human contact.
[2] From latin *nexus*: "the act of binding together; bond".

al., 1978; L'vovich, 1979; Shiklomanov, 1997), but modern estimates predominantly turn to modelling methods.

Macro-scale hydrological modelling is a relatively young field with the first model being developed in the 1980's (Vörösmarty et al., 1989). Hydrological models can be used to meet the need of estimating water resources when observational data are missing, in time or space, and for predictions of future water resources. Large-scale hydrological modelling offers a way to estimate trans-boundary water resources, but is also fraught with high uncertainties. These uncertainties stem from different parts of the modelling process and have not been properly recognised in hydrological modelling in general (Beven, 2009) and large-scale modelling in particular (Haddeland et al., 2011). Accounting for uncertainties should be a natural part of any hydrological study and is a necessity for drawing robust conclusions.

One of the main uncertainties in hydrological modelling is the lack of high-quality climate observations. Uncertainties in precipitation, the main driver of hydrological models, has been shown to have a significant impact on model results (Vörösmarty et al., 1989; Fekete et al., 2004; Biemans et al., 2009). However, precipitation products available on the global scale are often biased and quality varies over time and in space largely dependent on the monitoring networks and measurement techniques (Adler et al., 2011; Tapiador et al., 2012).

Over recent decades, the skill and resolution of global atmospheric models have increased significantly (Simmons and Hollingsworth, 2002; Palmer et al., 2007; Rodwell et al., 2010; Wedi, 2014). The land-surface components of these models have traditionally been developed with a strong focus on the energy balance (Overgaard et al., 2006), but with higher resolutions and better descriptions of hydrological processes they have become more similar to hydrological models over the years (Balsamo et al., 2011). With model outputs of key hydrological variables, such as precipitation, evaporation and runoff, being readily available, the applications of these models have been extended to fields which hydrological models have traditionally been used for (e.g. Alfieri et al., 2013). However, these models are commonly evaluated from a meteorological perspective and their performance on hydrologically relevant spatial and temporal scales need further investigation.

Not only do we lack quality climate observations, but approximately 50% of the global land surface has been estimated to be ungauged, i.e. lacks discharge data (Fekete et al., 2002). This implies that models can not be calibrated or evaluated in these basins and therefore knowledge gained from areas with measurements needs to be transferred to areas without measurements, so called regionalisation. Prediction in ungauged basins is a core research question in hydrology which manifested itself in the 10-year science plan of the International Association of Hydrological Sciences (IAHS), "Predictions in Ungauged Basins" (PUB; Sivapalan et al., 2003).

It is clear that estimation of global water resources is fraught with difficulties due to data limitations. These limitations, in terms of quality, quantity and availability, together with our limited understanding of the data uncertainties, make modelling over large spatial domains a challenging task. In addition, it is complex to parameterise hydrological processes that occur at time scales and spatial scales much finer than current model resolutions. All these issues contribute to the high uncertainties in global runoff estimates.

Aim of this thesis

The main aim of this thesis is to explore the uncertainties in data available for global-scale hydrological studies in terms of consistency between datasets and how these data, albeit their uncertainties, can be used to extend our knowledge from comparatively data-rich domains to data-poor domains. This can be broken down into the following specific research questions:

- I Given the high uncertainties in individual datasets for large-scale hydrological studies, how consistent are these data?
- II Outputs from numerical global weather prediction models include key hydrological variables such as precipitation, evaporation and runoff. How hydrologically representative are these and how useful are the data for hydrological applications?
- III To extend our knowledge from data-rich areas to data-sparse areas we need to transfer our knowledge between spatial units. How do methods used for regionalisation in local or regional studies perform on the global scale?

Large-scale hydrological modelling

An understanding of the variable nature of water resources is necessary for appropriate management of resources and for resilience against future possible changes in these resources. The necessity of estimating water resources has led to the development of hydrological models, which can be used to assess water resources in regions and times with sparse or no data. These models are aimed at quantifying the different fluxes and storages in the hydrological system and can, e.g., serve as tools for resource management or flood warning systems, but are also important for educational and research purposes (paper **I**). Many hydrological models have been developed for catchment-scale studies for specific regions, but the recognition of the global and continental nature of water-related issues and the strong impact of land-surface processes on climate, led to a call for large-scale hydrological models in the 1980's (Eagleson, 1986; Shuttleworth, 1988). Today, there exists a large number of models that simulate runoff over large spatial domains, but they differ in terms of complexity and modelling aims (paper **III**).

Land-surface models

Land-surface models (LSMs) form the surface component in atmospheric models and have traditionally been developed within the meteorological community with a focus on the energy balance rather than the water balance (Overgaard et al., 2006). In the beginning, the models were very simple, such as the bucket model developed by Manabe (1969). A big advancement in the land-surface modelling came with the first physically-based LSM developed by Deardorff (1978) with methods to take into account soil temperature, soil moisture and vegetation as layers (Sellers et al., 1986; Pitman, 2003). LSMs differ in complexity (Dirmeyer et al., 2006), operate at sub-daily time steps to solve the energy balance, are typically run at coarser horizontal resolutions than global hydrological models (GHMs), and include no routing (Sood and Smakhtin, 2014; paper **III**).

Dynamic global vegetation models (DGVMs) can be described as an extension of LSMs and include additional processes such as the carbon cycle (e.g. Foley et al., 1996; McGuire et al., 2001). However, with the continuous development of LSMs the difference between these models has been decreasing and many of the LSMs now include for example photosynthesis and respiration (Williams et al., 2009).

One LSM that has been applied in global runoff assessments is the VIC Variable Infiltration Capacity model (VIC; Wood et al., 1992; Liang et al., 1994). Nijssen et al. (2001) found discharge simulated with VIC to be somewhat lower for some continents compared to previous studies. Demaria et al. (2007) found that runoff simulations were sensitive to only three of ten parameters analysed in VIC, and concluded that the base flow formulation was over-parameterised and could be simplified.

Another LSM that has been applied in global water-resources studies is the H08 model developed by Hanasaki et al. (2008a; 2008b). The model could be described as a hybrid between a LSM and a GHM as it solves the surface energy and water balances and includes modules such as surface hydrology, crop growth, river routing, reservoirs, environmental flow and anthropogenic withdrawals. The model has been applied globally, e.g., to identify water-stressed regions (Hanasaki et al., 2008b) and to assess virtual water transfers (Hanasaki et al., 2010; Dalin et al., 2012).

The Hydrology-Tiled ECMWF Scheme for Surface Exchange over Land (H-TESSEL; Balsamo et al., 2009), the LSM of the numerical weather prediction model at the ECMWF, is another example of a LSM that is being used in applications, which have traditionally been done with hydrological models. Within the Global Flood Awareness System (GloFAS) it is forced with an ensemble of numerical weather predictions to issue flood warnings globally based on climatological thresholds (Alfieri et al., 2013).

Global hydrological models

The history of global hydrological models starts with the Water Balance Model and Water Transport Model (WBM) developed by Vörösmarty et al. (1989; 1998). Today there exists a number of global hydrological models, e.g., the Macro-scale Probability-Distributed-Moisture model (Macro-PDM; Arnell, 1999; 2003), Water - Global Analysis and Prognosis model (WaterGAP; Döll et al., 2003; Kaspar, 2004) and Water And Snow balance MODelling system - Macro-scale (WASMOD-M; Widén-Nilsson et al., 2007; 2009). These models were developed within the hydrological community and typically differ from LSMs in their resolution (temporal and spatial) and the detail in which hydrological processes are described (Sood and Smakhtin, 2014; paper **III**).

The snow schemes are typically simple degree-day methods, since the energy balance is not modelled. Potential evaporation is either part of the forcing or calculated internally with, e.g., the Penman-Monteith (Monteith, 1965) or the Priestley-Taylor (Priestley and Taylor, 1972) formulas.

Complexity varies between models (paper **III**): some include groundwater components (e.g. WBMplus; Wisser et al., 2010), some allow for subgrid variability in soil moisture (e.g. MacPDM) and some account for anthropogenic alterations to the systems such as reservoirs, irrigation and water withdrawals (e.g. WaterGAP).

Commonly, the resolution of input data has determined the spatial resolution of these models and many of them operate at a 0.5°×0.5° latitude-longitude grid (Sood and Smakhtin, 2014; paper **III**). However, higher resolution models have been developed, e.g., a 5′ version of WaterGAP (paper **III**), and there is ongoing research on developing hyper-resolution (0.1–1 km) GHMs (Bierkens et al., 2014).

Model intercomparisons

The different modelling traditions from which LSMs and GHMs have emerged mean that they share some common features within their groups, but developments over the years have to some extent closed the gap between them (Balsamo et al., 2011; paper **IV**).

The Project for the Intercomparison of Land-surface Parameterization Schemes (PILPS) was initiated in 1992 with the aim of improving LSMs especially in respect to hydrological, energy and momentum flux exchange descriptions (Henderson-Sellers et al., 1993). In a study within the PILPS project, 16 LSMs were run with the same forcing data for the Red and Arkansas River basins in the United States and model output compared. The study showed a large spread in the partitioning between evaporation and runoff, with runoff coefficients (i.e. quotient of runoff to precipitation) ranging from 0.02 to 0.41, compared to an observed quotient of 0.15 (Wood et al., 1998). Other intercomparison studies of LSMs have also shown large differences in the partitioning between evaporation and runoff (e.g. Polcher et al., 1996; Dirmeyer et al., 2006), indicating that structural and parameter differences play an important role for model results.

The Water Model Intercomparison Project (WaterMIP) was launched as a joint effort of the Water and Global CHange (WATCH) and the Global Soil Wetness Project (GSWP) to compare results from both LSMs and GHMs. In a study with six LSMs and five GHMs being run with the same forcing data, Haddeland et al. (2011) found that the models differed substantially in estimates of global water fluxes. Global mean land evaporation ranged from 415 to 586 mm yr^{-1} and runoff from 290 to 457 mm yr^{-1}. The LSM group showed a larger spread in global runoff coefficients compared to the GHMs, but both the mean and median of the LSMs were lower than that for the GHMs. They also found that for most regions the degree-day method employed by the GHMs resulted in higher mean snow-water equivalents (i.e. more snow storage) compared to the energy-balance approaches employed by the LSMs.

The authors conclude that, due to the complexity of the models, it is not feasible to explain all the differences noted between them and that studies, such as climate-impact studies, should not be based on a single model. Another study within WaterMIP, predicting climate-change impacts using seven different models, found that the differences between the models were more important than the differences between the emission scenarios or between the global climate models used for forcing since they resulted in different directions of change (Voß et al., 2008).

In all, there exist large differences between models, both between the LSMs and GHMs and within each group. These differences have an effect on model output, which manifests itself in the broad ranges of results using the same forcing data.

Uncertainties in hydrological modelling

Models are simplifications of a complex reality: “All models are wrong, but some are useful” (Box, 1979). Modelling is, as a result, inherently uncertain. The nature of these uncertainties can be either aleatory, i.e., random and caused by natural variability, or epistemic, i.e. systematic and caused by lack of, or ignorance of, knowledge. In addition, these uncertainties can change characteristics over time, e.g. with changes in monitoring networks or measurement techniques and when using models in different temporal or spatial domains. These uncertainties affect model performance and the inferences we can draw from model outcomes (paper **I**). Uncertainties in hydrological modelling stem mainly from four sources: input data, evaluation data, model structure and model parameters (e.g. Engeland et al., 2005; Refsgaard and Storm, 1996).

Input data

Precipitation

Precipitation has been measured with gauges for centuries (Strangeways, 2010), but one of the main features of precipitation is the large variability both in time and space which makes it difficult to capture the spatio-temporal variability accurately with such point measurements (Michaelides et al., 2009). The quality of precipitation data is strongly dependent on the gauge density, and the density needed differs depending on the local climate and topography. In addition to quality problems caused by insufficient gauge coverage, the quality is also affected by measurement errors (Strangeways, 2004; Tapiador et al., 2012). These errors include wind-induced undercatch, which can be especially pronounced in snow affected areas, and sublimation and evaporation from the gauge itself (Groisman and Legates, 1994). Errors for point measurements depend on the type of gauge used, but undercatch is typically 5–16% and can be much larger for snow measurements (McMillan et al., 2012). However, these errors are often overshadowed by interpolation/extrapolation errors, when estimating areal precipitation (McMillan et al., 2012).

Remote-sensing data, in terms of radar and satellite observations, provide a means to overcome some of the difficulties in capturing the spatio-temporal variability with gauge networks, but are affected by uncertainties stemming from limitations in both measurement techniques and algorithms (Michaelides et al., 2009). Several global precipitation datasets therefore merge remote-sensing data and ground-based observations to combine the relative merits of each type of product, e.g. the Global Precipitation Climatology Project (GPCP) datasets (Adler et al., 2003).

Precipitation is the most significant variable in water-balance calculations since it strongly determines the quality of the discharge predictions (Vörösmarty et al., 1989; 1998; Wisser et al., 2010). In a sensitivity study including six different precipitation datasets, Fekete et al. (2004) showed that uncertainties in precipitation data were transferred to the output with at least the same magnitude.

Evaporation

Despite its importance as a major driver of droughts (Seneviratne, 2012; Sheffield et al., 2012), current estimates of terrestrial evaporation are highly uncertain (Mueller et al., 2013). In hydrological modelling, actual evaporation is often simulated using some functional relationship between actual evaporation and available moisture and potential evaporation, which can differ in complexity (Zhao et al., 2013; paper **III**).

There exist a number of different formulas to estimate potential evaporation with very different complexity and data demand. For instance, Widén-Nilsson et al. (2007) used a simple function only requiring air temperature and relative humidity as input. Other methods, such as the Penman-Monteith combination formula (Monteith, 1965), may require a much larger number of input variables such as wind speed, temperature, radiation, albedo, humidity and more. Uncertainties are introduced in potential evaporation estimates both by uncertainties of the variables that the estimates are calculated from, but also by the choice of formula with its inherent assumptions. Studies of the performance (in terms of resultant performance from forcing of a GHM or based on criteria relating to observed basin water budgets) of different formulas have shown that no formula consistently performs better than the others (Federer et al., 1996; Vörösmarty et al., 1998; Arnell, 1999; Lu et al., 2005).

The land-cover impact on potential evaporation can be substantial, but many formulas neglect the effect of vegetation. For instance many researchers use the Food and Agricultural Organization (FAO) reference-crop evapotranspiration as potential evaporation, although it is supposed to represent the maximum evaporation from a hypothetical reference crop with a height of 0.12 m, an albedo of 0.23 and a surface resistance of 70 s/m (Allen et al., 1994). In global hydrological modelling it is not uncommon to use formulas that do not take land cover into account (e.g. Döll et al., 2003; Widén-Nilsson et al., 2007; 2009), or to consider effects only for irrigated crops (e.g. Wisser et al. 2010), although some consider vegetation type explicitly (e.g. Gosling and Arnell, 2011).

Other data

Recent advances in remote-sensing technologies have contributed to a tremendous increase in global-scale datasets available for large-scale modelling, including high-resolution hydrography datasets. However, the accuracy of these latter products depends on the topographic relief, with flat regions being more prone to errors (Lehner et al., 2008). In addition, for water-balance studies the topographic delineation can cause problems where the groundwater divides do not follow the topography (Eli, 1998).

Anthropogenic effects on the global water system are far reaching: dams alone are estimated to impact a majority of the world's large river systems (Nilsson et al., 2005). Datasets on reservoirs and irrigation have been developed, but are connected to uncertainties, e.g. in storage volume estimates (Lehner et al., 2011) and due to inconsistent definitions of irrigated land between data sources (Siebert et al., 2010). In addition, reservoir/irrigation operation schemes are often not known (Hanasaki et al., 2006).

Other sources of uncertainties in input data are for example static and coarse maps of vegetation and land use, but these likely have minor effects in comparison to the uncertainties in precipitation inputs.

Evaluation data

Discharge constitutes the most important variable for calibration and validation of hydrological models. It represents the aggregated response of the basin to precipitation input and reflects storage processes in the system. At the global level, the Global Runoff Data Centre (GRDC) in Koblenz, Germany, serving as one of the most important archives of discharge data records, likely holds the most complete compilation (Fekete and Vörösmarty, 2002). Nonetheless, both metadata and discharge data themselves are affected by uncertainties and those are poorly documented (paper **II**).

There exist many methods for discharge measurement, which are suitable for different conditions (e.g. flow rates and channel properties). Typically, stage is measured continuously and discharge is only measured occasionally to derive and update stage-discharge relationships. This introduces uncertainties not only in the measurements themselves (stage and discharge), but also through the rating curve used to derive the discharge from the stage measurements. Errors in the stage measurements are usually neglible compared to other factors affecting the discharge estimates (Baldassarre and Montanari, 2009). McMillan et al. (2012) provide typical relative error ranges of ±50–100% for low flows, and ±10–20% for medium to high flows based on a review of studies dealing with observational errors of daily discharge estimates. Such uncertainties affect model calibration and evaluation and should be taken into account when making inferences of model performance.

Other types of data have also been used for model calibration/evaluation, for instance groundwater data (e.g. Döll and Fiedler, 2008) and water storage data (e.g. Güntner, 2008) and new datasets developed from remote-sensing measurements like satellite altimetry show promise for extension of global discharge databases (Tourian et al., 2013). However, these data are also connected to issues of uncertainties and temporal or spatial coverage.

Model related uncertainties

Both model structure (i.e. the parameterisations used to simplify the complex hydrological processes occurring in nature) and model parameters cause uncertainties in the predicted variables of interest. The complexity of models predicting runoff on global and continental scales differs greatly (Sood and Smakhtin, 2014; paper **III**) and as has been shown in model-intercomparison studies, these structural differences can have a major effect on the partitioning of precipitation between evaporation and runoff (e.g. Haddeland et al., 2011).

On the global scale, anthropogenic alterations of the water system have large effects on the hydrological regimes (Nilsson et al., 2005). These effects are of great importance to incorporate in any large-scale model, but limited data make parameterisations of these processes uncertain (Wisser et al., 2010; Biemans et al., 2011). Overall, model specific uncertainties, e.g. due to simplifications of known processes or lack of knowledge of important processes, can have significant impacts on model output (e.g. Butts et al., 2004; Müller Schmied et al., 2014).

Uncertainty analysis

Since uncertainties are inherent in any modelling exercise, these need to be appropriately accounted for and disseminated (paper **I**). However, when dealing with global data the knowledge of the uncertainties affecting different datasets is often limited, both in terms of their nature and in terms of their magnitude. In addition, in many regions model evaluation is not possible due to lack of data, which introduces further complications in the uncertainty assessment of global models.

