Global-Scale Modelling of the Land-Surface Water Balance

Development and Analysis of WASMOD-M

ELIN WIDÉN-NILSSON
Dissertation presented at Uppsala University to be publicly examined in Axel Hambergsalen, Geocentrum, Villavägen 16, Uppsala, Friday, December 14, 2007 at 10:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Water is essential for all life on earth. Global population increase and climate change are projected to increase the water stress, which already today is very high in many areas of the world. The differences between the largest and smallest global runoff estimates exceed the highest continental runoff estimates. These differences, which are caused by different modelling and measurement techniques together with large natural variabilities need to be further addressed. This thesis focuses on global water balance models that calculate global runoff, evaporation and water storage from precipitation and other climate data.

A new global water balance model, WASMOD-M was developed. Already when tuned against the volume error it reasonable produced within-year runoff patterns, but the volume error was not enough to confine the model parameter space. The parameter space and the simulated hydrograph could be better confined with, e.g., the Nash criterion. Calibration against snow-cover data confined the snow parameters better, although some equifinality still persisted. Thus, even the simple WASMOD-M showed signs of being overparameterised.

A simple regionalisation procedure that only utilised proximity contributed to calculate a global runoff estimate in line with earlier estimations. The need for better specifications of global runoff estimates was highlighted.

Global modellers depend on global data-sets that can have low quality in many areas. Major sources of uncertainty are precipitation and river regulation. A new routing method that utilises high-resolution flow network information in low-resolution calculations was developed and shown to perform well over all spatial scales, while the standard linear reservoir routing decreased in performance with decreasing resolution. This algorithm, called aggregated time-delay-histogram routing, is intended for inclusion in WASMOD-M.

Keywords: Global, Water balance, Runoff, Regionalisation, Model uncertainty, Multi-objective, Parameter, Evaluation criteria, Routing, Climate change

Elin Widén-Nilsson, Department of Earth Sciences, Villav. 16, Uppsala University, SE-75236 Uppsala, Sweden

© Elin Widén-Nilsson 2007

ISSN 1651-6214
ISBN 978-91-554-7051-7
urn:nbn:se:uu:diva-8352 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-8352)
“We must devote more attention not only to the technical issues of hydrology raised by the model builders but also to encouraging and preparing more young hydrologists to build a career in this direction. He who controls the future of global-scale models controls the direction of hydrology”

Peter S. Eagleson, 1986, The Emergence of Global-Scale Hydrology, Water resources research, 22(9): 6S-14S

”’It’s a little anxious,’ he said to himself, ‘to be a Very Small Animal Entirely Surrounded by Water’”

Piglet in Winnie-the-Pooh, A.A. Milne, 1926
List of papers


The Editorial Office of Advances in Atmospheric Sciences gave permission to reprint paper I and Elsevier (paper II) include publication in thesis in the authors retained rights.

In paper I, I drafted the section about global water balance modelling and contributed with references for the other sections. In paper II I made the data preparation, model codes, simulations, analysis and I was responsible for the writing. In paper III the second author made the first version of the simulation code and some of the data preparation, while I developed the code, made the simulations, and was responsible for the analysis. The writing was shared with the third author. The practical work of paper IV was performed by the first author, in discussions with me and the other co-authors.
Contents

Introduction...................................................................................................11

Global and continental runoff ..............................................................................13
    Global and continental runoff estimates.......................................................13
    Models that calculate global runoffs ........................................................16
        Land-surface models and dynamic global vegetation models .............16
        Two LSMs with hydrological focus .....................................................17
        Routing models ................................................................................17
        Global water balance models............................................................19

Uncertainties of GCMs and LSMs and the effect in studies of hydrological impact of climate change ....................................................23

Methods ...........................................................................................................25
    Global data for WASMOD-M..................................................................25
        Simulation time periods, model windup times and initial values ...27
    WASMOD-M ..........................................................................................28
        Model parameter tuning of WASMOD-M in gauged areas ..........30
            Criteria ......................................................................................31
        Inter-station areas ...........................................................................33
    Regionalisation of WASMOD-M’s parameter values to ungauged areas34
    Evaluation and validation of WASMOD-M ............................................34

Routing of daily runoff produced by WASMOD ..............................................35
    Site and data description ......................................................................35
    Resolutions ..........................................................................................35
    Daily, distributed WASMOD ................................................................36
    Linear reservoir routing .......................................................................37
    Time-delay-histogram routing ...............................................................37
    Velocity calibration ............................................................................38
    Benchmark ..........................................................................................39

Results...........................................................................................................40
    Tuning and criteria ................................................................................40
    Parameter-value behaviour ..................................................................41
    Time aggregation ................................................................................42
    Regionalisation ....................................................................................43
    Routing ...............................................................................................43
Introduction

Water is essential for life on earth. Humans need water for drinking, cooking and hygiene and indirectly for irrigation of crops, watering cattle, for industry and power plant usage. We also benefit from well-functioning ecological systems and excessive water withdrawal from rivers can cause collapses of some ecosystems. Today 1.1 billion people lack access to sufficient drinking water (WHO/UNICEF, 2000). This number and the extent of water scarce areas will increase with population growth and climate change (Vörösmarty et al., 2000a; Alcamo and Henrichs, 2002; Arnell, 2003). In addition to the predicted indirect human influence on the global hydrological cycle by emissions of greenhouse gasses, humans are known to directly influence by irrigation, dams etc. (Vörösmarty et al., 2004; Gordon et al., 2005; Oki and Kanae, 2006, Haddeland et al., 2007). These global problems call for global models (Eagleson, 1986; Shuttleworth, 1988). The future is uncertain but projection of future water resources can be made and models are a useful tool for this. But can these models be trusted? What are their abilities to simulate current and future water resources? Do the models describe the reality in a realistic way, or are they too simplistic or overparameterised? Current models and data estimates differ around 30% in their estimates of the global runoff and up to 70% for individual continents (GRDC, 2004; paper II). Different time periods, varying methods of extrapolating data to ungauged areas and varying definitions on which runoff parts to include are some causes of these differences.