Uncertainty analysis has been largely missing in large-scale hydrological modelling (Haddeland et al., 2011; Sood and Smakhtin, 2014) and to some extent in much of the hydrological modelling (Beven, 2009). The aforementioned limitations of data and especially of metadata could be one of the reasons for the lack of uncertainty estimates. There is also a popular belief that the general public and decision makers want, and need, simple answers (Frewer et al., 2003; paper **I**).

Uncertainty analyses need to be carried out throughout the modelling chain, starting with scrutinisation of the input data. Beven et al. (2011) and Beven and Westerberg (2011) discuss the case when uncertainties in data actually render their combination disinformative. Disinformative data can result in biased or even wrong inferences from models. However, distinguishing informative data from disinformative ones is not straightforward. Beven et al. (2011) suggest that one way to detect disinformative input data is to analyse the consistency between data and classify physically inconsistent ones as disinformative. Such a screening should increase our possibility to be "right for the right reasons", but also to decrease the risk of being "wrong for the wrong reasons" (paper **I**; **II**).

Regionalisation

The ungauged basin

Many basins in the world are ungauged, i.e. lack sufficient (in terms of quantity or quality) observed data to make adequate hydrological analyses on the spatial and/or temporal scales of interest (Sivapalan et al., 2003). Fekete et al. (2002) report that approximately 50% of the continental landmass is ungauged. However, the ungauged proportion will depend on which time period is being considered and the observations might not be available at the scale of interest in a particular study (e.g. a sub-basin might lack observations although the basin outlet is gauged).

The global discharge-monitoring network has declined substantially over the last three decades in terms of number of gauging stations (The Ad Hoc Group et al., 2001; Fekete and Vörösmarty, 2002; Shiklomanov et al., 2002). In addition, there are significant delays in the reporting of new data (Fekete and Vörösmarty, 2002).

The scarcity of high-quality observed data in many regions of the world and the need for predictions of water resources in these areas therefore require some transfer of knowledge gained from data-rich basins to data-poor basins, so called regionalisation.

Traditional regionalisation

Regionalisation methods can be divided into two main groups: i) methods aimed at transferring model parameters calibrated in gauged basins to ungauged basins and ii) methods aimed at transferring signatures or indices of hydrological behaviour from gauged to ungauged basins. The former has a long-standing tradition in hydrology (e.g. Abdulla and Lettenmaier, 1997; Peel et al., 2000; Döll et al., 2003; Xu, 2003), but in more recent time there has been a shift towards the latter (Wagener and Montanari, 2011). Whether it is model parameters or hydrological indices/signatures that are being regionalised, the fundamental basis for regionalisation is that predictions need to be made without direct calibration. Instead, regionalisation builds on the assumption that basins with similar climatic and physiographic characteristics will have similar hydrological regimes.

The traditional regionalisation strategies of model parameters can be divided into the following main groups: i) spatial proximity, ii) *a priori* parameter estimation, iii) regression, and iv) hydrological similarity methods.

The simplest method for regionalisation is to assume that basins close to each other have similar parameter values/behaviour. Information is then simply transferred between basins in close spatial proximity (e.g. Vandewiele and Elias, 1995), i.e. the spatial proximity method (often referred to as the nearest neighbour method). However, basins can behave very differently although they are geographically close (Post et al., 1998; Beven, 2000).

A priori parameter estimates based on physical characteristics are common in global-scale hydrological studies as a way to avoid the need of calibration in ungauged basins (Arnell, 2003). However, the *a priori* definition of model parameters is not straight forward, since many model parameters are effective parameters and subjected to commensurability issues between measured properties and model parameters (Beven, 2000).

Regression methods include multivariate regression between calibrated model parameters from a pool of gauged basins and physical properties of the basins. These resulting regression models have then been used to derive parameter values for ungauged basins (e.g. Döll et al., 2003). Many of the studies using regression approaches have shown that the correlation between model parameters and basin characteristics can be low (e.g. Peel et al., 2000), which may be related to problems with data quality, over-parameterisation, and equifinality (Beven, 1993; 2006), but also to problems in identifying the important basins characteristics.

The last method, hydrological similarity, encompasses a variety of approaches to estimate the hydrological similarity between basins and transfer entire parameter sets rather than individual parameters as is done in the previous method (e.g. McIntyre et al., 2005; Reichl et al., 2009).

Regionalisation of hydrological indices

One big drawback of the traditional regionalisation methods is that they are model specific. In addition, many models suffer from poorly identifiable parameters and this can render regionalisation virtually impossible (Kuczera and Mroczkowski, 1998). One of the main aims of the IAHS initiative PUB was to improve the understanding of how hydrological functioning depends on catchment characteristics (Sivapalan et al., 2003; Hrachowitz et al., 2013) and this coordinated effort by the scientific community contributed to the change of focus on model-dependent parameter-transfer studies to studies predicting the dynamic behaviour of basins. Knowledge of this expected behaviour can in itself be utilised to reduce predictive uncertainties in ungauged basins through indirect calibration (Wagener and Montanari, 2011).

The predicted behaviour in terms of signatures can be used to constrain model parameters in ungauged basins (Westerberg et al., 2014).

Hydrological signatures have been regionalised in similar ways as model parameters. For example, Yadav et al. (2007) used regression models to estimate a number of signatures, including a base-flow index, high pulse count (a measure of the frequency of high flows) and slope of the flow-duration curve, in ungauged basins. Westerberg et al. (2011) showed that predicted uncertainty could be reduced in two hydrological models through calibration against flow-duration curves. In a later study, Westerberg et al. (2014) regionalised flow-duration curves in Central America based on proximity in the basin characteristics space and used the regionalised curves to constrain the predictive uncertainties of a hydrological model.

Global regionalisation studies are rare, but Beck et al. (2013) used neural networks to estimate two signatures (a base-flow index and a base-flow recession constant) worldwide based on a large set (3,394) of catchments covering wide ranges of conditions in terms of climate, physiography and hydrology.

Heterogeneity and regionalisation

Basins can exhibit a pronounced heterogeneity in terms of hydrological behaviour. Even so, no generally accepted catchment classification exists in hydrology (Wagener et al., 2007; Hrachowitz et al., 2013). Nonetheless, several studies have shown that different types of catchment groupings can improve regionalisation results. Laaha and Blöschl (2006) compared four different grouping techniques for 325 basins in Austria and showed that grouping of catchments based on seasonality performed the best in terms of regression regionalisation of a low-flow index (95^{th} percentile of daily flows). Sauquet and Catalogne (2011) found that grouping French catchments based on a regression tree approach on catchment characteristics considerably improved regression models for prediction of parameters describing flow-duration curves.

Data

The analyses in this thesis were based on freely available global data from a variety of sources. Observational climate and hydrographic data used in paper **II** are listed in Table 1 along with NWP model output and observational data for paper **IV**. Discharge data were retrieved from GRDC in June 2011 for paper **II** and in June 2013 for paper **IV** and **V**.

Table 1. *Data used in paper II and IV.*

Dataset	Temporal resolution	Spatial resolution	Reference
		Basin delineation	
DRT	N/A	0.5°	(Wu et al., 2012)
GIS polygons	N/A	~15″	(Lehner, 2012)
STN-30p	N/A	0.5°	(Vörösmarty et al., 2000b)
DDM30	N/A	0.5°	(Döll and Lehner, 2002)
		Observed precipitation	
CRU TS 3.10.01	Monthly	0.5°	(Harris et al., 2014)
GPCC v6	Monthly	0.5°	(Becker et al., 2013)
GPCP v2.1	Monthly	2.5°	(Adler et al., 2003)
$WATCH_{CRU}$	Daily	0.5°	(Weedon et al., 2011)
$WATCH_{GPCC}$	Daily	0.5°	(Weedon et al., 2011)
		Potential evaporation	
CRU TS 3.10.01	Monthly	0.5°	(Harris et al., 2014)
$WATCH_{PM}$	Daily	0.5°	(Weedon et al., 2011)
$WATCH_{PT}$	Daily	0.5°	(Weedon et al., 2011)
		Actual evaporation	
LandFlux-EVAL synthesis	Monthly	1°	(Mueller et al., 2013)
		NWP model output[1]	
ERA-40	Monthly	T159 (~1.125°)	(Uppala et al., 2005)
ERA-Interim	Monthly	T255 (~0.75°)	(Dee et al., 2011)
ERA-CM	Monthly	T159 (~1.125°)	(Hersbach et al., 2013)
O_{H1d}	Monthly	T1279 (~0.25°)	
O_{L10d}	Monthly	T639 (~0.5°)	
$Land_C$	Monthly	T255 (~0.75°)	(Balsamo et al., 2013)
$Land_U$	Monthly	T255 (~0.75°)	

1. Precipitation, actual evaporation and runoff were retrieved from each of the ECMWF model setups.

In paper **I**, three gridded flow networks and a GIS-polygon dataset (Lehner, 2012) were used for basin area definitions. The gridded products were

DDM30 (Döll and Lehner, 2002), STN-30p (Vörösmarty et al., 2000b) and an early version of the datasets developed by Wu et al. (2012) using the dominant-river-tracing technique (Wu et al., 2011), herein referred to as DRT. Gridded climate data (precipitation and potential evaporation) were retrieved from the Climate Research Unit (CRU; Harris et al., 2014), the Global Precipitation Climatology Centre (GPCC; Becker et al., 2013) and the WATCH project forcing data (Weedon et al., 2011). Both CRU and GPCC are gauge-based products, but the number of gauges in the GPCC product (~67,200) far exceeds that of the CRU product (~11,800) (Becker et al., 2013). The WATCH dataset contains two products, which were both derived from reanalysis data (ERA-40) but bias corrected using observed data (CRU and GPCC, respectively). Both products, from here on $\mathrm{WATCH}_{\mathrm{CRU}}$ and $\mathrm{WATCH}_{\mathrm{GPCC}}$, were corrected for gauge undercatch using correction factors from Adam and Lettenmaier (2003). Potential evaporation data were retrieved from CRU (FAO reference-crop Penman-Monteith estimates) and WATCH (both FAO reference-crop Penman-Monteith estimates and Priestley-Taylor estimates, from here on $\mathrm{WATCH}_{\mathrm{PM}}$ and $\mathrm{WATCH}_{\mathrm{PT}}$, respectively).

In paper **IV**, precipitation data from the merged satellite and rain gauge monthly precipitation-analysis product, version 2.2, from the Global Precipitation Climatology Project (GPCP; Adler et al., 2003) and the diagnostic evaporation-data product from the LandFlux-EVAL multi-data set synthesis (E_{LFE}; Mueller et al., 2013) were used as observational data together with discharge. The model-output data consisted of precipitation, evaporation, and runoff retrieved from seven different setups of the ECMWF model system.

The basis for the ECMWF modelling system includes an atmospheric general circulation model, which is coupled to an ocean-wave model, a land-surface model and a circulation model for the oceans (ECMWF, 2013). The analysis included data from the operational models, three reanalysis products, and two offline runs of the land-surface scheme H-TESSEL.

The operational model outputs were taken from the high-resolution deterministic short-range forecast (24 hour from the analysis, $\mathrm{O}_{\mathrm{H1d}}$ herein) and the unperturbed member of the lower-resolution ensemble forecast with a longer lead time (10 days from the analysis, $\mathrm{O}_{\mathrm{L10d}}$ herein).

The reanalysis data consisted of ERA-40 (Uppala et al., 2005), ERA-Interim (Dee et al., 2011) and ERA-CM (Hersbach et al., 2013), which represents different reanalysis generations with ERA-40 being the oldest and ERA-CM the newest. ERA-CM consists of an ensemble of 10 members and does not include any assimilation of atmospheric variables as opposed to the former two, which are deterministic and use full data assimilation.

The two offline runs were performed without data assimilation using raw ERA-Interim forcing in one of the runs and corrected precipitation forcing in the other (Balsamo et al., 2013). For the corrected run (herein $\mathrm{Land}_{\mathrm{C}}$, with uncorrected as $Land_U$), monthly precipitation totals were scaled to match totals in a predecessor of the GPCP product herein (for details see Balsamo et al., 2010).

In paper **V**, a range of climatic and physiographic basin descriptors, listed in Table 2, were used to evaluate similarity between basins and for regionalisation of parameters.

Table 2. *Basin descriptors used in paper V.*

No	Descriptor	Explanation	Source
	Climate		CRU apart from S_F (WATCH)
1	AI	Aridity index (P/E_P)	
2	P	Mean annual precipitation	
3	P_{SI}	$P_{SI} = P^{-1} \sum \left\| P_m - \frac{P}{12} \right\|$ where P_m is mean monthly precipitation	
4	E_P	Mean annual potential evaporation	
5	$E_{P,SI}$	As for P_{SI} but for E_P	
6	SEAS	Correlation between mean monthly P and E_P	
7	S_F	Fraction of precipitation falling as snow	
8	T	Mean air temperature	
	Topography		ISLSCP apart from A (GRDC)
9	Emean	Mean elevation	
10	Erange	Elevation range (max. elevation – min. elevation)	
11	Slope	Mean slope	
12	A	Basin area	
	Land cover		ISLSCP
13	Forest	% forest cover	
14	Shrub	% shrub cover	
15	Grass/crop	% grass cover	
16	Water	% water bodies	
	Soils		ISRIC-WISE
17	Gravel	% gravel	
18	Sand	% sand	
19	Silt	% silt	
20	Clay	% clay	
	Gauge Location		GRDC
21	Lat	Latitude in degrees	
22	Lon1	sin(Longitude in degrees + 180)	
23	Lon2	cos(Longitude in degrees + 180)	

Climate characteristics were calculated from the CRU TS 3.10.01 dataset (Harris et al., 2014) for the 1960–1990 standard period and included: the aridity index (AI), mean annual precipitation (P), seasonality index for precipitation (P_{SI}), mean annual potential evaporation (E_P), seasonality index for potential evaporation ($E_{P,SI}$), correlation between mean monthly precipitation and potential evaporation (SEAS) and mean air temperature (T). Climatic mean monthly precipitation correction factors derived by Legates (1987) were used to account for systematic errors in the precipitation data. Since no distinction is made between solid and liquid precipitation in the CRU dataset, the $WATCH_{CRU}$ dataset was used to derive the fraction of precipitation falling as snow (S_F).

Topographic and land-cover data were retrieved from the International Satellite Land-Surface Climatology Project (ISLSCP) II data collection, specifically the HYDRO1k Elevation-derived Products (Verdin, 2011) and the University of Maryland Land Cover classifications (DeFries and Hansen, 2010). Soil texture data (fractions of gravel, sand, silt and clay) were derived from the International Soil Reference and Information Centre - World Inventory of Soil Emission Potentials (ISRIC-WISE) version 3 dataset (Batjes, 2005). All three datasets on topography, land cover and soils were available at $0.5^\circ \times 0.5^\circ$ latitude-longitude.

In addition to the basin characteristics listed in Table 2, data on location and capacity of large dams and reservoirs were obtained from the Global Reservoir and Dam (GRanD) database version 1.1 (Lehner et al., 2011).

Methods

Pre-processing of discharge and climate data

Discharge data were screened for obvious errors, before any analyses were performed, and erroneous data, e.g. assignment of an incorrect missing data indicator (e.g. 9999 instead of the correct indicator -999) or apparent decimal errors (magnitude difference), were discarded.

In paper **II**, climate data were missing for some cells that were defined as land in the basin delineations. These cells were assigned the average of the closest eight surrounding cells. If data were still missing, this was repeated in an iterative manner until all land areas were covered. For the calculation of basin precipitation and evaporation based on the GIS-polygon dataset, the intersections with the 0.5˚ climate data grid cells were used to determine the fraction of cell precipitation or evaporation that contributed to the basin (Figure 1). Sub-grid variability was not taken into account, i.e. precipitation and potential evaporation were assumed to be evenly distributed over each grid cell.

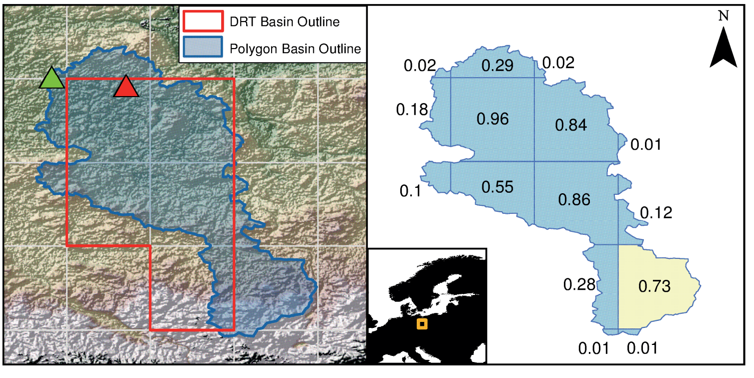

Figure 1. Example of treatment of gridded climate data for the polygon basin delineations. Basin outline based on DRT and GIS-polygon data for Berlin Mühlendamm UP discharge station on the Spree River (9,707 km^2) overlaid on 0.5° climate data grid (left). Basin intersections with the climate grid cells labelled with the fraction of the intersecting grid area (right). For each climate grid cell, only the intersecting fraction contributes to the basin average: e.g. for the yellow polygon 73% of the precipitation falling in the climate grid cell is assumed to fall within the basin. The red triangle indicates the location of the discharge station according to the GRDC archive and the green triangle the location after corrections made in the generation of the GIS-polygon dataset.

In paper **IV**, the resolution of different data sources differed (Table 1) and they were therefore bilinearly interpolated to a common 0.5°×0.5° latitude-longitude grid to facilitate comparisons. Basin averages for climate data were calculated based on the basin polygon intersections with the climate data as in paper **II**, but with the difference that Thiessen polygons based only on the land cells of the ECMWF landmask were used instead of grid cells to calculate the overlap. This was done to avoid any issues with mismatches between land and sea definitions in the different datasets.

Hydrography representation of basin area

An accurate description of the spatial extent of basins is important for the results of hydrological analyses. The performance of three different gridded hydrographic datasets, at 0.5°×0.5° latitude-longitude resolution, and one GIS polygon layer generated from 15″ hydrography, was evaluated in terms of representation the basin areas archived in the GRDC database. For this comparison, gauging stations need to be co-registered in the gridded networks, i.e. each gauging station must have been allocated to the most appropriate grid cell. For two of the gridded products, such co-registrations were already available with 663 GRDC stations for STN-30p (Fekete et al., 2002) and 1,235 GRDC stations for DDM30 (Hunger and Döll, 2008). A mixed automatic and manual registration procedure was performed for the third product.

First, each station was assigned to the cell corresponding to the latitude and longitude given for the station in the GRDC archive. In the second step, the station was automatically reassigned to any of the eight surrounding cells if its flow-accumulation area better matched the basin area in the GRDC archive. Finally, all stations that still displayed a relative area error of 10% or more after the automatic reassignment were manually inspected and reassigned if appropriate. The relative area error, ε_A, was defined as:

$$\varepsilon_A = \frac{A_{Acc} - A_{GRDC}}{\max(A_{Acc}, A_{GRDC})} \cdot 100\% \qquad (1)$$

A_{Acc} is the flow-accumulation area of the grid cell assigned to the station and A_{GRDC} the archived basin area. The same measure has been used previously (Fekete et al., 1999; Döll and Lehner, 2002), but is referred to as “symmetric error” in those studies. Cell areas for the gridded products were calculated as quadrangles based on the World Geodetic System 1984 ellipsoid.

Evaluation of consistency between model forcing and evaluation data

The method used to detect inconsistent data was similar to that of Beven et al. (2011), who identify disinformative data as those that violate the water balance. However, here the analysis was based on the long-term water balance rather than their event-based approach. In addition, the analysis also included transgressions of the potential-evaporation limit, similar to Peel et al. (2010). Since the long-term water balance was analysed, changes in basin storage could be ignored since these would only be significant for special cases such as melting glaciers. On the long term, the water-balance equation can be simplified to:

$$P = E_A + R \tag{2}$$

where P is precipitation, E_A is actual evaporation, and R is runoff. Two fundamental assumptions formed the basis for the consistency check between the forcing data and the evaluation data: i) For natural basins, runoff should not exceed the precipitation input to the system, and ii) actual evaporation, inferred as the difference between precipitation and runoff, should not exceed the potential evaporation (E_P).

The datasets used in the analysis are all affected by uncertainties of different types and from different sources. However, due to limited metadata and knowledge, it was not possible to estimate their nature and magnitude either temporally or spatially. Therefore, a relative uncertainty of ± 10% was assumed for the observed discharge data, resulting in a low, a high, and a sharp (i.e. the original discharge values) estimate for each discharge record. The climate data were used as they were.

The long-term quotient between runoff and precipitation, i.e. the runoff coefficient (RC), was calculated for each of the three discharge estimates for all basins with at least 10 years of data (the threshold was set after analysing the variation in RCs with regard to record length). The RCs were used to evaluate the first of the two criteria for the consistency between data, that runoff should not exceed precipitation input, i.e. the RC should not be higher than unity.

No screening of anthropogenic influences, such as reservoirs or interbasin transfers, was performed before the consistency checks. For some basins affected by such influences, the water-balance equation according to Eq. 2 should not be expected to be fulfilled, but these data would still be disinformative in any modelling unless anthropogenic effects are explicitly treated in the model.

Analysis of land-surface water budgets in an NWP scheme

Similarly to the analysis of consistency between observational datasets in paper **II**, the long-term water budgets were the focus of the analysis of the NWP model output in paper **IV**.

Firstly, global precipitation and land-surface evaporation were benchmarked against observational data, GPCP and E_{LFE} respectively. The balance between global precipitation and evaporation was analysed as a first measure of internal consistency or inconsistency of the models, since these are expected to be in approximate balance over a year or longer.

Long-term runoff was benchmarked against observed runoff for 611 basins worldwide as an indication of the hydrological representativity of the land-surface schemes. Basins were selected based on size (≥10,000 km^2) and available data for the common period (1995–2001) of the model outputs. Basins with implausible RCs (greater than unity) for the observational precipitation data were excluded from the analysis in line with paper **II**.

Lastly, the internal consistency of the models was analysed based on the long-term water balance equation for all land cells and on the basin scale. Specifically, evaporation should not exceed precipitation over the long term and runoff should balance the difference between precipitation and evaporation.

Regionalisation of flow-duration curves

For all basins with discharge records with at least 10 consecutive years of monthly average flows in the GRDC archive and basin areas of at least 5,000 km^2, an initial screening for inconsistent data was carried out allowing for a 10% uncertainty in the long-term average discharge following paper **II**. Empirical flow-duration curves were constructed based on observed data, for the records that passed the screening, by ranking each flow of the period of record in descending order so that the highest flow received a rank, i, of 1 and the lowest one a rank of N (corresponding to the number of observations in the record). Each of these ranked flows, q_i, was subsequently assigned with an exceedance probability, p_i:

$$p_i = P(Q > q_i) = \frac{i}{N+1} \tag{3}$$

For purposes of evaluation, ten exceedance percentiles (5, 15, 25, 35, 45, 55, 65, 75, 85 and 95%) were chosen for analysis of performance of fitted and regionalised FDCs. The 5% highest and lowest flows were excluded as these are likely the most uncertain (Westerberg et al., 2014).

The prediction methods of the regionalised FDCs comprised of two main approaches: direct estimation of regionalised FDCs (direct methods) and regionalisation of parameters of different approximations fitted to the empirical FDCs. Both the empirical FDCs and log-transforms thereof were tested in the fitting procedure. The log-transformed FDCs were used to limit the effects of high flows in the fitting (observed values smaller than 0.001 $m^3 s^{-1}$ were replaced with 0.001 $m^3 s^{-1}$ to avoid log-transforms of 0).

All basin descriptors in Table 2 were standardised by subtracting the mean and dividing by the standard deviation for each descriptor prior to the use in the different regionalisation approaches.

Direct methods

The simplest method used to regionalise the FDCs was the spatial proximity (nearest neighbour, NN) approach, which was used as a baseline minimum information case. Two other proximity methods were also tested based on the similarity of basin descriptors in Table 2: proximity in the catchment descriptor space (P1) and proximity in the catchment descriptor space including geographical proximity (P2). The regionalised FDCs were calculated following the method by Westerberg et al. (2014), which is shortly summarised here.

Proximity was calculated as the Euclidian distance in the descriptor/geographical space. An initial correlation analysis between the basin descriptors in Table 2 and the ten exceedance percentiles led to the following descriptors being adopted for the similarity calculation: AI, P, P_{SI}, Slope and Forest. Each FDC was standardised by dividing the raw data with the basin area, i.e. specific discharge was used, and a jack-knife approach, i.e. each basin in turn was considered ungauged, for the regionalisation.

Either the one closest or the ten closest basins (in the geographical and/or descriptor space sense) were used as donor basins. In the case with one donor, each basin was assigned the specific discharge FDC of the donor catchment and this was subsequently multiplied by the basin area of the ungauged basin to obtain the regionalised FDC. For the case where 10 donors were used, the regionalised FDCs were calculated based on a fuzzy membership function weighted according to the proximity between the donor and the ungauged basin (see Westerberg et al., 2014 for details of the methodology). The fuzzy membership function requires that an uncertainty band be attached to the observed FDCs for the calculations. In this thesis, due to limited knowledge of the real uncertainty in the discharge values from the GRDC archive, a generic rather conservative uncertainty was assumed for all FDCs: a relative uncertainty of ±20% for flows that were exceeded less than 50% of the time, and a linearly increasing uncertainty from ±20% for the median flow to ± 75% for the lowest flow. For flows smaller than or equal to

1 m^3 s^{-1} the uncertainty was assumed to be ± 0.75 m^3 s^{-1} (with a lower limit of zero) since relative uncertainties do not work well for small numbers.

The final direct estimation method builds on artificial neural networks (ANNs), which are computational models originally developed to mimic the data processing and learning by the brain (i.e. the neurons). Today ANNs are used for machine learning in a multitude of fields including hydrology. For instance, Beck et al. (2013) use them to estimate a base-flow index and recession constant worldwide and Mazvimavi et al. (2005) use ANNs for FDC and parameter estimation.

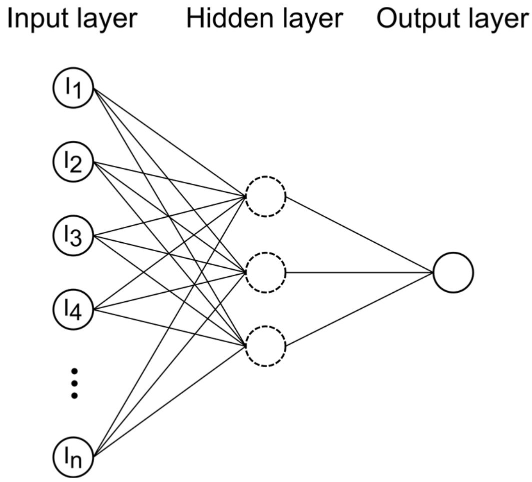

Figure 2. Schematic of a neural network. The input layer consists of input neurons $I_1, I_2, I_3, \ldots I_n$. In this thesis, n was 23 and each neuron corresponded to one of the 23 basin descriptors listed in Table 2. The hidden layer consists of a number of hidden neurons, here exemplified with three neurons (indicated with dashed lines), which are connected to each of the input neurons through weights, which are indicated by the lines. The final layer is the output layer, which is connected to the hidden layer and here exemplified with one neuron. In this thesis, the output neurons corresponded to either exceedance percentile flows or parameters of the FDC approximations.

In this thesis, the neural network architecture used was a three layer feed forward network (Figure 2) with the Levenberg-Marquardt back propagation learning algorithm, which has shown high efficiency (Hagan and Menhaj, 1994). The first layer, the input layer, consisted of 23 neurons corresponding to the 23 basin descriptors in Table 2. The second layer consisted of hidden neurons, which are necessary for the network to be capable of modelling nonlinear relationships, and the number of neurons in this layer was determined by trial and error. The final layer, the output layer, consisted of either one or ten neuron(s). In the case of using one neuron, each exceedance percentile flow was estimated in turn whereas in the ten neurons case all were estimated at the same time. Flows were log-transformed before training for normalisation purposes and back-transformed after training.

The networks were trained on a subsample of the data by adjustment of weights linking the input layer to the hidden layer and the hidden layer to the output layer so that the root-mean-square error between the output (predicted exceedance percentile flows) from the network and the targets (empirical exceedance percentile flows) was minimised. The data (basin descriptors and observed FDCs) were first split into two subsamples. The first subsample, containing 10% randomly selected basins, was completely excluded from the development phase of the networks (unseen data for the networks) to allow a proper evaluation of the networks. The remaining 90% of the data (development data) were used as training, validation and test data for the networks. A ten-fold cross-validation procedure was performed meaning an ensemble of networks was trained on 80% of the development data in turn and the performance, in terms of R^2-values for the test data (10% of the development data), was used to determine the number of hidden neurons needed. Early stopping was used to avoid overfitting, i.e. the training procedure was stopped when the error in the validation data (10% of the development data) started to increase. The median output of the ensemble of ten cross-validation networks was used as the prediction of the exceedance percentile flow(s).

FDC approximations

In many regionalisation studies, the empirical FDCs are approximated by some type of parametric function and regression models are developed between the fitted parameters and catchment descriptors to allow for prediction of parameters and ultimately FDCs in ungauged basins. Many analytical functions and statistical distributions have been suggested as approximate models for FDCs (e.g. Mimikou and Kaemaki, 1985; Castellarin et al., 2004).

In this study, three analytical functions, four statistical distributions and empirical orthogonal functions (EOFs) were tested for their ability to reproduce the empirical FDCs. The analytical functions included the exponential model (Eq. 4), the power law model (Eq. 5) and the logarithm model (Eq. 6):

$$\hat{Q}_p(i) = a(i)e^{b(i)p} \qquad a \geq 0, b < 0 \tag{4}$$

$$\hat{Q}_p(i) = a(i)p^{b(i)} \qquad a \geq 0, b < 0 \tag{5}$$

$$\hat{Q}_p(i) = a(i) + b(i)\ln(p) \qquad a \geq 0, b < 0 \tag{6}$$

$\hat{Q}_p(i)$ is the flow (in $m^3 s^{-1}$, either logged or in real data space) for basin i and exceedance percentile p, and a and b are parameters to be fitted according to the constraints given for each function. Parameters were estimated using ordinary least squares on the 10 exceedance percentiles of the empirical FDCs. Initially 2nd and 3rd order polynomials were also considered, but were excluded as it was not feasible to define parameter constraints that guaranteed monotone and non-negative curves in the parameter regionalisation procedure.

The statistical distributions included the two-parameter lognormal and gamma distributions (Eq. 7 and 8), which were fitted on the entire record of data using the method of maximum likelihood (referred to as LN2 and GA2 herein). In addition, a mixed version of each (Eq. 9 and 10) was included for which the no flow threshold parameter, p_0, was determined from the observed data and the respective distribution was fitted on the non-zero data:

$$\hat{Q}_p(i) = F^{-1}(1-p;\mu(i),\sigma(i)) \qquad \begin{matrix}\mu \in R \\ \sigma > 0\end{matrix} \tag{7}$$

$$\hat{Q}_p(i) = G^{-1}(1-p;k(i),\theta(i)) \qquad \begin{matrix}k > 0 \\ \theta > 0\end{matrix} \tag{8}$$

$$\hat{Q}_p(i) = \begin{cases} F^{-1}\left(1-\frac{p}{1-p_0(i)};\mu(i),\sigma(i)\right) & 0 \leq p \leq p_0(i) \\ 0 & p \geq p_0(i)\end{cases} \qquad \begin{matrix}\mu \in R \\ \sigma > 0 \\ p_0 \geq 0\end{matrix} \tag{9}$$

$$\hat{Q}_p(i) = \begin{cases} G^{-1}\left(1-\frac{p}{1-p_{0(i)}};k(i),\theta(i)\right) & 0 \leq p \leq p_0(i) \\ 0 & p \geq p_0(i)\end{cases} \qquad \begin{matrix}k > 0 \\ \theta > 0 \\ p_0 \geq 0\end{matrix} \tag{10}$$

F^{-1}, μ and σ are the inverse cumulative distribution function, the mean and the standard deviation of the lognormal distribution, respectively. G^{-1}, is the inverse cumulative distribution function, k is the shape parameter, and Θ is the scale parameter of the gamma distribution.

Finally, empirical orthogonal functions (EOFs) were used as a way to reproduce the FDCs. In the EOF analysis, the entire dataset of FDCs was decomposed to a limited number of patterns which explained as much of the variance in the dataset as possible. Each pattern, called basis function, was found by calculating the eigenvectors of the covariance matrix of the dataset. EOFs have been used in hydrology previously, e.g. for estimation of runoff, rainfall and temperature at ungauged sites (Rao and Hsieh, 1991) and for extrapolation of runoff records (Hisdal and Tveito, 1993). Sauquet and Catalogne (2011) showed that EOFs reproduced 1,080 FDCs from catchments in France better than a number of commonly used analytical functions (of which several are tested in this study). As in Sauquet and Catalogne (2011), the EOF analysis in this thesis describes the FDC at site i as a linear combination of orthogonal basis functions:

$$\hat{Q}_p(i) = \gamma(i) + \sum_{m}^{M} \alpha_m(i)\beta_m(p) \tag{11}$$

M corresponds to the number of exceedance percentiles considered (i.e. 10 in this application), and α_m is the m^{th} weight associated to the m^{th} basis function β_m. γ is a parameter related to the magnitude of the FDC since the basis functions are defined to have zero mean.

Fitted parameters for all FDC approximations were regionalised using stepwise linear regression (SR) and ANNs. The ANN architecture and development was the same as for the networks used to predict exceedance percentile flows but in this case the ANNs were trained on individual parameters of the approximations. To reduce the skewness, the parameters to be predicted were Box-Cox transformed (Box and Cox, 1964) prior to the training of the networks and output was back-transformed after training.

Two basin groupings were used to explore if grouping of basins prior to parameter estimation or EOF decomposition could improve prediction results. The first grouping represents a low-information case: basins were grouped only on the main Köppen-Geiger classes based on Peel et al. (2007), i.e. tropical, arid, temperate, cold and polar climates.

The second grouping followed the method of Sauquet and Catalogne (2011), who group catchments based on the so called index of concavity (IC). In their study, IC was calculated as the quotient between the 10^{th} percentile minus the 99^{th} percentile and the 1^{st} percentile minus the 99^{th} percentile, but since the highest and lowest 5% of the flows are excluded in this thesis, a modified IC for each basin i was calculated as:

$$IC(i) = \frac{Q_{15}(i) - Q_{95}(i)}{Q_5(i) - Q_{95}(i)} \tag{12}$$

IC ranges between 0 and 1 where a value close to zero represents catchments with low storage capacity and/or highly variable climate and values close to unity represent catchments for which the runoff response is dampened by large storage capacities (Sauquet and Catalogne, 2011). The grouping was performed using a regression-tree approach with IC as dependent variable and the basin descriptors in Table 2 as decision variables.

Results

Hydrography representation of basin area

At the time of collection of discharge data from the GRDC for paper **II** (June 2011), there were 7,763 gauging stations in the archive. After exclusion of gauging stations with insufficient metadata (missing basin area or coordinates), basin areas smaller than 5,000 km^2 or insufficient discharge data (no daily data available), 2,177 stations remained that were co-registered in the DRT flow network. Of those, 558 stations were available as co-registered stations in the DDM30 and STN-30p datasets and allowed for comparisons of the performances of the three datasets (Figure 3).

DDM30 displayed less scatter (standard deviation of the relative area error, $\varepsilon_{A,}$ 8.9%) compared to DRT (14.6%) and STN-30p (14.3%). Apart from a few largely under- and overestimated station areas in DDM30 and STN-30p, there was little consistency between the errors of the different datasets (Figure 3d–f).

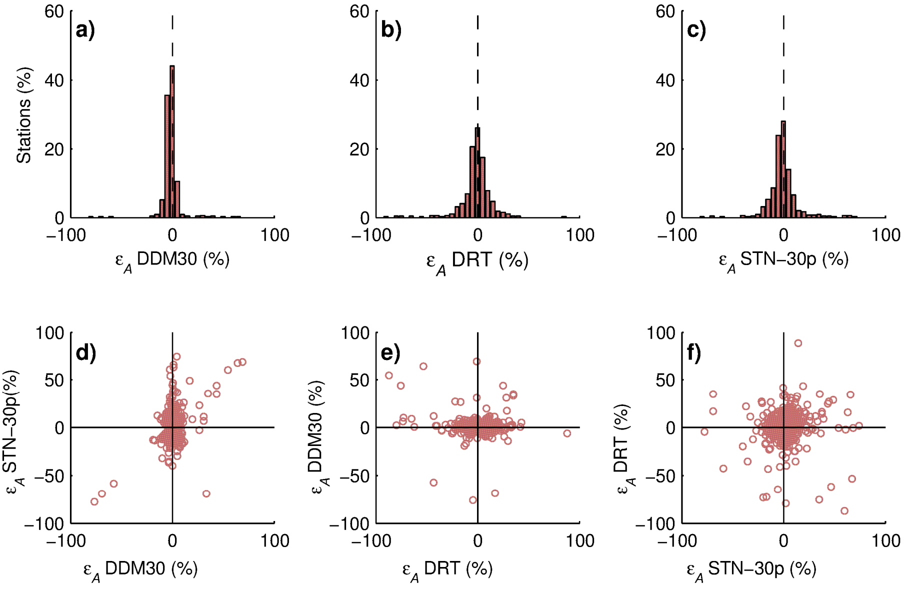

Figure 3. Histograms of relative area errors for the three gridded hydrographies: DDM30 (**a**), DRT (**b**) and STN-30p (**c**). Lower panel (**d**–**f**) shows comparisons of the relative area error for each basin in the different hydrographies.