This thesis focuses on global water balance models. These models calculate runoff in all rivers around the world from climatic input such as precipitation and temperature. Calculations of evaporation and runoff are performed independently for individual grid cells, often sized 0.5° latitude × 0.5° longitude, and routed along flow networks to the river outlets. Thus, the models focus on renewable water resources. Many global models only calculate the amount, not the quality, of the runoff.

Environmental models are a representation of the reality, but they can never include all processes and they cannot be verified to be true since natural systems never are closed (Oreskes et al., 1994). Oreskes et al. (1994) also argue against the usage of “validation”, but being such a commonly used word it is still used in this thesis, together with the more vague “evaluation”. Model simplifications is one source of model uncertainty, while others are measurement errors in input and validation data, miss-match between meas-
ured and model entities (the incommensurability problem), parameter errors and model structural errors. One effect of the uncertainties is equifinality, introduced to hydrological modelling community by Beven (1993), meaning that several parameter value combinations can produce equally good model results. Equifinality is an indication of an overparameterised model. It is also connected to the fact that a single representation of the system can not be found given the limitations of data (Beven, 2006). New processes are included in models to improve the results, but additional processes introduce additional parameters and these new parameters can often not be supported by the available data (Beven, 2006; 2007). The modelling done in this thesis was usually made with a wide range of parameter values to select sets that performed reasonably well, to take the equifinality into account.

The overall goal of the thesis was to calculate the global runoff variations in time and space. When the models get better in describing the current situations, they will also be more reliable tools to assess the effects on the global runoff of climate change, population growth, water withdrawal etc. The uncertainties in the global water balance models have thus been a focus of this work. The main tool has been the new, simple, global water balance model WASMOD-M. Compared to the catchment modelling community, there are few global water balance models, and the hypothesis was that a simple model would be useful to efficiently study model uncertainty. The construction of a new model could also be motivated by the advantages of model ensembles for future predictions. WASMOD-M will be able to contribute to analysis of present and future water resources. The overall objective was approached in a series of steps to:

- Review existing models (paper I; II) and their usage in climate studies (paper I) as well as to review model and data-based estimates of the global water runoff (paper II) and compile published model errors (paper III).
- Develop a simple global water balance model, WASMOD-M (paper II; III).
- Develop a simple regionalisation methodology to extrapolate parameter values to ungauged areas and given the regionalisation, report global water balance estimates, with clear information on which parts that are included and not (paper II).
- Evaluate the performance and parameter values of WASMOD-M. Can one, or a very small number of parameter-value sets, be used to produce acceptable global and continental water balances (paper II)? Is the model, despite its simplicity, overparameterised (paper III)?
- Study the influence of different model evaluation measures, of both runoff and snow, and time aggregations on model tuning (paper III).
- Develop a river routing algorithm that is independent of the spatial resolution (paper IV).
Global and continental runoff estimates

Estimates of the total global and continental runoffs can be based on gauge runoff data, precipitation minus evaporation (reanalysis) data, modelled runoff or remote sensing water estimates. Data and model-based runoff estimates are compiled in paper II and the complications in comparing these measurements are also discussed. One problem is that different continent boundaries are used. Oceania, e.g., is defined by different sources as including or excluding parts of Australia, New Zealand, Papua New Guinea and various small islands (e.g., Nijssen et al., 2001a). In addition to the border to Oceania, the borders of Asia with Europe often differ, e.g., by the assignment of Turkey (Nijssen et al., 2001a; Döll et al., 2003). The inclusion of whole Russia in Europe in the runoff estimate by FAO (2003) doubles the continents runoff although any influence on the higher Asian runoff in comparison with other estimates cannot be seen. Arctic areas (especially Antarctica and Greenland) are commonly excluded in modelling studies but included in data-based runoff estimates.

Runoff from rivers to oceans (exorheic) makes up the major part of the global runoff. Other contributions come from groundwater flow to oceans, river flow to inland basins (endorheic) and groundwater flow to inland basins. Korzun et al. (1978) estimate the amount of river flow which evaporates and comes from percolation into rivers. It is not always clear which of these components are included in various global and continental runoff estimates (Table 1). Global runoff also varies with time (Probst and Tarty, 1987; Shiklomanov, 1997; Labat, 2006) and not all estimates include information about the time period for the calculation (Table 1). It was shown in paper II that the low runoff estimates of the LSMS (land-surface models) in the compilation by Oki et al. (2001) to a large degree is caused by the climate during their short simulation period. Regional and global trends in runoff and their possible relations to different forcings related to climatic warming are analysed and debated (Labat et al., 2004; 2005; Legates et al., 2005; Milly et al., 2005; Shiklomanov et al., 2006; Gedney et al., 2006a; 2006b; Peel and McMahon, 2006; Piao et al., 2007).
The difference between the largest and smallest global runoff estimates in Table 1 exceeds the highest continental runoff estimate. Even if there is a general tendency that global runoff estimates based only on measurements are higher than modelled ones (Table 1), there are at least three reasons why the difference should be even larger (paper II). Runoff to internal basins (around 1000 km$^3$yr$^{-1}$) is commonly excluded in compilations of runoff measurements. There should also be a difference between those measurement compilations that include groundwater runoff into oceans and those that do not. Some of the measurement compilations include evaporation from rivers and dams, which is not included in any model with the exception of WGHM (Döll et al., 2003).

Apart from runoff from different continents, in km$^3$yr$^{-1}$ or mm, runoff can instead be presented as their ocean outlet (e.g., Fekete et al., 2002) or by latitude bands. In the latter case it is important to differentiate sums over the interior runoff generating cells and flow accumulated outlet cells. The first is smoothed and the latter peaked, especially at the outlet of the Amazon river. Baumgartner and Reichel (1975) present sums over interior runoff generating cells, but Dai and Trenberth (2002) compare these with the summation over outlet cells.