Of the 2,177 stations co-registered in GRDC, 2,005 were available in the GIS polygon dataset too. The GIS dataset displayed small errors compared to the DRT dataset (though some stations showed marked differences compared to the archived areas) and there was little consistency between the errors in the datasets (Figure 4a–b). Visual inspection showed that none of the four datasets analysed displayed any spatial pattern of the relative area errors.

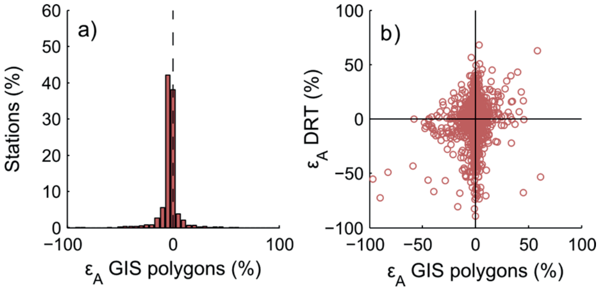

Figure 4. Histogram of relative area errors for the GIS polygon dataset (**a**) and comparison of the relative area error for each basin in the DRT and the GIS dataset (**b**).

Evaluation of consistency between model forcing and evaluation data

Based on the results from the analysis of basin-area representation, only the GIS dataset was used in the evaluation of the consistency between forcing and evaluation data to minimise effects of area errors. Long-term runoff coefficients could be determined for 1,561 of the basins available in the polygon dataset and the general distribution of RCs did not differ much between datasets (Figure 5).

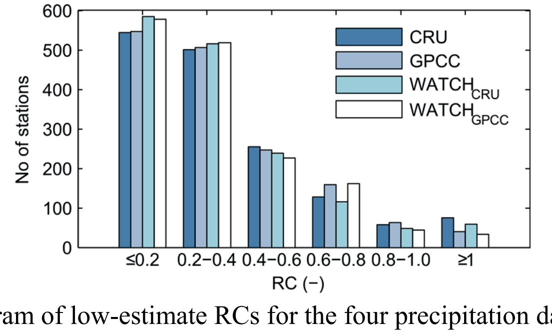

Figure 5. Histogram of low-estimate RCs for the four precipitation datasets.

RCs violating the first consistency criteria, i.e. RCs higher than unity, were found for all precipitation datasets even when using the conservative low RC estimate. Implausible RCs were more common for CRU and $\text{WATCH}_{\text{CRU}}$

than for the other two datasets. Basins with RCs higher than unity were mainly found in Alaska and north-western Canada.

A simplified version of the Budyko (1974) curve was used to graphically analyse the second consistency test. Inferred actual evaporation (P-R) was plotted against potential evaporation for each basin and for all combinations of precipitation and potential evaporation datasets.

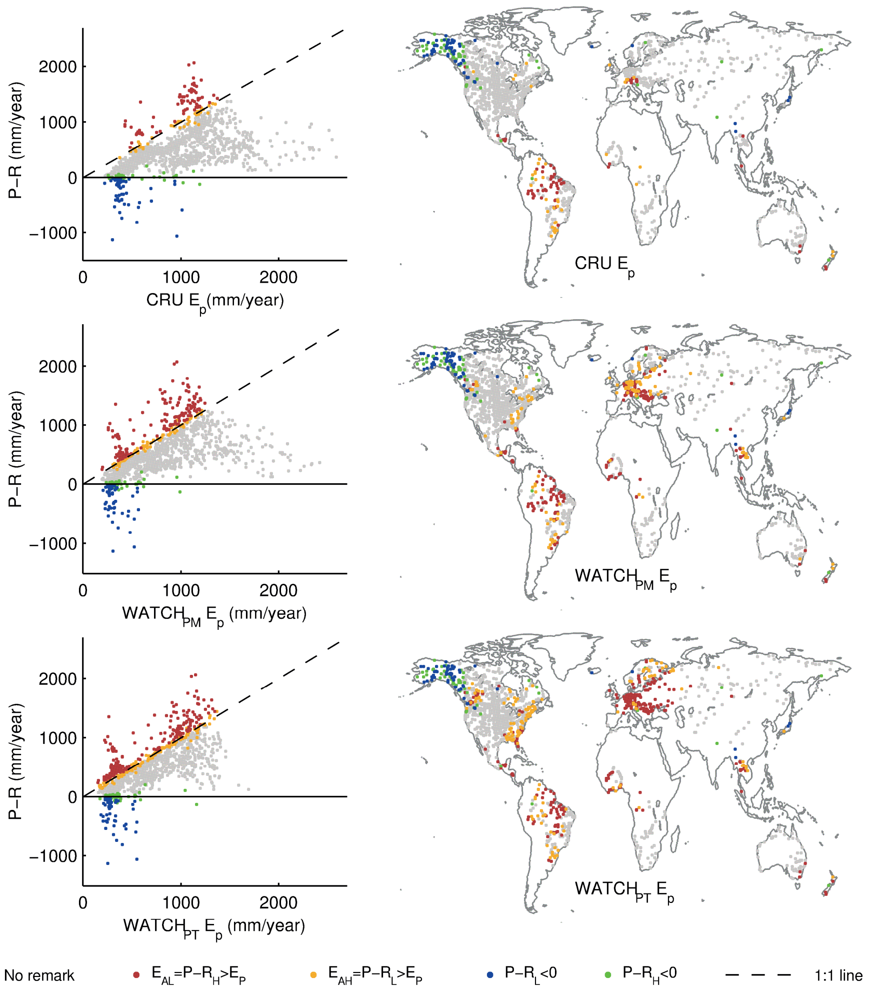

Figure 6. Mean annual actual evaporation (estimated as P-R using CRU precipitation data) versus potential evaporation from CRU, WATCH Penman-Monteith and WATCH Priestley-Taylor (left panel). Potential evaporation is plotted against actual evaporation estimated using the sharp runoff estimate, i.e. the y-value of each dot represents the sharp evaporation estimate. The colour coding is based on the high runoff estimate (R_H, giving low estimate of E_A) and low runoff estimate (R_L, giving high estimate of E_A) as indicated in the legend. The right panel shows the geographical distributions.

The uncertainty of the observed runoff was accounted for as a color-coding of each basin (example in Figure 6), where red corresponds to basins for which the inferred actual evaporation exceeds potential evaporation even for the high runoff estimate (i.e. for the conservative, low, estimate of actual evaporation). Orange corresponds to basins for which the inferred actual evaporation exceeds the potential evaporation for the high or for the high and sharp estimate of actual evaporation, but not for the low estimate. Basins were also colour coded for violations of the first consistency criteria: blue dots represent basins where the actual evaporation was negative (i.e. RC>1) even for the high (low) actual evaporation (runoff) estimate and green if this occurred only for the low estimates of actual evaporation.

Analysis of the graphs showed that the frequency of basins with inferred actual evaporation estimates higher than potential evaporation was noticeably higher for the two WATCH datasets compared to CRU (example shown for CRU precipitation in Figure 6 and Table 3). For all three potential evaporation datasets, too high actual evaporation frequently appeared in the Amazon basin, and for the WATCH datasets frequently on the east coast of North America, in Europe, equatorial Africa and South East Asia (Figure 6).

Table 3. *Percent of basins exhibiting potential data inconsistencies. A basin is only accounted for in the worst category that applies to it, e.g. if the lowest actual evaporation estimate exceeds E_P it is accounted for in column $E_{AL}>E_P$, but not $E_{AH}>E_P$.*

Precipitation	Potential Evaporation	No remark	$E_{AL}>E_P$	$E_{AH}>E_P$	$P-R_L<0$	$P-R_H<0$
CRU	CRU	85.6	4.5	2.9	3.6	3.4
CRU	$WATCH_{PM}$	71.7	12.2	9.1	3.6	3.4
CRU	$WATCH_{PT}$	62.6	19.8	10.6	3.6	3.4

Analysis of land-surface water budgets in an NWP scheme

All model outputs displayed positive biases in precipitation compared to the observational benchmark GPCP (Figure 7). ERA-40 showed a strong bias that increased during the period whereas the strong biases noted in the operational models decreased during the period. For the operational models, the high precipitation rates led to high evaporation rates, but not in ERA-40 (Figure 7d–f). The high uncertainties in current global land evaporation estimates are shown by the wide range of the estimates in the observation-based E_{LFE} dataset (Figure 7e). Although precipitation and evaporation were expected to be in approximate balance over the global domain, substantial differences were found for all models but ERA-CM (Figure 7g).

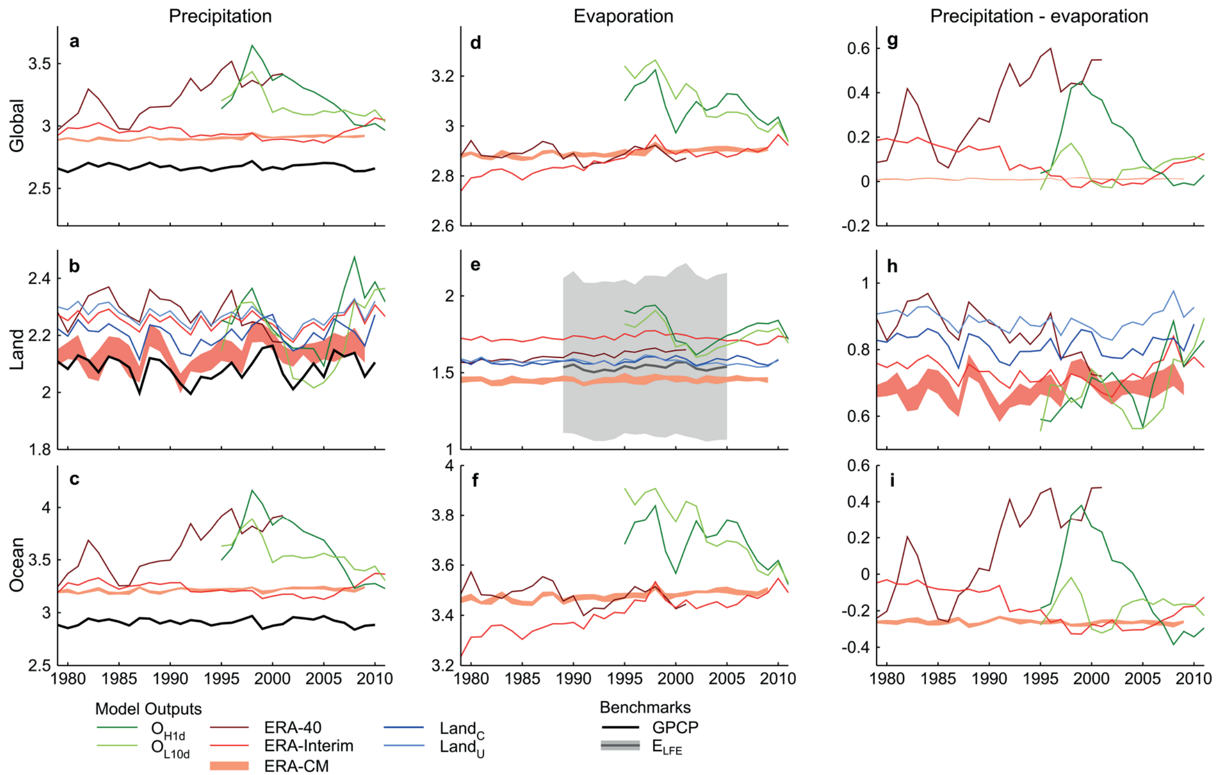

*Figure 7***.** Global annual average precipitation and evaporation rates (mm d^{-1}). Rates of average precipitation (P), evaporation (E_A) and P-E_A for the global domain (top), land areas (middle) and oceans (bottom). The shading for ERA-CM and E_{LFE} denotes the span of the ensembles. The land area in **e** is the one for which E_{LFE} data were available.

The long-term (1995–2001) averaging allowed travel times for flow to be neglected in the analysis of runoff and model estimates were benchmarked against observed discharge for 611 basins (Figure 8).

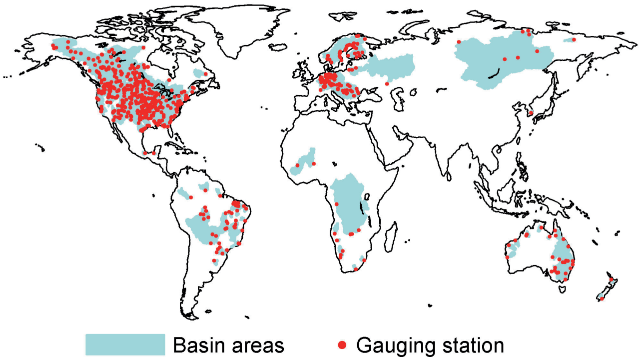

Figure 8. Geographical coverage of the 611 basins in the study.

Substantial differences in terms of relative errors between modelled and observed runoff were found for many basins (Figure 9). For all models, a majority of the basins displayed relative errors larger than ±20%: ERA-40 83% of the basins, ERA-Interim 75%, ERA-CM ensemble members 61–65%, O_{H1d} 76%, O_{L10d} 73%, $Land_C$ 62% and $Land_U$ 56%.

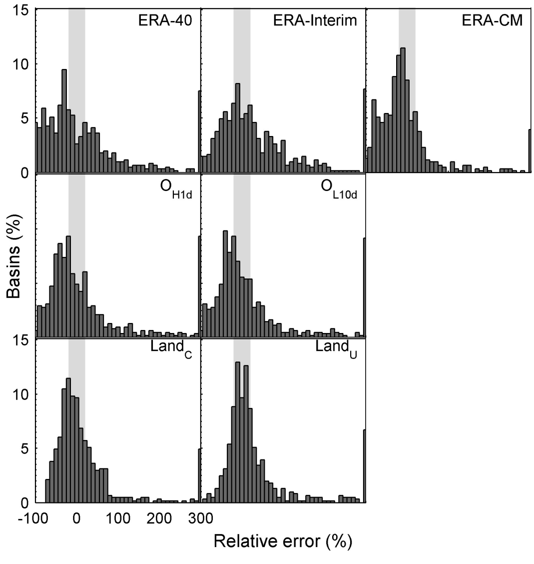

Figure 9. Relative errors between simulated and observed runoff 1995–2001. Error calculations for ERA-CM were based on the median long-term runoff of the 10 ensemble members. Light grey shading corresponds to ±20% error band.

In addition to the global water-budget imbalances (Figure 7), land-surface water-budget imbalances were found both on cell and basin scale when analysing individual models (example shown in Figure 10). For extensive areas, ERA-Interim exhibited long-term evaporation exceeding precipitation (Figure 10a) and similar patterns were found for the operational models and ERA-40. ERA-CM also displayed this issue but to a smaller extent.

Runoff should equal the difference between precipitation and evaporation over long-term (assuming storage changes can be neglected, Eq. 2). However, imbalances (i.e. $E_A>P$ and $P-E_A \neq R$) were seen for large areas in ERA-Interim (Figure 10b) and translated to basin scale (Figure 10c).

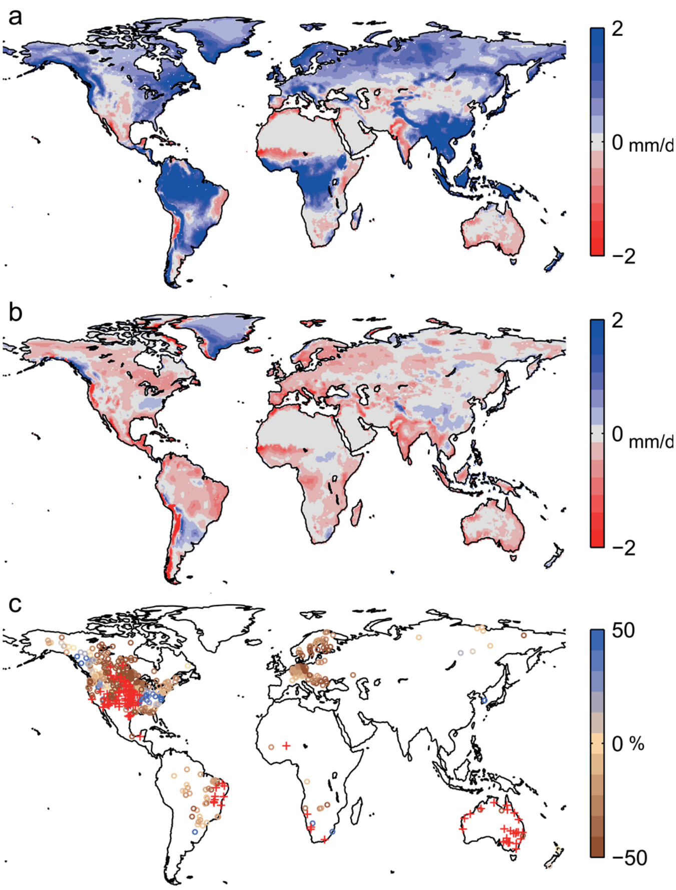

Figure 10. ERA-Interim water budgets 1995–2001. Precipitation minus evaporation (**a**), precipitation minus evaporation minus runoff (**b**), relative errors between simulated precipitation minus evaporation and runoff (**c**). Red crosses in **c** indicate basins for which long-term evaporation exceeds precipitation.

ERA-40 performed the worst in terms of percent of basins that exhibited long-term evaporation more than 10% higher than precipitation (Table 4).

For the model setups using full data assimilation (ERA-40, ERA-Interim and the operational models), 44–66% of all basins displayed an imbalance between runoff and the difference between precipitation and evaporation that exceeded 10% of the average precipitation.

Table 4. *Percent of basins with water budget issues.*

Dataset	$E_A/P>1.1$	$\|P-E_A-R\|/P>0.1$
ERA-40	16	66
ERA-Interim	10	50
ERA-CM	0–2	5–9
O_{H1d}	12	55
O_{L10d}	8	44
$Land_C$	0	0
$Land_U$	0	0

Regionalisation of flow-duration curves

The basins (n=2,102) included in paper **V** (Figure 11) showed a large spread in mean runoff values (0.1–5,500 mm year^{-1}) and included 138 (6.6%) basins that were dry for 10% of the time or more. Basin areas varied between 5,000 and 4,700,000 km^2, with 21,000 km^2 being the median area. The basins covered all main Köppen-Geiger climates following the classification by Peel et al. (2007).

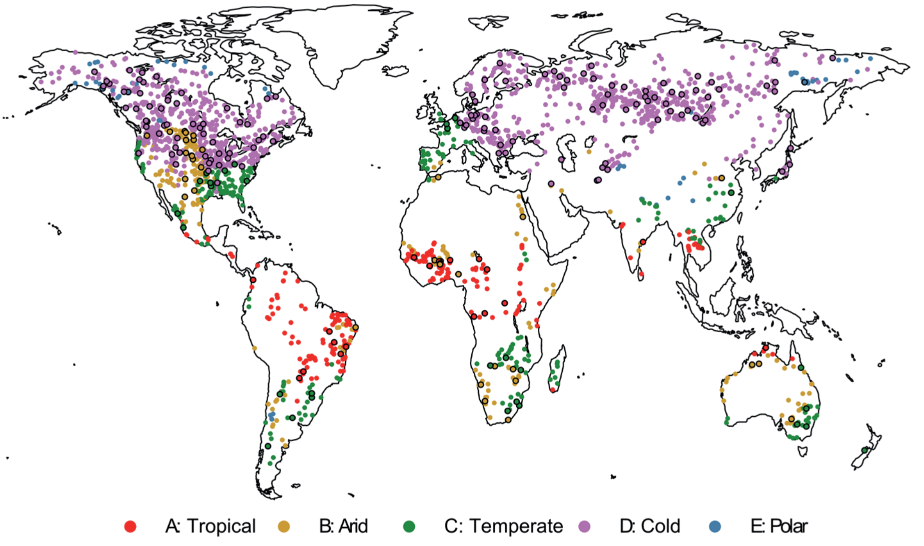

Figure 11. Gauging stations included in the study colour coded according to dominant main Köppen-Geiger classes of the basins. Stations with black outlines belong to the randomly selected 10% of evaluation basins.

Direct methods

Direct estimation of the FDCs using proximity measures (NN, Pv1 and Pv2), resulted in high relative errors for many basins and for all exceedance percentiles (Figure 12a). Pv1 and Pv2 only performed slightly better than NN. Using the 10 closest (in geographical and/or catchment descriptor space) basins as donors, compared to only the one closest, improved the results somewhat, but resulted in unrealistic non-monotone FDCs for some basins for all proximity measures. Estimation of the FDCs using ANNs (both with one and 10 output neurons) resulted in similar relative errors to the proximity methods (example in Figure 12b), but non-monotone FDCs were generated for all network architectures tested (i.e. one and 10 output neurons and different numbers of hidden neurons).

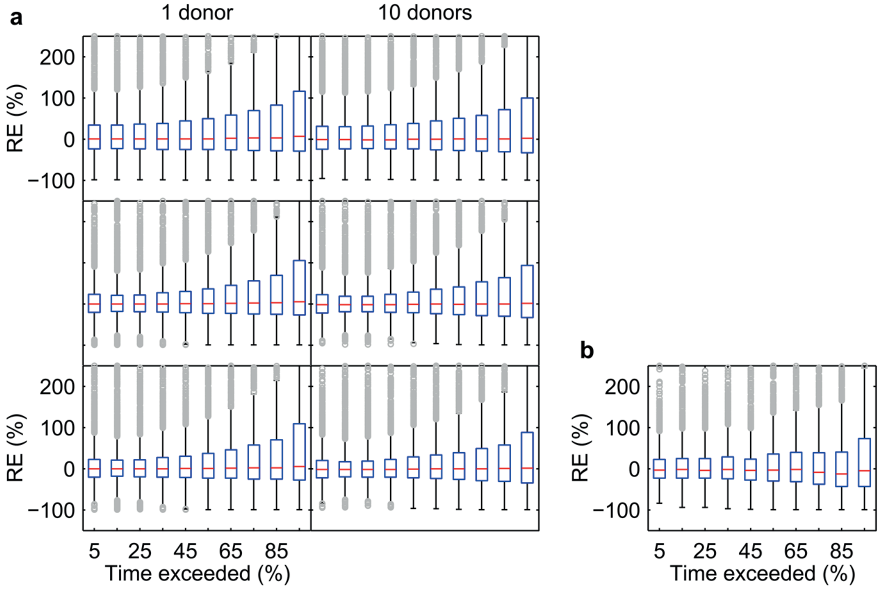

Figure 12. Boxplots of relative errors between empirical and regionalised FDCs: **a**, estimated with proximity techniques, nearest neighbour (NN, top), proximity in the catchment descriptor space (Pv1, middle) and proximity in the catchment descriptor space including geographical proximity (Pv2, bottom) using 1 and 10 donors (left and right panel respectively), and **b**, estimated with ANNs with one output neuron. Red lines in the boxes indicate the median, the edges the 25th and 75th percentile, the whiskers extend to 1.5 times the interquartile range and circles indicate outliers.

FDC approximations

The EOF decomposition led to theoretically unsound (non-monotone) FDCs if more than the first basis function were included in the reconstruction of the FDCs. Therefore, only the first basis function was retained, which explained 95% of the variance of the empirical FDCs and 78% of the log-transformed empirical FDCs.

In terms of relative errors between the empirical and fitted exceedance flows, the lognormal distribution and EOF performed the best (Figure 13). Given the parameter constraints for the exponential, power law and logarithmic model (Eq. 4–6) they are bounded by 0, which resulted in a poor fit of the log-transformed data for basins with low flows. Relaxing the parameter constraints, to encompass negative values in the log-transformed data, resulted in unrealistic (negative) flows in the fitted FDCs and/or unrealistic curves (non-monotone) in the regionalisation.

Based on this initial analysis, EOF (based on log-transformed data), LN2 and LN3 were considered the best candidates for regionalisation.

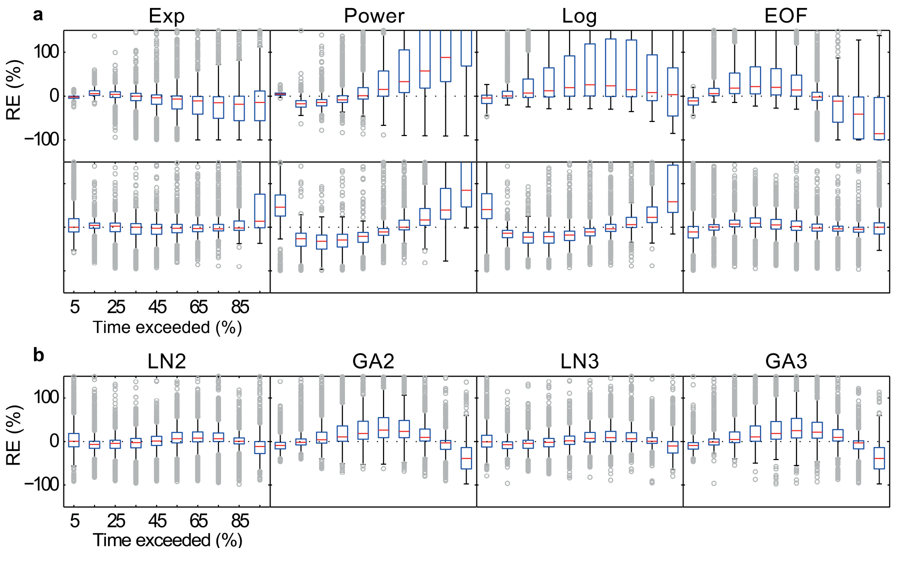

Figure 13. Relative errors for the 10 exceedance percentiles: **a**, for different analytical approximations (left to right) of the empirical FDCs (top panel) and log-transformed empirical FDCs (lower panel), and **b**, for different statistical approximations of empirical FDCs. Boxes are defined as in Figure 12.

The location parameters (γ and μ) were better predicted than the shape parameters (α and σ) for all approximations and by both SR and ANN, but ANNs outperformed SR in parameter estimation in terms of R^2-values and NRMSE for all parameters (example for LN3 in Table 5). This is likely because ANNs do not assume any specific functional relationship between the input and output whilst the SR technique assumes a linear relationship.

However, this flexibility in ANNs mean they can easily be overfitted. The early stopping technique used in this thesis is one method to avoid overfitting, but the generalisation capability of the networks should nonetheless be tested.

Table 5. *R^2-values and NRMSE for parameter estimates with SR and ANN. Numbers in brackets refer to the unseen data.*

Parameter	R^2		NRMSE	
	ANN	SR	ANN	SR
LN3 μ	0.94 (0.94)	0.78 (0.82)	0.12 (0.13)	0.22 (0.21)
LN3 σ	0.74 (0.60)	0.41 (0.42)	0.27 (0.34)	0.41 (0.40)
LN3 p_0	0.82 (0.79)	0. 39 (0.44)	1.75 (1.55)	3.08 (2.55)

In Figure 14, the predicted and fitted parameters for the development data (i.e. basins used as training, testing or validation data in the ANN development process) and the unseen data (i.e. the 10% completely excluded from the development) are shown separately for EOF. The performances for the two subsets was similar (as confirmed by small losses in R^2-values and small increases of NRMSE in Table 5).

Heteroscedasticity is seen for both parameters, but especially the shape parameter (α). The shape parameter for LN2 also shows strong heteroscedasticity. Many of the basins with low α-values for EOF and high σ-values for LN2 are intermittent basins for which the fit of both approximations is generally poor.

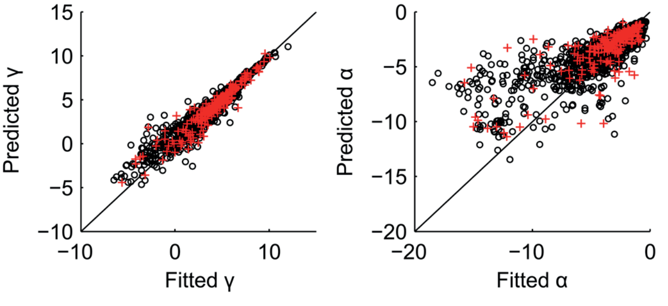

Figure 14. Predicted parameters versus fitted parameters for γ (left) and α (right) of EOF. Each black circle represents a basin that was used in the neural network development (training, validation, testing) and, each red cross one of the basins that were excluded completely from the development phase (unseen data).

The analysis of performance using the two groupings, based on main Köppen-Geiger climates and the regression tree on IC gave mixed results. Cluster size differed for the five main climates (A: 233, B: 256, C: 296, D: 1258 and E: 59) and for clusters based on IC (13 clusters with 105–218 basins in each).

EOF decomposition of the individual groups of the two clustering methods did not result in any marked improvement. However, using the groupings improved the parameter estimates through SR for all parameters apart from the threshold parameter in LN3 when using clustering on the main climates. On the other hand, using ANNs to estimate parameters for individual clusters performed worse than the ANNs trained on the entire dataset. Nevertheless, the ANN estimates with no grouping applied still performed better than SR with grouping. Therefore, the parameters predicted by ANNs trained on the entire dataset were used.

In terms of relative errors for the regionalised curves, all three approximations performed similarly with LN3 representing low flows somewhat better (Figure 15). The results were on par with the direct estimation techniques (Figure 12), but somewhat better for low flows, and with the advantage that all these curves were monotone.

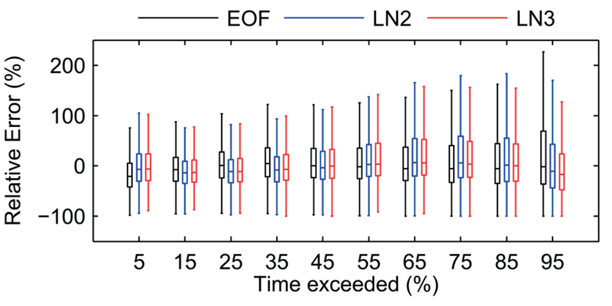

Figure 15. Relative errors of the 10 exceedance percentiles for the regionalised FDCs. Boxes are defined as in Figure 12, but outliers have been removed for better visibility.

Geographically, large mean relative errors for high, intermediate and low flows were scattered throughout the global domain for both the fitted curves and the regionalised ones. The fit of LN3 was good with small relative errors for many basins, but was rather poor in some basins e.g. in eastern and central Russia. The regionalised curves displayed very large relative errors for basins scattered throughout the world, with a tendency to perform poorly in arid regions, e.g. central U.S., and the eastern and central parts of Russia (which also had poor fits).

Examples of the empirical FDCs shown in Figure 16 together with the fitted and regionalised FDCs using LN3, highlight the very mixed performance of the method. In some basins, the fit between the empirical FDC and the fitted FDC was poor (e.g. Alsek River). Some basins exhibited large discrepancies between the empirical FDC and the regionalised one, e.g. the Amazon and Murrumbidgee rivers. Some basins showed very good results on the other hand, e.g. the Danube, Huang He and Rhine River.

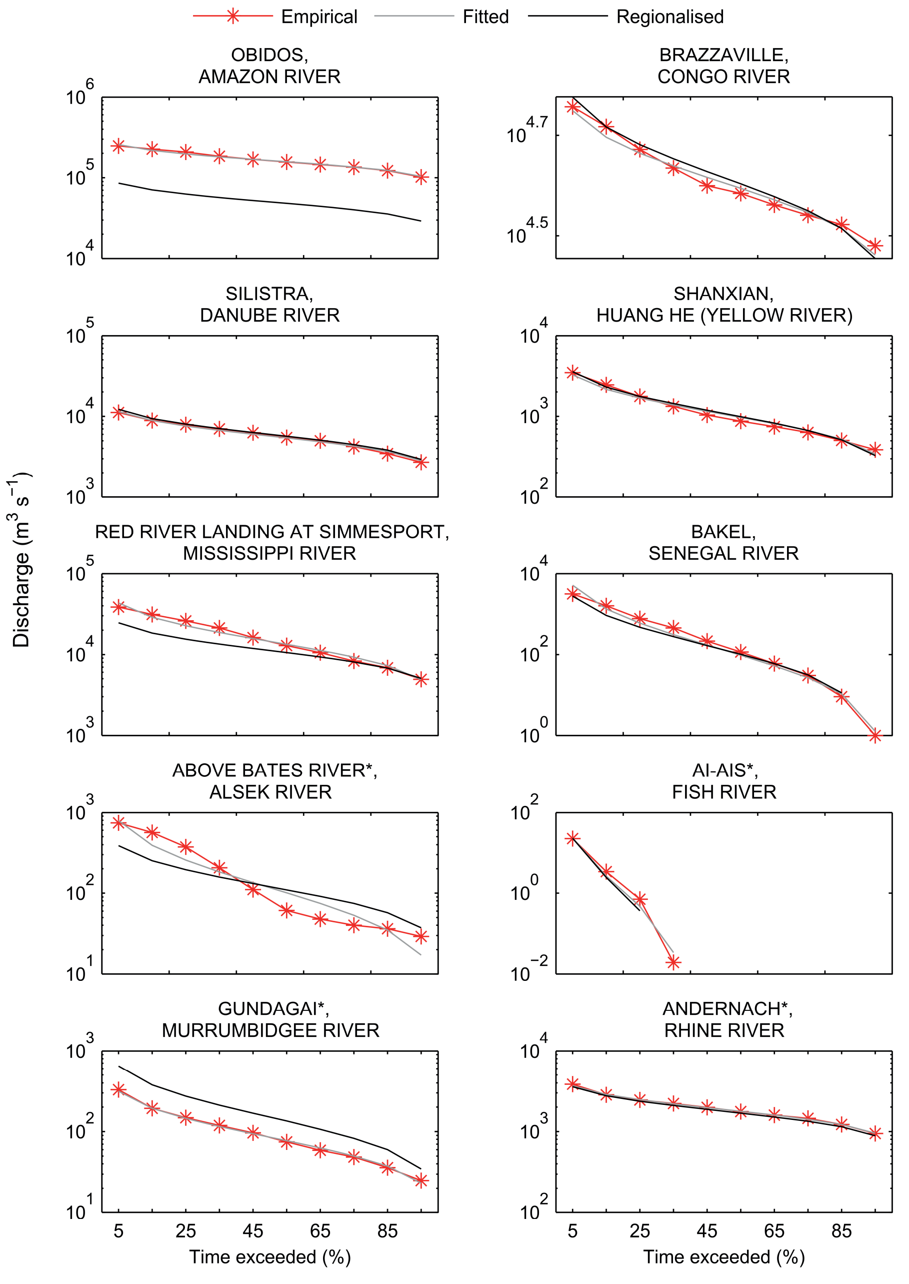

Figure 16. Examples of empirical, fitted and regionalised FDCs. Stations marked with an asterisk were part of the unseen data. Note different ordinate scales.

Discussion

The focus of this thesis was to address challenges in large-scale hydrology introduced by disinformative and uncertain data. In paper **II**, the identification of disinformative data in a number of datasets was carried out through a screening procedure aimed at detecting inconsistencies between datasets. This work was carried on in paper **IV,** in which the internal consistency of a number of model versions of a numerical weather prediction scheme was analysed. Both paper **III** and **V** deal with predictive tools to overcome the lack of measured data, in terms of a technical review of models for flood forecasting in paper **III,** and in terms of spatial extrapolation through regionalisation in paper **V**. The importance of all these aspects of large-scale hydrology is to some extent summarised in the seven reasons for doing uncertainty analysis pointed out in paper **I**.

Representation of basin areas

A correct delineation of the basin is an important starting point for water-balance analyses. The results in paper **II** from the comparisons of basin areas in the hydrographic datasets with the areas reported in the GRDC archive showed that many basins exhibited large relative area errors. However, these discrepancies are likely not only a result of deficiencies in the hydrographic datasets, but also to varying quality of the metadata in the archive. Basin areas in the archive are reported by the data providers and contain no information of their accuracy (U. Looser, head of GRDC, personal communication, October 2011). Comparison of archived basin areas at the time of retrieval from GRDC for paper **II** (June 2011) and the areas reported from the archive in the GIS polygon dataset showed that reported basin areas had changed by over 100% for a few stations.