Recent additions to the global water estimates come by remote sensing technologies. Levels of larger water bodies measured by satellite altimetry can be related to river flows through observed rating curves (Calmant and Seyler, 2006). The errors are in the order of decimetres or centimetres and the main disadvantage of the method is the low temporal resolution for many sites (Calmant and Seyler, 2006). The area of surface water extent is also measured and combined with altimetry measurements (Alsdorf and Lettenmaier, 2003). River discharge can be calculated with reasonable accuracy without ground measurements from a combination of several space-based measures (Bjerklie et al., 2003). A more recent approach is measuring the temporal and spatial changes in Earth’s gravity by mass fluxes close to the earth surface and the GRACE (Gravity Recovery and Climate Experiment) satellite was launched 2002 (Güntner et al., 2007a). The seasonal changes in total water storage can be tracked for continents and large river basins, but the division into separate compartments, such as snow, ice, lakes, soil moisture and groundwater is complicated (Güntner et al., 2007a). Simulations with simpler and more advanced LSMs/hydrological models and WGHM have been used to compare the GRACE data and to separate the different storage compartments (e.g., Andersen and Hinderer, 2005; Rodell et al., 2004; Ramillien et al., 2006; Schmidt et al., 2006; Swenson and Milly, 2006). The GRACE publications focus on changes in water storage and no global runoff estimate has yet been published to the author’s knowledge, but evaporation estimates, calculated with the usage of measured precipitation and simulated runoff, are reported for some basins (Rodell, et al., 2004; Ramillien et al., 2006).
Table 1. Global and continental runoff estimates in km³/year (many data may not be directly comparable because of different continental boundaries and averaging periods). K78: Korzun et al. (1978), Table 157, time period not specified; L79: L’vovich (1979), Table 20, time period not specified; BR75: Baumgartner and Reichel (1975), Table XXXV, time period not specified; S97: Shiklomanov (1997), Table 4.2, time period 1921-1985; GRDC04: GRDC (2004), time period “approximately” 1961-1990; F02: Fekete et al. (2002), Table 3, model WBM and data, almost equal to Fekete et al. (1999b), time period not specified; F99bm: Fekete et al. (1999b), model WBM, time period not specified; O01: Oki et al. (2001), Table 2, land-surface models and TRIP routing model, time period 1987-1988; D03: Döll et al. (2003), Table 1, model WGHM, time period 1961-1990; G04: Gerten et al. (2004), Table 2, model LPJ, time period 1961-1990, WN07: paper II, WASMOD-M, time period 1961-1990

<table>
<thead>
<tr>
<th>Continent</th>
<th>Area (10⁶ km²)</th>
<th>Data-based estimates</th>
<th>Combined estimates</th>
<th>Model estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>K78</td>
<td>L79</td>
<td>BR75</td>
</tr>
<tr>
<td>Europe</td>
<td>10.1</td>
<td>2,970</td>
<td>3,110</td>
<td>2,564</td>
</tr>
<tr>
<td>Asia</td>
<td>44.0</td>
<td>14,100</td>
<td>13,190</td>
<td>12,467</td>
</tr>
<tr>
<td>Africa</td>
<td>30.1</td>
<td>4,600</td>
<td>4,225</td>
<td>3,409</td>
</tr>
<tr>
<td>North America</td>
<td>22.4</td>
<td>8,180</td>
<td>5,960</td>
<td>5,840</td>
</tr>
<tr>
<td>South America</td>
<td>17.9</td>
<td>12,200</td>
<td>10,380</td>
<td>11,039</td>
</tr>
<tr>
<td>Oceania</td>
<td>8.7</td>
<td>2,510</td>
<td>1,965</td>
<td>2,394</td>
</tr>
<tr>
<td>Global (except Antarctica)</td>
<td>133</td>
<td>44,560</td>
<td>38,830</td>
<td>37,713</td>
</tr>
</tbody>
</table>

N: endorheic basins included
X: exorheic basins only

a Grid-cell area and continental boundaries from Fekete et al. (1999b), endorheic basins included. These areas are only valid for the W estimate (F02 and F99bm is calculated for 1.6% fewer grid cells). Oceania is defined as Australia, New Zealand, Papua New Guinea and some small Islands. W simulation is only made for a minor part of Greenland, which is included in the North America values.

b It is assumed, that this estimate excludes endorheic basins (not clearly stated in the original publication)

It is assumed that these estimates include endorheic basins (not clearly stated in the original publication)
Models that calculate global runoffs

Global water balance models have runoff as its main simulated variable, but several other types of models simulate global runoff as an important by-product (paper I; II). These are the general circulation models (GCMs) used to simulate climate change with their land-surface models (LSMs), dynamic global vegetation models (DGVMs) and routing models. DGVMs and routing models can also be coupled to GCMs, both off-line and online, while the DGVMs and LSMs can be applied stand-alone as well. Routing models need runoff as input from either of these models.

Land-surface models and dynamic global vegetation models

The hydrological description of GCMs with the LSMs started with the simple bucket model by Manabe (1969) where runoff only occur when the soil water content is at field capacity. The soil is thus described as a bucket and runoff occurs when the bucket is filled. The bucket approach together with the evaporation formulation caused too dry summers and the LSM algorithms developed from 1980’s and onward (Ducharne et al., 1998; Pitman, 2003). The next generation of LSMs were often called soil-vegetation-transfer schemes (SVATs) (Koster and Suarez, 1994). SVATs are one-dimensional, describing the exchanges of water and heat in a vertical column between soil, vegetation and atmosphere, where the soil is divided into several layers. The terrestrial biosphere/biophysical models SiB and BATS (Dickinson, 1987; Sellers et al., 1986) are early SVATs. LSM is a wide concept and also the hydrological model Sacramento Soil Accounting model and the global water balance model WGHM have been classified as LSMs (Sheffield et al., 2003; Ramillien et al., 2006), but usually only energy-balance-based LSMs are considered (Overgaard et al., 2006). An overview of many current LSMs is given by Dirmeyer et al. (2006). The LSMs are continuously developed to include more and more processes like photosynthesis and other plant processes together with the explicit simulation of carbon such that some now have turned into DGVMs (Pitman, 2003). The DGVMs are very detailed models that simulate the carbon cycle and evaporation changes without changes in climate forcing (Gerten et al., 2004). These models simulate, e.g., photosynthesis, plant and soil respiration, plant growth and competition between plants, fires and carbon and nutrient cycling (Kucharik et al., 2000; Krinner et al., 2005). DGVMs are not only developed from the biophysical models, but also from biogeochemical, biogeographical and vegetation dynamic models (Cramer et al., 2001; Sitch et al., 2003). Another notation of the DGVMs is ecosystem models, to which also the biogeochemical models are counted (Prentice et al., 2000). A com-
plex Earth System model is formed when an atmospheric GCM is coupled with an ocean general circulation model, a DGVM, a marine biogeochemistry model and an ice model (Mikolajewicz et al., 2007). A DGVM of special interest is LPJ (Sitch et al., 2003) since its hydrological simulations have been compared to three water balance models by Gerten (2004).