DDM30 outperformed the other two gridded products in terms of representing the archived basin areas, but due to extensive manual corrections (Döll and Lehner, 2002) it is difficult to use to derive high-resolution topographic basin information, e.g., for routing. The GIS polygon dataset outperformed the gridded dataset DRT in terms of representing basin areas: 94% of the basins displayed absolute relative area errors of 25% or less and 84% displayed errors less than 10%. Corresponding numbers for DRT were 80% and 45%, respectively.

Consistency between forcing and evaluation data

Runoff coefficients higher than unity were encountered for all precipitation datasets in paper **II**, even when conservatively accounting for the uncertainty in runoff. Several other global studies have reported similar problems (e.g. Fekete et al., 2002; Widén-Nilsson et al., 2009; Peel et al., 2010), which can result from a number of issues, e.g., poor representation of spatial and temporal patterns in precipitation, measurement errors, and in some basins inter-basin transfers (Peel et al., 2010).

It is theorised that wind-induced snow undercatch, which can have substantial effects on measurements in high-latitude areas (Adam and Lettenmaier, 2003), is one explanation to the fact that the majority of the basins exhibiting these implausible runoff coefficients were located in snow affected areas. The fact that the WATCH datasets had been corrected for solid undercatch, yet still exhibited this inconsistency, is indicative that the corrections were not sufficient or that other errors (e.g. poor representation of orographic effects) are present.

Transgressions of the potential-evaporation limit also occurred in all datasets analysed. The issue was most pronounced for the WATCH datasets and can most likely not only be attributed to effects of irrigation or inter-basin transfers not accounted for, but to a large part probably stems from inadequate potential evaporation estimates. All three datasets analysed ignore vegetation differences and the results indicate the importance of considering land cover.

When using the sharp discharge estimate, 8–43% of all basins exhibited data inconsistencies depending on how datasets were combined and the corresponding number when accounting conservatively for discharge uncertainties was 6–35%. This indicates that a substantial amount of the data analysed in paper **II** can be disinformative for model evaluation (Beven and Westerberg, 2011; Beven et al., 2011; paper **I**).

Land-surface water budgets in an NWP scheme

Biases and imbalances in the global and land-surface water-budgets were found in the analysis of NWP-model output in paper **IV**. The precipitation rates exceeded the observed for all model outputs. For ERA-40, the strong bias has been attributed to issues in the humidity assimilation (Uppala et al., 2005; Dee et al., 2011). The operational models (O_{H1d} and O_{L10d}) also showed strong biases, but decreasing with time, which is partly explained by an improved convection parameterisation (Bechtold et al. 2012).

The analysis of the evaporation rates showed that the model output fell within the estimates of the diagnostic dataset, but the uncertainties in both observational and modelled data were high as indicated by a large spread in the estimates. Globally, there was an imbalance in precipitation and evaporation of the models (apart from ERA-CM). For ERA-40 and O_{H1d} this was to a large extent due to a spurious exceedance of ocean precipitation over evaporation. Trenberth et al. (2011) have reported similar issues for other reanalysis datasets.

For hydrologically more relevant scales it was found that runoff was biased for a majority of the basins even when allowing for a relatively large uncertainty (±20%) in the long-term mean runoff. In addition, the long-term water budgets did not balance (E>P and P-E≠R). These imbalances are unlikely to be attributed to e.g. glacier melting not considered here, but are more likely an effect of the data assimilation. In the assimilation, soil moisture and snow storages are used as "nudge factors" to improve the forecasts, which results in a disrupted hydrological cycle (Dee et al. 2011).

The results indicate that care must be taken when using these data for hydrological purposes, but the results must also be seen in the perspective of these models being aimed at producing the best possible weather forecasts, not at predicting runoff or at representing global hydrological regimes.

Regionalisation of flow-duration curves

Using direct techniques for regionalisation of FDCs, i.e. proximity in the geographical/descriptor space or ANNs for percentile flow estimation, resulted in non-monotone FDCs unless only one donor was used for the proximity approaches (which resulted in high errors for low flows). The monotonicity issue could be considered negligible if only very low flows were concerned, but the issue appeared for intermediate flows too. Therefore, approximation of the FDCs with a given monotonic model seemed to be a good alternative for prediction of theoretically sound FDCs in this thesis. It was not possible to find one approximation, among the ones tested herein, that encompassed all the different shapes of the FDCs, but LN3 performed relatively well for most basins. Other functions or distributions should be tested to find a more appropriate approximation. However, the biggest loss in performance was found in the parameter regionalisation step.

The regionalisation of the fitted parameters showed that ANNs outperformed SR in terms of R^2-values and NRMSE. Performance of SR was improved by clustering of similar catchments prior to parameter estimation, but performance was still poorer than ANNs. A drawback with ANNs is that they represent somewhat of black box models, offering little insight to how the hydrological regimes are related to basin climate and physiographic characteristics. Although outside the scope of this thesis, some knowledge about the importance of different input variables can be gained by analysis of the trained weights (Olden et al., 2004) which could be interesting to explore. However, whether using ANNs or SR for the parameter estimation, the regionalised FDCs displayed large relative errors for high, intermediate as well as low flows. For instance, the mean discharge of the Amazon River was poorly predicted (μ 10.8 ln $m^3 s^{-1}$ compared to the empirical 12.0 ln $m^3 s^{-1}$), which led to massive relative errors (average -70%). On the other hand, the Murrumbidgee River was overestimated by ~75% due to an overestimation of the mean discharge (μ 5.0 ln $m^3 s^{-1}$ compared to the empirical 4.4 ln $m^3 s^{-1}$). The Murrumbidgee River is heavily impacted by regulation and water diversions, which could partly explain the poor predictions, but for the Amazon River, the reason for the underestimation was harder to explain. This indicated that even though rudimentary sensitivity tests to regulation (exclusion of basins with maximum capacities of 10, 20 and 30% of annual discharge) did not show any marked effects, anthropogenic influences can be very important on individual basin level, and need to be accounted for.

Including more basins (e.g. smaller than 5,000 km^2 or with shorter discharge records) may help in the training of the networks, but it is also possible that there are important characteristics of the basins that have been overlooked in the input data selection. In addition, the basin average values might not contain enough information to distinguish between different hydrological functioning of basins. It is also possible that the parameters that the networks were trained on actually introduced disinformation in the cases where the fit of a particular approximation was poor. For instance, low α-values in EOF were typically indicative of a poor fit in intermittent basins and may add little (or even wrong) information during the training process.

The parameter, and ultimately FDC, regionalisation may possibly improve with some of the aforementioned measures, but nonetheless a high uncertainty will remain in the predictions. Accounting for this uncertainty, including the uncertainty in the empirical FDCs themselves, the uncertainty introduced by the approximations and finally the uncertainty in the parameter estimation, will be of outmost importance for the regionalised FDCs to be useful for e.g. indirect model calibration.

Conclusions

Uncertainties in hydrological modelling and analysis is not an issue reserved for large-scale studies, but is often exacerbated by limited data availability and quality over large domains. The results in this thesis have shown that the uncertainties in global hydrological data can be high, and that these high uncertainties often lead to physically implausible discrepancies when combining data.

The analysis in paper **II** clearly showed that many of the data used in hydrological modelling are inconsistent. The method for screening forcing and evaluation data to detect inconsistencies is in no way limited to large-scale studies, but can be applied at any scale. Performing such a screening prior to analyses can, to some extent, prevent drawing the wrong conclusions in subsequent modelling, by highlighting basins/regions for which model results should be carefully considered before accepting or rejecting them (paper **I**).

However, not only data uncertainties affect global hydrological models, but differences in model structure and parameters can have a major effect on runoff estimates. The many differences between models reviewed in paper **III** have been shown to result in large spreads in model output in model intercomparison studies. This indicates that more data are needed in order to falsify or confirm the process parameterisations and in that way improve global water-resources estimates and reduce the predictive uncertainties.

In the analysis of the land-surface water budgets of the seven ECMWF model versions (paper **IV**), it was clear that they suffered from serious limitations in terms of imbalances between long-term precipitation, evaporation and runoff, but also biases compared to observational data. The data assimilation process led to disrupted hydrological cycles. This limits the usefulness of these data for hydrological purposes and bias correction, although fraught with its own problems, will be necessary until these issues are resolved.

In paper **V**, several issues were identified in the process of prediction of FDCs for a wide range of climatic, physiographic and ultimately hydrological conditions globally. Unrealistic FDCs, i.e. non-monotonic or with negative flows, were produced through several of the methodologies employed. In order to ensure theoretically sound regionalised FDCs, it was found necessary to approximate the empirical FDCs with some monotonic function or distribution. The mixed lognormal distribution was found to perform reasonably well for most basins in terms of reproducing the empirical FDCs. Regionalisation of the parameters through ANNs were found to perform better than SR, but many FDCs were poorly predicted using regionalised parameters.

Acknowledgements

It goes without saying that this work would not have been possible without the help and support by many people around me. First of all, I would like to thank my supervisors Sven Halldin, Allan Rodhe and Chong-Yu Xu. Thank you Sven, for being so supportive and for your invaluable writing advice: the papers would not have been half as good without your red pen and the scholarship applications not nearly as successful. Allan, thank you for your insightful comments on the work I've done, and especially for always being able to take a step back and look at the big picture when I've gotten lost in the swamp of details. Thank you Chong-Yu, for teaching me the joys of statistics, for comments on my work and for the memorable hat party in Vienna!

I would like to thank Florian Pappenberger and Fredrik Wetterhall for the great opportunity to come to the ECMWF (and the Sportsman...) in Reading as a visiting scientist. I learnt a lot and really enjoyed my stay at the centre and came home with both ECMWF kitsch and inspiration to finish off my PhD (the latter perhaps more needed than the former).

I would like to thank my co-authors Ida Westerberg, Hannah Cloke, Peter Salamon, Jutta Thielen, John Juston, Beatriz Quesada, Jan Seibert and Keith Beven. Especially Ida, for the fun and interesting herring seminars and for being a good friend and roommate in Uppsala and elsewhere. I would also like to thank Keith for his comments on several of my papers and for not laughing at me at our first meeting, on my second day or so in Uppsala, when Monte Carlo only referred to casinos in my mind.

Not many people would stand to see my tired face every Monday morning or my loud swearing when battling ArcGIS, but José-Luis, Ida, Liang and Nina have clearly shown some strength in that respect and for such lovely office mates I'm very grateful. José-Luis, I've really missed your "words of wisdom".

Although reading scientific papers is a truly enjoyable pastime, it has been very refreshing to sometimes indulge in other literature and I am very grateful for the book club meetings with Eva, Petra, Johan, Olof and Katrin (even though their choice of books sometimes left a lot to desire - who was responsible for the "masterpiece" by Mr. Noll?).

I would also like to thank all the other colleagues and friends at the department: Agnes, Diana, José, Dennis, Lebing, Marc, Jean-Marc, Andreas, Julia, Maria, Fritjof, Thomas, Claudia, Cecilia, Peter, Simon, Giuliano, Elias, Eduardo, Carmen, Tito, Estuardo, Zhibing, Nino and many more. I would especially like to thank Martin and Alvaro for taking the discussion on whether my humour is bad or not to a new level by getting it in print.

I want to thank Erik, Therese, Henrik, Börje, Maria and Linda in Kalasgruppen for being such a fun gang and for letting me focus on the important aspects of working at Geocentrum: the parties!

Thank you to everyone at LUVAL for making it a fun place to work and special thanks to Veijo Pohjola, Anna Rutgersson and Roger Herbert for making it possible for me to engage in work at the ECMWF and for the AGU technical committee on surface waters. I would also like to thank Thomas Nord for having the patience with me whenever I blew up another computer.

Finally, I would like to thank Rob, my family, the Sturegatan family and all my friends for their support, especially for believing in me all those times when I was in doubt. At last, this thesis can put an end to your sleepless nights wondering what I really spent my time on all these years when I was "doing something with water".

Sammanfattning på svenska
Summary in Swedish

Desinformativa och osäkra data i global hydrologi – utmaningar för modellering och regionalisering

Vatten är en förutsättning för människan och för välfungerande ekosystem, men många människor lever utan tillgång till rent vatten och goda sanitära förhållanden. Globalt sett ökar behovet av färskvatten genom att befolkningen ökar och ekonomierna växer. Även förändringar i klimat och markanvändning förväntas bidra till att belastningen på de redan ansträngda vattenresurserna ökar. För att kunna trygga vattenförsörjningen och därmed matförsörjningen krävs kunskap om dessa resurser. Vid bedömning av vattenresursers storlek är avrinningen i vattendragen en nyckelfråga. Den säger hur mycket och hur fort färskvatten nybildas av nederbörden och därmed hur mycket vatten som finns tillgängligt för människa och ekosystem. Trots den avgörande roll som vatten spelar är våra kunskaper om hur de globala vattenresurserna varierar i tid och rum begränsade och olika uppskattningar av den globala avrinningen uppvisar stora skillnader.

En av utmaningarna med att skatta vattenresurser är att det saknas data av god kvalitet i många områden. Tidiga skattningar av de globala vattenresurserna byggde enbart på mätdata, men under 1980-talet började man utveckla hydrologiska modeller att användas i global skala. Sådana modeller kan användas för att skatta vattentillgången när data saknas, antingen i tid eller i rum, och för projektioner inför framtiden. Globala modeller är utvecklade med ett storskaligt fokus och lämpar sig för studier av kontinental och regional natur, men är inte lämpliga för studier av lokala vattenresurser.

Alla hydrologiska modeller, och kanske särskilt globala hydrologiska modeller, är förknippade med osäkerhet. Detta är oundvikligt vid all modellering då våra matematiska modeller är förenklingar av en komplex verklighet. Den förenklade processbeskrivningen i hydrologiska modeller är en bidragande orsak till osäkerheten i modellresultaten, men även osäkerhet i indata, parameterosäkerhet och osäkerhet i utvärderingsdata bidrar. Trots detta har analyser av dessa osäkerheter och deras konsekvenser för modellresultat nästintill helt saknats i global modellering.

Det finns idag ett flertal modeller som används för storskalig hydrologisk modellering. En del av dessa modeller har sin bakgrund inom meteorologisk forskning och fungerar framförallt som randvillkor för storskaliga atmosfärsmodeller. De utvecklades med fokus på energibalansen snarare än på vattenbalansen. Ofta är den rumsliga upplösningen av dessa modeller låg, men den tidsmässiga relativt hög (<1 dag), och lateral transport av vatten mellan beräkningsrutor saknas. Globala hydrologiska modeller, å andra sidan, utvecklades inom det hydrologiska forskningsfältet och fokuserar helt på vattenbalansen. Med undantag för hybridmodeller är t.ex. snöparametriseringen enklare i de globala hydrologiska modellerna än i de som är kopplade till klimatmodellerna eftersom energibalansen inte simuleras. Stora variationer i komplexitet, upplösning och parametrisering gör de hydrologiska modellerna lämpliga för olika tillämpningar, men ger även upphov till de stora variationer i modellresultat som visats när det gäller skattningar av den globala avrinningen.

Osäkerheten i data påverkar både modellresultat och hur dessa korrekt ska utvärderas. Osäkerheten hos nederbördsdata dominerar ofta osäkerheten i de indata som används till hydrologiska modeller, men osäkerheten i beräknad potentiell avdunstning kan också vara viktig. Oftast utvärderas modellresultat mot observerad vattenföring, men i de många avrinningsområden på jorden där vattenföringsmätningar saknas är sådan direkt kalibrering och utvärdering av modeller inte möjlig.

Denna avhandling berör både osäkerheter i data, genom jämförelser av globala data och modellresultat, och avsaknaden av data genom en inledande studie för regionalisering av varaktighetskurvor. I en studie omfattande flera uppsättningar indata (nederbörd och potentiell avdunstning) och vattenföringsdata från ett stort antal avrinningsområden världen över kunde det fastläggas att 8–43 % av alla avrinningsområden uppvisade desinformativa data (fysiskt orimliga) om den observerade vattenföringen användes vid jämförelsen. Om hänsyn togs till osäkerheten i vattenföringsdata gällde detta för 6–35 % av avrinningsområdena. För alla uppsättningar nederbördsdata observerades avrinningskoefficienter (kvoten mellan avrinning och nederbörd) högre än ett för många avrinningsområden, vilket är orimligt om inte vatten tillförs genom glaciärsmältning eller genom grundvattenflöden som inte tagits hänsyn till. De flesta avrinningsområden med dessa höga avrinningskoefficienter ligger i områden där den uppmätta nederbörden sannolikt är betydligt mindre än den verkliga på grund av snö eller orografiska effekter. Den skattade verkliga avdunstningen (nederbörd minus avrinning) översteg den potentiella i många områden, vilket är fysiskt orimligt och tyder på desinformativa data. En orsak till detta bedömdes vara att den potentiella avdunstningen i de använda datauppsättningarna är beräknad utan hänsyn till vegetationens inverkan. Förekomsten av desinformativa data varierade beroende på vilka data som kombinerades, även om det fanns vissa likartade geografiska mönster för alla kombinationer.

Analysen av vattenbudgetarna i de olika versionerna av den numeriska vädermodellen från det europeiska centret för medellånga väderprognoser

(European Centre for Medium-Range Weather Forecasts, ECMWF) visade stora avvikelser från hydrologiskt rimliga resultat. Globalt sett fanns obalans mellan långtidsvärden av nederbörd och avdunstning och avdunstningen var i många fall större än nederbörden i både modellrute- och avrinningsområdesskala. I många avrinningsområden balanserades inte långtidsvärden för avrinningen av skillnaden mellan nederbörd och avdunstning och den uppvisade för majoriteten av avrinningsområdena relativa fel på över ±20 % av observerad avrinning. Modeller av denna typ är ämnade att producera väderprognoser av hög kvalitet och för att nå detta mål uppdateras modellerna genom assimilering av olika observerade variabler och under denna process används bland annat markvatten som en justeringsfaktor, vilket till stor del kan förklara de obalanserade vattenbudgetarna.

Den sista delen i avhandlingen angriper problemet med avsaknad av vattenföringsdata, vilken förhindrar kalibrering och utvärdering av hydrologiska modeller i många avrinningsområden i världen. För att skatta avrinningen i områden utan mätdata krävs metoder för att överföra kunskap från områden med data till områden utan data, så kallad regionalisering. Traditionellt har man inom den hydrologiska forskningen ofta försökt att regionalisera modellparameterar, som kalibrerats i områden med vattenföringsdata, till områden utan mätdata genom att exempelvis utveckla regressionssamband mellan modellparametrar och olika egenskaper hos avrinningsområden (såsom klimat, geologi, vegetation m.m.).