Two LSMs with hydrological focus are the VIC model (Liang et al., 1994; Nijssen et al., 1997; Wood et al., 1992) and global model by Hanasaki (2007a; 2007b). The VIC model requires special attention as it has been applied globally (Nijssen et al., 2001a, 2001b), been used by many and been compared to several global runoff models. It can operate in full energy and water balance mode as well as a simpler water balance mode only (O’Donnell et al., 2000). It can be classified both as macro-scale hydrological model, SVAT and LSM. The model, which exist in several versions, can be applied both locally and globally and either lumped or distributed with several resolutions (Demaria et al., 2007). The within-cell and between-cell routing is described by Lohmann et al. (1996) and Nijssen et al. (1997). The VIC model can be run both with and without calibration (Nijssen et al., 2001a; 2001b). The first model version had 4 parameters (Wood et al., 1992) while Demaria et al. (2007) studied a selection of 10 streamflow parameters in VIC 4.0.4 and found overparameterisation in the baseflow formulation. Other model tests include the evaluation of a global regionalisation procedure (Nijssen et al., 2001a), model intercomparison in the PILPS project (e.g., Wood et al., 1998), and comparison with independent datasets such as snow (Nijssen et al., 2001b; Hillard et al., 2003; Pan et al., 2003; Sheffield et al., 2003). The recent model by Hanasaki (2007a; 2007b) also includes the energy balance but focuses on the water balance with separate modules for water withdrawal and dam operation.

Routing models

Flows in rivers are delayed by lakes and wetlands, as well as by dams, and long and meandering river stretches causes delays themselves. The procedure of modelling the runoff from its point of generation to an outlet is called routing, and is normally connected with calculations of these delays. A flow network is needed for the routing. Several global runoff networks exist (Table 2), from the 1 km resolution Hydro1k (USGS, 1999) to the $4^\circ \times 5^\circ$ network by Miller et al. (1994). The low-resolution networks were used to route the runoff of the older, very low-resolution GCMs. Measured runoff is an integrated measure of the conditions in the whole basin and is useful for GCM validation. A routing scheme applied to runoff produced by GCMs grid cells can be compared to gauge data (Russell and Miller, 1990; Liston et al., 1994; Hagemann and Dümenil, 1998; Coe, 2000; Arora, 2001).
Table 2. Examples of global runoff networks. “www availability” indicates the data-
sets that are directly downloadable over internet whereas other datasets are avail-
able after contact with the responsible researcher.

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Name</th>
<th>Reference</th>
<th>www availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 km</td>
<td>Hydro1k</td>
<td>USGS (1999)</td>
<td>Yes</td>
</tr>
<tr>
<td>5'</td>
<td>Coe (1998)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>5', 0.5°, 1°</td>
<td>Graham et al. (1999)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5°</td>
<td>Hagemann and Dümenil (1998)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5°</td>
<td>Renssen and Knoop (2000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5°</td>
<td>STN-30p</td>
<td>Vörösmarty et al. (2000b)</td>
<td>Yes</td>
</tr>
<tr>
<td>0.5°</td>
<td>DDM30</td>
<td>Döll and Lehner (2002)</td>
<td></td>
</tr>
<tr>
<td>1°</td>
<td>TRIP</td>
<td>Oki and Sud (1998)</td>
<td>Yes</td>
</tr>
<tr>
<td>2° × 2.5°, 4° × 5°</td>
<td>Miller et al. (1994)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are two main groups of routing models, the cell-to-cell models and the
source-to-sink models (Olivera et al., 2000). The cell-to-cell models, where
flow is accumulated along the flow net are the most common. Examples of
such models (paper I) are the models of Miller et al. (1994), Liston et al.
(1994), the HD model (Hagemann and Dümenil, 1998), TRIP (Oki et al.,
1999), the model by (Arora and Boer, 1999) and RTM (Branstetter and
Erickson, 2003). One or several linear reservoirs (Eq. 1) are commonly used
(Liston et al., 1994; Miller et al., 1994; Hagemann and Dümenil, 1998):

\[ S = k Q_{out} = \frac{d}{v} Q_{out} \]  

(1)

The proportionality constant \( k \), that regulates the outflow \( Q_{out} \), from the stor-
age \( S \) has the dimension of time and is related to travel distance, slope and
roughness of the streambed and the length, width and depth of the stream
(Liston et al., 1994). The distance \( d \) between cells is not constant and \( k = d/v \),
where \( v \) is the speed, is used since outflow can occur to one of eight possible
directions in the flow network (Miller et al., 1994; Kaspar, 2004). The dis-
tance is often multiplied by a meandering factor to get a representation of the
true river lengths in large grid cells (Oki et al., 1999; Lucas-Picher et al.,
2003). The speed \( v \), which is in the order of 1 ms\(^{-1}\) can be set constant or
vary between basins (Miller et al., 1994, Kaspar, 2004). Variable velocities
were explored by Arora and Boer (1999), Arora et al. (1999) and Lucas-
Picher et al. (2003) using the Mannings equation that relates the velocity to
the hydraulic radius of the river and the slope. Geomorphological relation-
ships between width and discharge and slope and discharge, were used to
estimate the model parameters. Sushama et al. (2004) found that a simpler
scheme that varies \( k \) instead of \( v \) could be used instead. A short time-step is
needed in cell-to-cell routing to get numerical stable problems (Liston et al.,
Global source-to-sink routing models are represented by Naden (1999) and Olivera (2000). In cell-to-cell models, accumulated runoff can be picked out in any cell at the cost of long computation times, while in source-to-sink models the destination cells, such as the basin outlet or a cell with a runoff gauge, are decided beforehand and the time delays from the source to the sink is pre-processed. Within-cell routing can also be simulated in large grid cells, and this procedure is related to the source-to-sink routing (Olivera et al., 2000).

Routing is only necessary for very few rivers in monthly global-scale models if only river transport is considered (Sausen et al., 1994; Kleinen and Petschel-Held, 2007). Lakes, wetlands and dams are much more important for the delays (Vörösmarty et al., 1997; Coe, 2000). Ice damming can influence the river flows as well (Graham, 2004; Shiklomanov et al., 2006). Recently Haddeland et al. (2006) and Hanasaki (2006) published algorithms for calculating operation strategies of dams. Irrigation dams are the focus in both studies.

Global water balance models

The global water-balance models are simple models, transferring precipitation to evaporation and runoff. WASMOD-M, as presented here, is a new model in this category (paper II). One of the first global water balance model studies was published by Vörösmarty et al. (1997). Döll et al. (2003) describe this kind of models as “global hydrological models”. That notation can also be used for LSMS (Andersen and Hinderer, 2005).

The simplest global water balance model is probably the one by Kleinen and Petschel-Held (2007) which focuses on flooding. It uses a monthly time-step, only snow storage and no soil moisture storage. Other global water balance models are the model by Klepper, shortly described in (RIVM/UNEP, 1997), the model by Miller et al. (2003), WBM (Vörösmarty et al., 1989, 1996, 1998), Macro-PDM (Arnell, 1999c, 2003) and WGHM (Döll et al., 2003; Kaspar, 2004), a submodel of WaterGAP (Döll et al., 1999; Alcamo et al., 2003a). The latter three are the most well-known global models. Some of their main similarities and differences are listed in Table 3. WASMOD-M is simpler than these three models. WGMH is the most complex of the three. All four models work on a 0.5°×0.5° resolution but Macro-PDM is applied with 10’ resolution in Europe (Schröter et al., 2005) and the coming WaterGAP3 will have 5’ minute resolution (Martina Weiß, pers. comm.). Three previous models use pseudo-daily climate data or calculations. WBM uses a flow network but does only have routing with the additional WTM water transport model. Macro-PDM has within-cell routing only and the runoff in the whole basin is summed and the discharge in a
specific measurement point could not be used until Schröter et al. (2005) added flow-net data to the basin boundaries. Routing is used in the version of Macro-PDM implemented by Meigh et al. (1999). WGHM calculates routing in rivers as well as in lakes, wetlands and reservoirs. Water withdrawal is also simulated with WaterGAP.


<table>
<thead>
<tr>
<th></th>
<th>WBM</th>
<th>Macro-PDM</th>
<th>WGHM/WaterGAP2</th>
<th>WASMOD-M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial resolution</td>
<td>0.5° x 0.5°</td>
<td>0.5° x 0.5°</td>
<td>0.5° x 0.5°</td>
<td>0.5° x 0.5°</td>
</tr>
<tr>
<td>Time-step</td>
<td>Pseudo-daily (month)</td>
<td>Daily, output monthly</td>
<td>Daily, output monthly</td>
<td>Monthly</td>
</tr>
<tr>
<td>Climate data</td>
<td>Precip., temp. (mean, ranges, dewpoint), wind, percent sunshine</td>
<td>Pseudo-daily precip. and temp., vapour pressure, cloud, wind</td>
<td>Pseudo-daily precip., temp., wet days, cloud, and average sunshine hours</td>
<td>Precip., temp., vapour pressure</td>
</tr>
<tr>
<td>Within-cell distribution</td>
<td>-</td>
<td>Soil moisture statistically, open water and land</td>
<td>Open water and land</td>
<td>-</td>
</tr>
<tr>
<td>Routing</td>
<td>In WTM</td>
<td>Within-cell</td>
<td>Within-cell, between-cells, lakes/wetlands</td>
<td>-</td>
</tr>
<tr>
<td>Flownet, basin orders</td>
<td>STN-30p</td>
<td>Watersheds from RIVM</td>
<td>DDM30</td>
<td>STN-30p</td>
</tr>
<tr>
<td>Water withdrawal</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Parameters</td>
<td>? (&gt;4)</td>
<td>13</td>
<td>≈ 36</td>
<td>6</td>
</tr>
<tr>
<td>Calibrated parameters</td>
<td>In WTM only</td>
<td>Simple tuning</td>
<td>1, sometimes 1-2 additional correction factors</td>
<td>5</td>
</tr>
<tr>
<td>Regionalisation</td>
<td>-</td>
<td>-</td>
<td>Multiple regression</td>
<td>Window</td>
</tr>
<tr>
<td>Required physiographic datasets</td>
<td>Soil map, land cover, topography, rooting depth</td>
<td>Soil map, land cover</td>
<td>Soil map, land cover (several), leaf mass, albedo, water holding capacity, slope, hydrogeology, digital chart of the world</td>
<td>-</td>
</tr>
</tbody>
</table>

The three main global water-balance models have different approaches to calibration or tuning (paper II). Calibration should ideally be avoided as a substantial part of the world is ungauged and since calibration is questionable when a model should be used for climate change studies (Arnell, 1999c; Hanasaki et al., 2007a), but because of limited data quality, complexity of processes, and subgrid spatial heterogeneity it can be considered necessary (Döll et al., 2003). Model performance of the land-surface models compared in the PILPS project was better for the calibrated models (Wood et al., 1998). This was true for conceptual as well as for physically-based models.
WBM has parameter values assigned *a priori* (Vörösmarty *et al.*, 1998) and it is not calibrated. The parameter values are related to vegetation and soil properties, or assumed constant. Fekete *et al.* (1999a) apply runoff-correction factors in gauged cells to make inflow to downstream areas equal to measured flow. WTM, the water-transport model of WBM is however calibrated (Vörösmarty *et al.*, 1989; Vörösmarty and Moore 1991; Vörösmarty *et al.*, 1996).

Arnell’s (1999c; 2003) approach is to avoid calibration as much as possible. Macro-PDM parameter values are estimated from spatial databases or assumed constant from literature values and using knowledge of previous applications of the model (Arnell 1999c; 2003). Six of the 13 parameters are globally uniform and 7 are functions of soil texture and vegetation (Arnell, 2003). When developing the model, Arnell (1999c) did some tuning to set values and test model sensitivity. The tuning process included tests of precipitation datasets and potential-evaporation calculations and was done against long-term average runoff and long-term average within-year runoff patterns.