I avhandlingen utforskas istället möjligheten att regionalisera varaktighetskurvor, vilket har fördelen att de är modelloberoende och kan (om regionaliseringen fungerar väl) användas för indirekt kalibrering av hydrologiska modeller. Ett flertal metoder för regionalisering prövades, både direkt regionalisering av kurvorna och regionalisering av parametrar för funktioner anpassade till kurvorna. Ingen av anpassningarna kunde representera alla de former som de empiriska kurvorna visade och de direkta metoderna resulterade i teoretiskt orimliga kurvor (icke-monotona). Användandet av neurala nät för regionaliseringen av parametrarna gav bättre resultat än multipel linjär regression, men de regionaliserade kurvorna visade i många fall på stora systematiska fel. Även om en del kurvor förutsas väl visade studien att mer forskning behövs både för att förbättra resultaten, men också för att på ett adekvat sätt kunna skatta osäkerheten i de regionaliserade kurvorna.

Resultaten i denna avhandling visar att det finns stor osäkerhet i data för storskalig hydrologisk modellering. Denna osäkerhet kan få en avgörande betydelse i modelleringsprocessen och bör tas i beaktande. Avhandlingen leder till slutsatsen att data bör kontrolleras innan modelleringen sker och att desinformativa data i möjligaste mån bör exkluderas från kalibrering och utvärdering. Fortsatt arbete med hur osäkerhet i indata, utvärderingsdata och modellresultat kan kvantifieras och kommuniceras för storskaliga modeller är en förutsättning för mer robusta och trovärdiga skattningar av de globala vattenresurserna.

References

Abdulla, F.A. and D.P. Lettenmaier (1997). Development of regional parameter estimation equations for a macroscale hydrologic model. *Journal of Hydrology,* 197(1–4), p. 230–257.

Adam, J.C. and D.P. Lettenmaier (2003). Adjustment of global gridded precipitation for systematic bias. *Journal of Geophysical Research,* 108(D9), 4257.

Adler, R.F., G. Gu and G.J. Huffman (2011). Estimating climatological bias errors for the Global Precipitation Climatology Project (GPCP). *Journal of Applied Meteorology and Climatology,* 51(1), p. 84–99.

Adler, R.F., G.J. Huffman, A. Chang, R. Ferraro, P.-P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin and E. Nelkin (2003). The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *Journal of Hydrometeorology,* 4(6), p. 1147–1167.

Alcamo, J., M. Flörke and M. Märker (2007). Future long-term changes in global water resources driven by socio-economic and climatic changes. *Hydrological Sciences Journal,* 52(2), p. 247–275.

Alfieri, L., P. Burek, E. Dutra, B. Krzeminski, D. Muraro, J. Thielen and F. Pappenberger (2013). GloFAS - global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences,* 17(3), p. 1161–1175.

Allen, R.G., M. Smith, L.S. Pereira and A. Perrier (1994). An update for the calculation of reference evapotranspiration. *ICID Bulletin,* 43(2), p. 35–92.

Arnell, N.W. (1999). A simple water balance model for the simulation of streamflow over a large geographic domain. *Journal of Hydrology,* 217(3–4), p. 314–335.

Arnell, N.W. (2003). Effects of IPCC SRES emissions scenarios on river runoff: a global perspective. *Hydrology and Earth System Sciences,* 7(5), p. 619–641.

Baldassarre, G.D. and A. Montanari (2009). Uncertainty in river discharge observations: A quantitative analysis. *Hydrology and Earth System Sciences,* 13(6), p. 913–921.

Balsamo, G., C. Albergel, A. Beljaars, S. Boussetta, H. Cloke, D. Dee, E. Dutra, J. Muñoz-Sabater, F. Pappenberger, P. de Rosnay, T. Stockdale and F. Vitart (2013).

ERA-Interim/Land: A global land water resources dataset. *Hydrology and Earth System Sciences Discussions,* 10(12), p. 14705–14745.

Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi and A.K. Betts (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system. *Journal of Hydrometeorology,* 10(3), p. 623–643.

Balsamo, G., S. Boussetta, P. Lopez and L. Ferrati (2010). *Evaluation of ERA-Interim and ERA-Interim-GPCP-rescaled precipitation over the U.S.A.* [Online]. ERA Report Series, No. 5. Reading, United Kingdom: ECMWF. [Accessed 20 November 2014]. Available from:
http://old.ecmwf.int/publications /

Balsamo, G., F. Pappenberger, E. Dutra, P. Viterbo and B. van den Hurk (2011). A revised land hydrology in the ECMWF model: A step towards daily water flux prediction in a fully-closed water cycle. *Hydrological Processes,* 25(7), p. 1046–1054.

Batjes, N.H. (2005). ISRIC-WISE global data set of derived soil properties on a 0.5 by 0.5 degree grid (version 3.0). [Online]. Report 2005/08. Wageningen, The Netherlands: ISRIC – World Soil Information. [Accessed 20 November 2014]. Available from: http://www.isric.org/

Baumgartner, A. and E. Reichel (1975). *Die Weltwasserbilanz*. München, Germany: R. Oldenbourg Verlag GmbH.

Bechtold, P., P. Bauer, P. Berrisford, J. Bidlot, C. Cardinali, T. Haiden, M. Janousek, D. Klocke, L. Magnusson, A. McNally, F. Prates, M. Rodwell, N. Semane and F. Vitart (2012). *Progress in predicting tropical systems: The role of convection.* ECMWF Technical Memoranda, No. 686. Reading, United Kingdom: ECMWF. [Accessed 21 November 2014]. Available from:
http://old.ecmwf.int/publications /

Beck, H.E., A.I.J.M. van Dijk, D.G. Miralles, R.A.M. de Jeu, L.A. Bruijnzeel, T.R. McVicar and J. Schellekens (2013). Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research,* 49(12), p. 7843–7863.

Becker, A., P. Finger, A. Meyer-Christoffer, B. Rudolf, K. Schamm, U. Schneider and M. Ziese (2013). A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth System Science Data,* 5(1), p. 71–99.

Beven, K., P.J. Smith and A. Wood (2011). On the colour and spin of epistemic error (and what we might do about it). *Hydrology and Earth System Sciences,* 15(10), p. 3123–3133.

Beven, K. and I. Westerberg (2011). On red herrings and real herrings: Disinformation and information in hydrological inference. *Hydrological Processes,* 25(10), p. 1676–1680.

Beven, K.J. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources,* 16(1), p. 41–51.

Beven, K.J. (2000). Uniqueness of place and process representations in hydrological modelling. *Hydrology and Earth System Sciences,* 4(2), p. 203–213.

Beven, K.J. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology,* 320(1–2), p. 18–36.

Beven, K.J. (2009). *Environmental modelling: An uncertain future?* London, United Kingdom: Routledge.

Biemans, H., I. Haddeland, P. Kabat, F. Ludwig, R.W.A. Hutjes, J. Heinke, W. von Bloh and D. Gerten (2011). Impact of reservoirs on river discharge and irrigation water supply during the 20th century. *Water Resources Research,* 47(3), W03509.

Biemans, H., R.W.A. Hutjes, P. Kabat, B.J. Strengers, D. Gerten and S. Rost (2009). Effects of precipitation uncertainty on discharge calculations for main river basins. *Journal of Hydrometeorology,* 10(4), p. 1011–1025.

Bierkens, M.F.P., V.A. Bell, P. Burek, N. Chaney, L. Condon, C.H. David, A. de Roo, P. Döll, N. Drost, J.S. Famiglietti, M. Flörke, D.J. Gochis, P. Houser, R. Hut, J. Keune, S. Kollet, R. Maxwell, J.T. Reager, L. Samaniego, E. Sudicky, E.H. Sutanudjaja, N. van de Giesen, H. Winsemius and E.F. Wood (2014). Hyper-resolution global hydrological modelling: What's next? *Hydrological Processes*, doi: 10.1002/hyp.10391.

Box, G.E.P. (1979). Robustness in scientific model building. In: Launer, R.L. and G.N. Wilkinson (eds.) *Robustness in statistics.* New York, USA: Academic Press, p. 201–236.

Box, G.E.P. and D.R. Cox (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 211–252.

Budyko, M.I. (1974). *Climate and life*. London, United Kingdom: Academic Press.

Butts, M.B., J.T. Payne, M. Kristensen and H. Madsen (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology,* 298(1–4), p. 242–266.

Castellarin, A., G. Galeati, L. Brandimarte, A. Montanari and A. Brath (2004). Regional flow-duration curves: Reliability for ungauged basins. *Advances in Water Resources,* 27(10), p. 953–965.

Dalin, C., M. Konar, N. Hanasaki, A. Rinaldo and I. Rodriguez-Iturbe (2012). Evolution of the global virtual water trade network. *Proceedings of the National Academy of Sciences,* 109(16), p. 5989–5994.

Deardorff, J.W. (1978). Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. *Journal of Geophysical Research,* 83(C4), p. 1889–1903.

Decharme, B. and H. Douville (2006). Uncertainties in the GSWP-2 precipitation forcing and their impacts on regional and global hydrological simulations. *Climate Dynamics,* 27(7–8), p. 695–713.

Dee, D.P., S.M. Uppala, A.J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M.A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A.C.M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A.J. Geer, L. Haimberger, S.B. Healy, H. Hersbach, E.V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A.P. McNally, B.M. Monge-Sanz, J.J. Morcrette, B.K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.N. Thépaut and F. Vitart (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society,* 137(656), p. 553–597.

DeFries, R.S. and M. Hansen. 2010. *ISLSCP II University of Maryland Global Land Cover Classifications, 1992-1993*. In Hall, Forrest G., G. Collatz, B. Meeson, S. Los, E. Brown de Colstoun, and D. Landis (eds.). ISLSCP Initiative II Collection. Data set. Available online [http://daac.ornl.gov/] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A.

Demaria, E.M., B. Nijssen and T. Wagener (2007). Monte Carlo sensitivity analysis of land surface parameters using the variable infiltration capacity model. *Journal of Geophysical Research,* 112(D11), D11113.

Dirmeyer, P.A., X. Gao, M. Zhao, Z. Guo, T. Oki and N. Hanasaki (2006). GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bulletin of the American Meteorological Society,* 87(10), p. 1381–1397.

Döll, P. and K. Fiedler (2008). Global-scale modeling of groundwater recharge. *Hydrology and Earth System Sciences,* 12(3), p. 863–885.

Döll, P., F. Kaspar and B. Lehner (2003). A global hydrological model for deriving water availability indicators: Model tuning and validation. *Journal of Hydrology,* 270(1–2), p. 105–134.

Döll, P. and B. Lehner (2002). Validation of a new global 30-min drainage direction map. *Journal of Hydrology,* 258(1–4), p. 214–231.

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

Eagleson, P.S. (1986). The emergence of global-scale hydrology. In *1986, Trends and directions in hydrology*. Collected reprint series (1986). Washington, DC, US: American Geophysical Union. 6S–14S.

ECMWF (2013). *User guide to ECMWF forecast products*. [Online]. Reading, United Kingdom: ECMWF. [Accessed 20 November 2014]. Available from: http://www.ecmwf.int/

Eli, S. (1998). Investigating catchment hydrology and low flow characteristics using GIS. *Nordic hydrology,* 29(2), p. 105–128.

Engeland, K., C.Y. Xu and L. Gottschalk (2005). Assessing uncertainties in a conceptual water balance model using bayesian methodology. *Hydrological Sciences Journal,* 50(1), p. 45–63.

Federer, C.A., C. Vörösmarty and B. Fekete (1996). Intercomparison of methods for calculating potential evaporation in regional and global water balance models. *Water Resources Research,* 32(7), p. 2315–2321.

Fekete, B.M. and C.J. Vörösmarty (2002). The current status of global river discharge monitoring and potential new technologies complementing traditional discharge measurements. In D. Schertzer et al. (eds.): *Predictions in Ungauged Basins: PUB kick-off, 20–22 November 2002, Brasilia*. IAHS Publication 349. Paris: International Association of Hydrological Sciences. p. 129–136.

Fekete, B.M., C.J. Vörösmarty and W. Grabs (1999). *Global composite runoff fields of observed river discharge and simulated water balances*. GRDC Report series, No. 22. Koblenz, Germany: Global Runoff Data Centre.

Fekete, B.M., C.J. Vörösmarty and W. Grabs (2002). High-resolution fields of global runoff combining observed river discharge and simulated water balances. *Global Biogeochemical Cycles,* 16(3), 1042.

Fekete, B.M., C.J. Vörösmarty, J.O. Roads and C.J. Willmott (2004). Uncertainties in precipitation and their impacts on runoff estimates. *Journal of Climate,* 17(2), p. 294–304.

Finley, J.W. and J.N. Seiber (2014). The nexus of food, energy, and water. *Journal of Agricultural and Food Chemistry,* 62(27), p. 6255–6262.

Foley, J.A., I.C. Prentice, N. Ramankutty, S. Levis, D. Pollard, S. Sitch and A. Haxeltine (1996). An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochemical Cycles,* 10(4), p. 603–628.

Frewer, L., S. Hunt, M. Brennan, S. Kuznesof, M. Ness and C. Ritson (2003). The views of scientific experts on how the public conceptualize uncertainty. *Journal of Risk Research,* 6(1), p. 75–85.

Gleick, P.H. and M. Palaniappan (2010). Peak water limits to freshwater withdrawal and use. *Proceedings of the National Academy of Sciences of the United States of America,* 107(25), p. 11155–11162.

Gosling, S. N. and N. W. Arnell (2011). Simulating current global river runoff with a global hydrological model: Model revisions, validation, and sensitivity analysis. *Hydrological Processes*, 25(7), p. 1129–1145.

Groisman, P.Y. and D.R. Legates (1994). The accuracy of United States precipitation data. *Bulletin of the American Meteorological Society,* 75(2), p. 215–227.

Güntner, A. (2008). Improvement of global hydrological models using GRACE data. *Surveys in Geophysics,* 29(4), p. 375–397.

Haddeland, I., D.B. Clark, W. Franssen, F. Ludwig, F. Voß, N.W. Arnell, N. Bertrand, M. Best, S. Folwell, D. Gerten, S. Gomes, S.N. Gosling, S. Hagemann, N. Hanasaki, R. Harding, J. Heinke, P. Kabat, S. Koirala, T. Oki, J. Polcher, T. Stacke, P. Viterbo, G.P. Weedon and P. Yeh (2011). Multimodel estimate of the global terrestrial water balance: Setup and first results. *Journal of Hydrometeorology,* 12(5), p. 869–884.

Hagan, M.T. and M.B. Menhaj (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks,* 5(6), p. 989–993.

Hanasaki, N., T. Inuzuka, S. Kanae and T. Oki (2010). An estimation of global virtual water flow and sources of water withdrawal for major crops and livestock products using a global hydrological model. *Journal of Hydrology,* 384(3–4), p. 232–244.

Hanasaki, N., S. Kanae and T. Oki (2006). A reservoir operation scheme for global river routing models. *Journal of Hydrology,* 327(1–2), p. 22–41.

Hanasaki, N., S. Kanae, T. Oki, K. Masuda, K. Motoya, N. Shirakawa, Y. Shen and K. Tanaka (2008a). An integrated model for the assessment of global water resources - Part 1: Model description and input meteorological forcing. *Hydrology and Earth System Sciences,* 12, p. 1007–1025.

Hanasaki, N., S. Kanae, T. Oki, K. Masuda, K. Motoya, N. Shirakawa, Y. Shen and K. Tanaka (2008b). An integrated model for the assessment of global water resources - Part 2: Applications and assessments. *Hydrology and Earth System Sciences,* 12, p. 1027–1037.

Harris, I., P.D. Jones, T.J. Osborn and D.H. Lister (2014). Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset. *International Journal of Climatology,* 34(3), p. 623–642.

Henderson-Sellers, A., Z.L. Yang and R.E. Dickinson (1993). The Project for Intercomparison of Land-Surface Parameterization Schemes. *Bulletin of the American Meteorological Society,* 74(7), p. 1335–1349.

Hersbach, H., C. Peubey, A. Simmons, P. Poli, D. Dee and P. Berrisford (2013). *ERA-20CM: A twentieth century atmospheric model ensemble*. [Online]. ERA Report Series, No. 16. Reading: ECMWF. [Accessed 20 November 2014]. Available from: http://old.ecmwf.int/publications

Hisdal, H. and O.E. Tveito (1993). Extension of runoff series using empirical orthogonal functions. *Hydrological Sciences Journal,* 38(1), p. 33–49.

Hrachowitz, M., H.H.G. Savenije, G. Blöschl, J.J. McDonnell, M. Sivapalan, J.W. Pomeroy, B. Arheimer, T. Blume, M.P. Clark, U. Ehret, F. Fenicia, J.E. Freer, A. Gelfan, H.V. Gupta, D.A. Hughes, R.W. Hut, A. Montanari, S. Pande, D. Tetzlaff, P.A. Troch, S. Uhlenbrook, T. Wagener, H.C. Winsemius, R.A. Woods, E. Zehe and C. Cudennec (2013). A decade of predictions in ungauged basins (PUB) - a review. *Hydrological Sciences Journal,* 58(6), p. 1198–1255.

Hunger, M. and P. Döll (2008). Value of river discharge data for global-scale hydrological modeling. *Hydrology and Earth System Sciences,* 12(3), p. 841–861.

Jiménez Cisneros, B.E., T. Oki, N.W. Arnell, G. Benito, J.G. Cogley, P. Döll, T. Jiang and S.S. Mwakalila (2014). Freshwater resources. In: Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea and L.L. White (eds.) *Climate change 2014: Impacts, adaptation, and vulnerability. Part a: Global and sectoral aspects. Contribution of working group ii to the fifth assessment report of the intergovernmental panel of climate change.* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, p. 229–269.

Kaspar, F. (2004). *Entwicklung und Unsicherheitanalyse eines globalen hydrologischen Modells*. Ph.D. thesis. Kassel University: Kassel University Press.

Kirchner, J.W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research,* 42(3), W03S04.

Korzun, V.I., A.A. Sokolow, M.I. Budyko, K.P. Voskresensky, G.P. Kalininin, A.A. Konoplyanstev, E.S. Korotkevich, P.S. Kuzin and M.I. Lvovich (1978). *World water balance and water resources of the earth*. USSR National Committee for the International Hydrological Decade. English translation. Studies and Reports in Hydrology, No. 25. Paris, France: UNESCO.