The approach of Döll *et al.* (2003) is to calibrate only the runoff regulation parameter of WGHM. Calibration is carried out against measured long-term average runoff to get a simulated maximum error of 1%. This goal is achieved for all 724 stations after applying one or two correction factors in 339 cases to ensure that the downstream stations get simulated runoff similar to the measured. Döll *et al.* (2003) regionalise their calibrated parameter with multiple regression. The other WGHM parameter values are globally uniform or related to land cover and other properties. A non-calibrated version of WGHM also exists (Martina Weiß, pers. comm.).

The models have been used for future global water resources assessments (paper I; Arnell, 1999a; 2003; 2004; Vörösmarty, *et al.*, 2000a; Alcamo and Henrichs, 2002; Alcamo *et al.*, 2003b) as well as for regional assessments (Arnell, 1999b; 2005; Schröter *et al.*, 2005; Lehner *et al.*, 2006). Kaspar (2004) and (Lehner *et al.*, 2006) have investigated the validity of using WGHM for climate change studies. Lehner *et al.*, (2006) studied WGHM’s ability to reproduce high and low flows before model simulations of hydrological extremes in a global change scenario and Kaspar (2004) found that the impact of climate scenarios on runoff was larger than the variations caused by parameter uncertainty.

WGHM/WaterGAP is probably the most well-tested model of the three. Model efficiency has been calculated with volume error (Eq. 19) and Nash (Eq. 20) efficiency measures for both monthly and annual time series, and observed and simulated high- and low-flow characteristics have been compared. Validation of the calibration and regionalisation are reported for 6 stations upstream of a calibration station and for 3 regionalised basins (Döll *et al.*, 2003). Monte Carlo test of WGHM are made by Kaspar (2004) and Güntner *et al.* (2007b). Kaspar (2004) found, for most of the 35 tested ba-
sins, that the parameters related to lakes and wetlands were most sensitive, and that the impact of climate change scenarios is stronger than the uncertainty caused by the uncertainty in the parameters. Güntner et al., (2007b) investigate the sensitivity of the total water storage (sum of snow, soil moisture, groundwater and surface water storage) in 22 basins and find strong regional variations in parameter sensitivity depending on which processes that are important in different basins.

Although only few parameter values were tested, Arnell (1999c), in developing Macro-PDM, show the effect on the average runoff of varying some parameter values between a few fixed values for 24 European model grid cells, before deciding upon a common best value. Arnell (2005) made a similar analysis for the Arctic rivers. Arnell (2003) compared volume errors of Macro-PDM with the calibrated and uncalibrated volume errors of Nijssen (2001a). Kleinen and Petschel-Held (2007) use these values for comparison too, as they are easy to find and among the first to be published. Simulated results for all 663 validation basins are reported for WBM by Fekete et al. (1999b). These basins are also used by WASMOD-M.

All three models have been used to test different ways to calculate potential evaporation (Vörösmarty et al., 1998; Arnell, 1999c; Kaspar, 2004) as well as precipitation with and without correction or to use different precipitation datasets (Arnell, 1999c; 2005; Döll et al., 2003; Fiedler and Döll, 2007; Fekete et al., 2004).

Model performance can be compared to measurements other than runoff as a check of the model’s internal consistency. WGHM simulations are compared with water storage variations from the GRACE satellite observations by Schmidt et al. (2006) whereas Werth et al. (2007) use these data for calibration. Comparison against satellite derived snow cover and runoff measurements are presented by Schulze and Döll (2004) in a test of a new, subgrid, snow routine for WGHM. Modified versions of WBM are used in comparison with remotely sensed snow (Rawlins et al., 2005; 2007) and isotope data (Fekete et al., 2006).

The global water balance models can be applied for a specific region as well (Vörösmarty and Moore, 1991; Vörösmarty et al., 1996, Arnell, 1999b; Arnell, 2005; Schröter et al., 2005; Lehner et al., 2006). Additionally, there exist several water balance models that have been applied over large areas but not globally. To identify the complete list is beyond the scope of this thesis, but some examples are the models by (Yates, 1997; Ewen et al., 1999; Graham, 1999; Guo et al., 2002; Yang and Musiake, 2003; Collischonn et al., 2007).
Uncertainties of GCMs and LSMs and the effect in studies of hydrological impact of climate change

The development of the LSMs have raised the question of overparameterisation (Hogue et al., 2006; Demaria et al., 2007) and the question if the LSMs despite their complexity perform as well as expected (Abramowitz, 2005). Even if the complexity sometimes can be motivated (Desborough, 1999) it does not help to have a good LSM, when the GCMs have big problems in simulating the precipitation which is the main driving force of runoff (Mearns et al., 1995; Arora, 2001; Paper I). Several scenarios have been set up that project future carbon dioxide levels based on, e.g., “business-as-usual” or more environmental-friendly world development together with population scenarios, for prediction of future climate conditions (Arnell, 2004). The differences between individual GCMs can be larger than between the scenarios (Graham, 2004). Because of the uncertainties in the GCMs the “delta change” approach is often used instead of the direct GCM output (e.g., Graham, 2004; Fronzek and Carter, 2007; paper I). This approach adds the relative change between a GCM’s climate prediction and its present climate simulations to a measured time series before feeding it to a hydrological model. It is assumed that relative climate changes are more reliably simulated by the GCMs than their absolute values (Hay et al., 2000).

Paper I describes six approaches to hydrological impact studies of climate change. These approaches are 1) direct usage of GCM output, 2) downscaling with regional climate models, 3) global water balance models, 4) continental-scale water balance and hydrological models (including LSMs) 5) hypothesised scenarios, and 6) statistical downscaling. The interest in hydrological impact is often on local and regional scale, which the GCM resolution cannot depict. Downscaling methods are thus needed to get higher resolution projections of climate change impact. Regional climate models and statistical downscaling focus on an area of interest and translate the low-resolution GCM results to this area. These methods are seldom or never applied globally. Hypothesized scenarios are quick solutions to modify existing climate time series without accessing GCM output data and considering the large errors in the GCM output. The approach is used in local and regional studies (paper I) but not in global water balance models, although it would be possible. Usage of “delta change” GCM output that gives some geographical pattern should be slightly more realistic than, e.g., +2° for the whole globe, as long as several GCMs are used. It should be kept in mind that the GCMs are driven by global scenarios themselves, of CO₂ and other emissions based on, e.g., assumed population changes (Arnell, 2004).

Impact of climate change needs to be studied locally and regionally, but the global analysis is also valuable. Global and continental water balance models, coupled or stand-alone LSMs and direct GCM output with routing can be used to study impacts on the continental scale. The uncertainty of the
GCMs is the main limitation of all approaches since it is the basis for the impact studies. Additional uncertainties come from errors in the regional climate models, errors in the hydrological models and lack of data (paper I). Further climate-hydrology feedback studies, including vegetation and further dialogue between modellers, the remote sensing community and other measurement communities will probably decrease the uncertainties (paper I; Gerten 2004; Overgaard et al., 2006).
Methods

The models used in this thesis are based on WASMOD\(^1\) (Water And Snow balance MODeling system) catchment model (Xu, 2002). While the original WASMOD is lumped, the two models used here are distributed. WASMOD-M, with M for macro-scale, is the monthly, global water balance model that is the focus of the thesis (paper II; III). The model version applied in paper IV, is a daily, distributed model, but it is still denoted WASMOD. It was developed to study the impact of spatial resolution for future versions and input-data sets of WASMOD-M. The simple structure and the few parameters of the original WASMOD were reasons why it was chosen for the first version of the new global water balance model. Additionally, different versions of WASMOD have shown a high level of generality from applications at variously-sized catchments in Europe, Asia and Africa. Studies have shown that its model parameters can be correlated with catchment characteristics (Xu, 1999a; 2003; Muller-Wohlfeil et al., 2003) such as land-cover fractions and soil texture. WASMOD (Xu, 1999b), together with the models of Refsgaard and Knudsen (1996) and Donnelly-Makowecchia and Moore (1999), are also among the few water-balance models that passed the full split-sample-proxy-basin test proposed by Klemeš (1986).

Global data for WASMOD-M

WASMOD-M requires gridded monthly precipitation, temperature and potential evaporation as input. A 0.5° × 0.5° latitude-longitude grid was used because data was easily and freely available for this projection and resolution, and because it is used by other global water balance models (Arnell, 1999c; Döll et al., 2003; Vörösmarty et al., 1998). Additional data needed are flow networks with basin boundaries and areas. Compilation of gridded runoff results for continents requires their boundaries too. Measured runoff, at stations co-registered to the flow network, and gridded snow cover measurements have been used in the model tuning.

\(^1\) Not to be confused with WASMOD (Water and Substance Simulation Model) developed by Ernst-Walter Reiche, and described by, e.g., Rinker (2001) in Beschreibung der Wasser- und Stoffhaushaltsdynamik devastierter Flächen mit dem Simulationsmodell WASMOD am Beispiel des Braunkohletagebaus Espenhain, Der Fakultät für Geowissenschaften, Geotechnik und Bergbau der Technischen Universität Bergakademie Freiberg
Climate-input data to WASMOD-M comprised gridded monthly values of precipitation, temperature and water-vapour pressure from CRU TS 2.02 (Mitchell et al., 2004) for paper II and CRU TS 2.10 (Mitchell and Jones, 2005) for paper III. Data covering 1901-2000 have been used. Precipitation-correction factors were calculated for gauge losses as quotients for each cell between corrected and uncorrected precipitation (Willmott et al., 1998; Legates and Willmott, 1990) as a long-term average for each month.

Gridded potential evaporation $ep$ was calculated from CRU air temperature $T_a$ (°C) and water vapour pressure $e_a$ (mbar = hPa) through the relative humidity $RH$ (%) (Dingman, 1994, Xu 2002):

$$RH = \frac{100 \times e_a}{e_{sat}(T_a)} = \frac{100 \times e_a}{6.108 \exp \left( \frac{17.27 \times T_a}{T_a + 237.3} \right)}$$  \hspace{1cm} (2)

$$ep = E_c \times \left( T_a^+ \right)^2 \times \left( 100 - RH \right)$$  \hspace{1cm} (3)

where $T_a^+ = \max (T_a, 0)$, $e_{sat}(T_a)$ is the saturated vapour pressure at temperature $T_a$, and $E_c$ (mm month$^{-1}$ °C$^{-2}$) is a non-calibrated model parameter. The calculation of $e_{sat}(T_a)$ is an empirical formula (Dingman, 1994) which holds for momentary values. Theoretically it should not be used for monthly averages, but such errors are compensated for by the $E_c$ adjustment factor. In paper II $E_c$ was set to 0.018 globally, while in paper III distributed values of $E_c$ were used. They were set in an inverse process, to get the average annual potential evaporation equal to the highest value in two potential evaporation datasets. This produced $E_c$ values ranging from $2.61 \times 10^{-4}$ (arid areas) to 0.999 (cold areas), with a median of 0.011.

Basin boundaries and areas, flow paths, and continent boundaries were taken from STN-30p (Vörösmarty et al., 2000b). 663 GRDC runoff stations have been co-registered to the flow network in "UNH/GRDC Composite Runoff Fields V1.0" (Fekete et al., 1999a; 1999b; 2002). These stations have basin or inter-station areas sufficiently large to fit the 0.5° grid, and some stations with inconsistencies in the inter-station runoff were removed, while 28 stations with negative inter-station runoff more likely caused by evaporation and water withdrawal rather than measurement inconsistencies were kept (Fekete et al., 1999a). These stations belong to 257 basins discharging to oceans or large lakes. The gauged areas cover 50% of the 59132 STN-30p land-surface cells. Related to the actively discharging area only, the coverage is 72% instead of 50% (Fekete et al., 2002).

It would have been useful to have regulation data since almost all large rivers are regulated (Vörösmarty et al., 2004; Nilsson et al., 2005) such that natural peak- and low-flow characteristics are modified. Regulation data are
often unavailable since dam and reservoir operation commonly involve competing economic interests and can also be in conflict with downstream water-supply needs (Brakenridge et al., 2005) and there exist no global database of naturalised streamflow, with flows recalculated to the natural behaviour where anthropogenic effects like abstraction and damming have been subtracted. Neither regulation, nor natural routing delays by long river stretches, lakes and wetlands have been included in WASMOD-M. The long-term average runoff was used for model tuning to overcome this effect in paper II. We assumed that regulation did not affect average flow volumes, a non-valid assumption in warm and arid areas with substantial evaporation loss from river stretches and reservoirs, or where large withdrawals leads to losses. The average runoff of the 663 stations were taken from “UNH/GRDC Composite Runoff Fields V1.0” (Fekete et al., 1999b). The long-term average within-year variations were also reported in the dataset and used in validation. These averages are based on different time periods during the 20th century and some include data from the 19th century.