Kuczera, G. and M. Mroczkowski (1998). Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resources Research,* 34(6), p. 1481–1489.

Laaha, G. and G. Blöschl (2006). A comparison of low flow regionalisation methods - catchment grouping. *Journal of Hydrology,* 323(1–4), p. 193–214.

Legates, D.R. (1987). *A climatology of global precipitation*. Publications in Climatology, vol. XL, No. 1. Centerton, New Jersey: C. W. Thorntwaite Associates Laboratory of Climatology and Newark, Delaware: University of Delaware Center for Climatic Research.

Lehner, B. (2012). *Derivation of watershed boundaries for GRDC gauging stations based on the HydroSHEDS drainage network*. [Online]. GRDC Technical Report. No. 41. Koblenz, Germany: Global Runoff Data Centre. [Accessed 20 November 2014]. Available from: http://www.bafg.de/GRDC/

Lehner, B., C. Reidy Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, J. Magome, C. Nilsson, J.C. Robertson, R. Rödel, N. Sindorf and D. Wisser (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.

Lehner, B., K. Verdin and A. Jarvis (2008). New global hydrography derived from spaceborne elevation data. *Eos Transactions American Geophysical Union,* 89(10), p. 93–94.

Liang, X., D.P. Lettenmaier, E.F. Wood and S.J. Burges (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research,* 99(D7), p. 14415–14428.

Lu, J.B., G. Sun, S.G. McNulty and D.M. Amatya (2005). A comparison of six potential evapotranspiration methods for regional use in the southeastern United States. *Journal of the American Water Resources Association,* 41(3), p. 621–633.

L'vovich, M.I. (1979). *World water resources and their future* (1974 in russian, English translation in 1979 by R.L. Nace [Ed.]). Chelsea, MI, U.S.: Lithocrafters Inc.

Manabe, S. (1969). Climate and the ocean circulation: I. The atmospheric circulation and the hydrology of the earth's surface. *Monthly Weather Review,* 97(11), p. 739–774.

Mazvimavi, D., A.M.J. Meijerink, H.H.G. Savenije and A. Stein (2005). Prediction of flow characteristics using multiple regression and neural networks: A case study in Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C,* 30(11–16), p. 639–647.

McGuire, A.D., S. Sitch, J.S. Clein, R. Dargaville, G. Esser, J. Foley, M. Heimann, F. Joos, J. Kaplan, D.W. Kicklighter, R.A. Meier, J.M. Melillo, B. Moore, III, I.C. Prentice, N. Ramankutty, T. Reichenau, A. Schloss, H. Tian, L.J. Williams and U. Wittenberg (2001). Carbon balance of the terrestrial biosphere in the Twentieth Century: Analyses of CO_2, climate and land use effects with four process-based ecosystem models. *Global Biogeochemical Cycles,* 15(1), p. 183–206.

McIntyre, N., H. Lee, H. Wheater, A. Young and T. Wagener (2005). Ensemble predictions of runoff in ungauged catchments. *Water Resources Research,* 41(12), W12434.

McMillan, H., T. Krueger and J. Freer (2012). Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality. *Hydrological Processes*, 26(26), 4078–4111.

Michaelides, S., V. Levizzani, E. Anagnostou, P. Bauer, T. Kasparis and J.E. Lane (2009). Precipitation: Measurement, remote sensing, climatology and modeling. *Atmospheric Research,* 94(4), p. 512–533.

Mimikou, M. and S. Kaemaki (1985). Regionalization of flow duration characteristics. *Journal of Hydrology,* 82(1–2), p. 77–91.

Monteith, J.L. (1965). Evaporation and environment. *Symposia of the Society for Experimental Biology,* 19, p. 205–234.

Mueller, B., M. Hirschi, C. Jimenez, P. Ciais, P.A. Dirmeyer, A.J. Dolman, J.B. Fisher, M. Jung, F. Ludwig, F. Maignan, D.G. Miralles, M.F. McCabe, M. Reichstein, J. Sheffield, K. Wang, E.F. Wood, Y. Zhang and S.I. Seneviratne (2013). Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis. *Hydrology and Earth System Sciences,* 17(10), p. 3707–3720.

Müller Schmied, H., S. Eisner, D. Franz, M. Wattenbach, F.T. Portmann, M. Flörke and P. Döll (2014). Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration. *Hydrology and Earth System Sciences,* 18(9), p. 3511–3538.

Nijssen, B., R. Schnur and D.P. Lettenmaier (2001). Global retrospective estimation of soil moisture using the variable infiltration capacity land surface model, 1980–93. *Journal of Climate,* 14(8), p. 1790–1808.

Nilsson, C., C.A. Reidy, M. Dynesius and C. Revenga (2005). Fragmentation and flow regulation of the world's large river systems. *Science,* 308(5720), p. 405–408.

Oki, T. and S. Kanae (2006). Global hydrological cycles and world water resources. *Science,* 313(5790), p. 1068–1072.

Olden, J.D., M.K. Joy and R.G. Death (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling,* 178(3–4), p. 389–397.

Overgaard, J., D. Rosbjerg and M.B. Butts (2006). Land-surface modelling in hydrological perspective - a review. *Biogeosciences,* 3(2), p. 229–241.

Palmer, T.N., R. Buizza, M. Leutbecher, R. Hagedorn, T. Jung, M. Rodwell, F. Vitart, J. Berner, E. Hagel, A. Lawrence, F. Pappenberger, Y.-Y. Park, L. von Bremen and I. Gilmour (2007). *The ensemble prediction system: Recent and ongoing* *developments*. ECMWF Technical Memoranda, No. 540. Reading, United Kingdom: ECMWF [Accessed 20 November 2014]. Available from: http://old.ecmwf.int/publications/

Peel, M.C., F.H.S. Chiew, A.W. Western and T.A. McMahon. (2000). *Extension of unimpaired monthly streamflow data and regionalisation of parameter values to estimate streamflow in ungauged catchments.* [Online]. Report prepared for the National Land and Water Resources Audit Melbourne, Australia: The University of Melbourne.

Peel, M.C., B.L. Finlayson and T.A. McMahon (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences Discussions,* 4(2), p. 439–473.

Peel, M.C., T.A. McMahon and B.L. Finlayson (2010). Vegetation impact on mean annual evapotranspiration at a global catchment scale. *Water Resources Research,* 46(9), W09508.

Pitman, A.J. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology,* 23(5), p. 479–510.

Polcher, J., K. Laval, L. Dümenil, J. Lean and P.R. Rowntree (1996). Comparing three land surface schemes used in general circulation models. *Journal of Hydrology,* 180(1–4), p. 373–394.

Post, D.A., J.A. Jones and G.E. Grant (1998). An improved methodology for predicting the daily hydrologic response of ungauged catchments. *Environmental Modelling and Software,* 13(3–4), p. 395–403.

Priestley, C.H.B. and R.J. Taylor (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review,* 100(2), p. 81–92.

Rao, A.R. and C.H. Hsieh (1991). Estimation of variables at ungaged locations by empirical orthogonal functions. *Journal of Hydrology,* 123(1–2), p. 51–67.

Refsgaard, J.C. and B. Storm (1996). Construction, calibration and validation of hydrological models. In: Abbott, M.B. and J.C. Refsgaard (eds.) *Distributed hydrological modelling.* Water Science and Technology Library, vol. 22. Dordrecht, The Netherlands: Kluwer Academic Publishers, p. 41–54.

Reichl, J.P.C., A.W. Western, N.R. McIntyre and F.H.S. Chiew (2009). Optimization of a similarity measure for estimating ungauged streamflow. *Water Resources Research,* 45, W10423.

Rodwell, M.J., D.S. Richardson, T.D. Hewson and T. Haiden (2010). A new equitable score suitable for verifying precipitation in numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society,* 136(650), p. 1344–1363.

Sauquet, E. and C. Catalogne (2011). Comparison of catchment grouping methods for flow duration curve estimation at ungauged sites in France. *Hydrology and Earth System Sciences,* 15(8), p. 2421–2435.

Sellers, P., Y. Mintz, Y. Sud and A. Dalcher (1986). A simple biosphere model (SIB) for use within general circulation models. *Journal of the Atmospheric Sciences,* 43(6), p. 505–531.

Seneviratne, S.I. (2012). Climate science: Historical drought trends revisited. *Nature,* 491(7424), p. 338–339.

Sheffield, J., E.F. Wood and M.L. Roderick (2012). Little change in global drought over the past 60 years. *Nature,* 491(7424), p. 435–438.

Shiklomanov, A.I., R.B. Lammers and C.J. Vörösmarty (2002). Widespread decline in hydrological monitoring threatens Pan-Arctic research. *Eos Transactions American Geophysical Union,* 83(2), p. 13–17.

Shiklomanov, I.A. (ed.) (1997). *Assessment of water resources and water availability in the world.* In: Comprehensive assessment of the freshwater resources of the world. World Meteorological Organization and Stockholm Environmantal Institute.

Shuttleworth, W.J. (1988). Macrohydrology - the new challenge for process hydrology. *Journal of Hydrology,* 100(1–3), p. 31–56.

Siebert, S., J. Burke, J.M. Faures, K. Frenken, J. Hoogeveen, P. Döll and F.T. Portmann (2010). Groundwater use for irrigation - a global inventory. *Hydrology and Earth System Sciences,* 14, p. 1863–1880.

Simmons, A.J. and A. Hollingsworth (2002). Some aspects of the improvement in skill of numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society,* 128(580), p. 647–677.

Sivapalan, M., K. Takeuchi, S.W. Franks, V.K. Gupta, H. Karambiri, V. Lakshmi, X. Liang, J.J. McDonnell, E.M. Mendiondo, P.E. O'Connell, T. Oki, J.W. Pomeroy, D. Schertzer, S. Uhlenbrook and E. Zehe (2003). IAHS decade on predictions in ungauged basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal,* 48(6), p. 857–880.

Sood, A. and V. Smakhtin (2014). Global hydrological models: A review. *Hydrological Sciences Journal*, DOI:10.1080/02626667.2014.950580.

Strangeways, I. (2004). Improving precipitation measurement. *International Journal of Climatology,* 24(11), p. 1443–1460.

Strangeways, I. (2010). A history of rain gauges. *Weather,* 65(5), p. 133–138.

Tapiador, F.J., F.J. Turk, W. Petersen, A.Y. Hou, E. García-Ortega, L.A.T. Machado, C.F. Angelis, P. Salio, C. Kidd, G.J. Huffman and M. de Castro (2012). Global precipitation measurement: Methods, datasets and applications. *Atmospheric Research,* 104–105(0), p. 70–97.

The Ad Hoc Group, C. Vörösmarty, A. Askew, W. Grabs, R.G. Barry, C. Birkett, P. Döll, B. Goodison, A. Hall, R. Jenne, L. Kitaev, J. Landwehr, M. Keeler, G. Leavesley, J. Schaake, K. Strzepek, S.S. Sundarvel, K. Takeuchi and F. Webster (2001). Global water data: A newly endangered species. *Eos Transactions American Geophysical Union,* 82(5), p. 54–58.

Tourian, M.J., N. Sneeuw and A. Bárdossy (2013). A quantile function approach to discharge estimation from satellite altimetry (ENVISAT). *Water Resources Research,* 49(7), p. 4174–4186.

Trenberth, K. E., J. T. Fasullo, and J. Mackaro (2011). Atmospheric moisture transports from ocean to land and global energy flows in reanalyses, *Journal of Climate*, *24*(18), 4907–4924.

Uppala, S.M., P.W. Kållberg, A.J. Simmons, U. Andrae, V.D.C. Bechtold, M. Fiorino, J.K. Gibson, J. Haseler, A. Hernandez, G.A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R.P. Allan, E. Andersson, K. Arpe, M.A. Balmaseda, A.C.M. Beljaars, L.V.D. Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B.J. Hoskins, L. Isaksen, P.A.E.M. Janssen, R. Jenne, A.P. McNally, J.F. Mahfouf, J.J. Morcrette, N.A. Rayner, R.W. Saunders, P. Simon, A. Sterl, K.E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo and J. Woollen (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society,* 131(612), p. 2961–3012.

Vandewiele, G.L. and A. Elias (1995). Monthly water balance of ungauged catchments obtained by geographical regionalization. *Journal of Hydrology,* 170(1–4), p. 277–291.

Verdin, K. L. 2011. *ISLSCP II HYDRO1k Elevation-derived Products*. In Hall, Forrest G., G. Collatz, B. Meeson, S. Los, E. Brown de Colstoun, and D. Landis (eds.). ISLSCP Initiative II Collection. Data set. Available online [http://daac.ornl.gov/] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. doi:10.3334/ORNLDAAC/1007

Voß, F., J. Alcamo, N. Arnell, I. Haddeland, S. Hagemann, R. Lammers, T. Oki, N. Hanasaki and H. Kim (2008). *First results from intercomparison of surface water availability modules*. [Online]. WATCH Technical Report, No. 1. [Accessed 20 November 2014]. Available from: http://www.eu-watch.org/

Vörösmarty, C.J., C.A. Federer and A.L. Schloss (1998). Potential evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling. *Journal of Hydrology,* 207(3–4), p. 147–169.

Vörösmarty, C.J., B.M. Fekete, M. Meybeck and R.B. Lammers (2000b). Global system of rivers: Its role in organizing continental land mass and defining land-to-ocean linkages. *Global Biogeochemical Cycles,* 14(2), p. 599–621.

Vörösmarty, C.J., P. Green, J. Salisbury and R.B. Lammers (2000a). Global water resources: Vulnerability from climate change and population growth. *Science,* 289(5477), p. 284–288.

Vörösmarty, C.J., B. Moore III, A.L. Grace, M.P. Gildea, J.M. Melillo, B.J. Peterson, E.B. Rastetter and P.A. Steudler (1989). Continental scale models of water balance and fluvial transport: An application to South America. *Global Biogeochemical Cycles,* 3(3), p. 241–265.

Wagener, T. and A. Montanari (2011). Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. *Water Resources Research,* 47(6), W06301.

Wagener, T., M. Sivapalan, P. Troch and R. Woods (2007). Catchment classification and hydrologic similarity. *Geography Compass,* 1(4), p. 901–931.

Wedi, N.P. (2014). Increasing horizontal resolution in numerical weather prediction and climate simulations: Illusion or panacea? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 372(2018), doi: 10.1098/rsta.2013.0289.

Weedon, G.P., S. Gomes, P. Viterbo, W.J. Shuttleworth, E. Blyth, H. Österle, J.C. Adam, N. Bellouin, O. Boucher and M. Best (2011). Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *Journal of Hydrometeorology,* 12(5), p. 823–848.

Westerberg, I., J. Guerrero, P. Younger, K. Beven, J. Seibert, S. Halldin, J. Freer and C. Xu (2011). Calibration of hydrological models using flow-duration curves. *Hydrology and Earth System Sciences,* 15, p. 2205–2227.

Westerberg, I.K., L. Gong, K.J. Beven, J. Seibert, A. Semedo, C.Y. Xu and S. Halldin (2014). Regional water balance modelling using flow-duration curves with observational uncertainties. *Hydrology and Earth System Sciences,* 18(8), p. 2993–3013.

Widén-Nilsson, E. (2007). *Global-scale modelling of the land-surface water balance: Development and analysis of WASMOD-M.* Ph.D. thesis. Uppsala University, Uppsala: Acta Universitatis Upsaliensis.

Widén-Nilsson, E., L. Gong, S. Halldin and C.-Y. Xu (2009). Model performance and parameter behavior for varying time aggregations and evaluation criteria in the WAMOD-M global water balance model. *Water Resources Research,* 45, W05418.

Widén-Nilsson, E., S. Halldin and C.-Y. Xu (2007). Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation. *Journal of Hydrology,* 340(1–2), p. 105–118.

Williams, M., A.D. Richardson, M. Reichstein, P.C. Stoy, P. Peylin, H. Verbeeck, N. Carvalhais, M. Jung, D.Y. Hollinger, J. Kattge, R. Leuning, Y. Luo, E. Tomelleri, C.M. Trudinger and Y.P. Wang (2009). Improving land surface models with FLUXNET data. *Biogeosciences,* 6(7), p. 1341–1359.

Wisser, D., B.M. Fekete, C.J. Vörösmarty and A.H. Schumann (2010). Reconstructing 20th century global hydrography: A contribution to the Global Terrestrial Network - Hydrology (GTN-H). *Hydrology and Earth System Sciences,* 14(1), p. 1–24.

Wood, E.F., D.P. Lettenmaier, X. Liang, D. Lohmann, A. Boone, S. Chang, F. Chen, Y. Dai, R.E. Dickinson, Q. Duan, M. Ek, Y.M. Gusev, F. Habets, P. Irannejad, R. Koster, K.E. Mitchel, O.N. Nasonova, J. Noilhan, J. Schaake, A. Schlosser, Y. Shao, A.B. Shmakin, D. Verseghy, K. Warrach, P. Wetzel, Y. Xue, Z.-L. Yang and Q.-C. Zeng (1998). The Project for Intercomparison of Land-Surface Parameterization Schemes (PILPS) phase 2(c) Red–Arkansas River basin experiment: 1. Experiment description and summary intercomparisons. *Global and Planetary Change,* 19(1–4), p. 115–135.

Wood, E.F., D.P. Lettenmaier and V.G. Zartarian (1992). A land-surface hydrology parameterization with subgrid variability for general circulation models. *Journal of Geophysical Research,* 97(D3), p. 2717–2728.

Wu, H., J.S. Kimball, H. Li, M. Huang, L.R. Leung and R.F. Adler (2012). A new global river network database for macroscale hydrologic modeling. *Water Resources Research,* 48(9), W09701.

Wu, H., J.S. Kimball, N. Mantua and J. Stanford (2011). Automated upscaling of river networks for macroscale hydrological modeling. *Water Resources Research,* 47(3), W03517.

WWAP (2014). *The United Nations world water development report 2014: Water and energy*. United Nations World Water Assessment Programme (WWAP). Paris: UNESCO.

Xu, C.-Y. (2003). Testing the transferability of regression equations derived from small sub-catchments to a large area in central Sweden. *Hydrology and Earth System Sciences,* 7(3), p. 317–324.

Yadav, M., T. Wagener and H. Gupta (2007). Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources,* 30(8), p. 1756–1774.

Zhao, L., J. Xia, C.-y. Xu, Z. Wang, L. Sobkowiak and C. Long (2013). Evapotranspiration estimation methods in hydrological models. *Journal of Geographical Sciences,* 23(2), p. 359–369.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1211

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014

Distribution: publications.uu.se
urn:nbn:se:uu:diva-236864