Model tuning against monthly time series was tried in paper III to better restrict the parameter values and avoid equifinality, despite the lack of routing in WASMOD-M. Very few rivers have regulation delays longer than one year (Vörösmarty et al. 1997), and the model was also tuned against annual data to overcome the regulation delay. The time-series for the 663 stations were provided directly by GRDC (2007). This dataset contains two monthly time-series for some stations, which more or less overlap. These are “original” monthly data, and “calculated” monthly data. Only the “original” monthly data were used in paper III, restricting the number of time-series to 654.

The combination of precipitation and runoff data showed that some basins had physically unreasonable runoff coefficients (total runoff divided by total precipitation). Only stations with runoff coefficients between 1% and 85% were used in paper II and regions with abundant runoff coefficient problems were excluded altogether, e.g., Alaska and the Colorado River basin. All stations were included in paper III, even if they had data problems.

Calibration of snow parameters was performed against the 1986–1995 “Northern Hemisphere Monthly Snow Cover Extent” 0.5° × 0.5° latitude-longitude dataset from ISLSCP (2003) and the National Snow and Ice Data Center (Armstrong and Brodzik, 2005).

Simulation time periods, model windup times and initial values
All simulations started with globally uniform initial values of land moisture and snowpack. Simulations in paper II always started 1901 and had at least a 14-year warm-up period. The parameter tuning procedure was applied for 1915–2000 and the simulations from the time period where gauge data were available for station were extracted. The time periods could not match when
the measured average runoffs included data from 1914 and earlier. Compari-
sions with other estimates were calculated for the standard period 1961-1990.
Gauge data before 1901 and after 2000 were shortened to match the climate
data in paper III. Model warm-up always started at least 5 years before. The
simulated and measured runoff time-series were divided in two for tuning
and validation purposes, making the total windup time much longer for the
validation. Snow calibration was made for the snow data period, i.e., 1986–
1995, with a windup of 5 years.

**WASMOD-M**

Two versions of WASMOD-M have been used. The first, used in paper II,
took its model equations directly from the code by Xu (2002). The changes
in the latter version, used in paper III, were made to 1) avoid non-closure of
the water balance by preventing the fast flow to exceed available water, 2)
be able to regulate all evaporation by a parameter, by removing the non-
parameter-governed direct loss. The model equations specific for paper II
are denoted “a”, and those for paper III “b”.

WASMOD-M calculates, with a monthly time step, snow accumulation
and melt, actual evaporation, and separates runoff into a fast and a slow
component for each grid cell. Given the monthly time step, the model simult-
aneously allows snowfall, rainfall, and snowmelt to occur in the same
month. The model has 6\(^2\) parameters of which 5 are calibrated and one is
pre-set (Table 4). All units are mm month\(^{-1}\) unless stated otherwise.

Snow- and rainfall (snow and rain) as well as snowmelt (melt) and accu-
mulation (sp) vary exponentially between two temperature thresholds
\((T_m, T_s, ^\circ C)\):

\[
snow = P \times \left(1 - \exp \left(-\left(\frac{(T_a - T_s)}{(T_s - T_m)}\right)^2\right)\right)
\]

\[
(4)
\]

\[
\text{rain} = P - snow
\]

\[
(5)
\]

\[
melt = (sp_{old}/\Delta t + snow) \times \left(1 - \exp \left(-\left(\frac{(T_m - T_a)}{(T_s - T_m)}\right)^2\right)\right)
\]

\[
(6)
\]

\[
sp = sp_{old} + (snow - melt) \times \Delta t
\]

\[
(7)
\]

where \(P\) is precipitation and \(\Delta t = 1\) month and \(\{x\}^-\) means \(\min(x, 0)\).

\(^2\) In paper II 4-6 parameters as snow only was calculated for the “cold” cells, while the second
model version (paper III) did not include this difference.
Direct water loss \((dl)\) to the atmosphere was calculated from potential evaporation and rainfall in paper II. The variable “land moisture” \((lm, \text{ in mm})\) represents the storage of water available for evaporation and runoff in the next time step. The name is introduced to make a clear distinction to the well-defined, local-scale, soil-physical entity “soil moisture”. Available water \((aw)\) is calculated from active rain, \(i.e.,\) rainfall minus direct loss, and snowmelt. Actual evaporation \((evaptot)\) is calculated from land moisture, potential evaporation \((ep)\), and direct loss:

\[
\begin{align*}
\begin{cases}
  dl = ep \times (1 - \exp(-\text{rain}/ep)) & ep > 1 \\
  dl = ep & ep \leq 1
\end{cases}
\end{align*}
\]

\(8a\)

\[
activerain = \{\text{rain} - dl\}^+
\]

\(9a\)

\[
aw = lm_{old}/\Delta t + activerain + melt
\]

\(10a\)

\[
\begin{align*}
\begin{cases}
  evappart = \min\{(ep - dl) \times (1 - A_c^{aw/(ep-dl)}), aw\} & ep - dl > 1 \\
  evappart = ep - dl & ep - dl \leq 1
\end{cases}
\end{align*}
\]

\(11a\)

\[
evap = \{evappart + dl\}^+
\]

\(12a\)

where \(A_c\) \((-\) is a tuneable parameter and \(\{x\}^+\) means \(\max(x,0)\). The direct loss was removed to let \(A_c\) control all evaporation in paper III. Thus equations 8a to 12a are substituted by 13b-14b

\[
aw = lm_{old}/\Delta t + \text{rain} + \text{melt}
\]

\(13b\)

\[
evap = \min\{ep \times (1 - A_c^{aw/ep}), aw\}
\]

\(14b\)

The slow runoff is a base flow, provided by land moisture whereas the fast runoff is provided by both land moisture and water added \(i.e.,\) melt + (active)rain) during a time step. A water availability check was introduced in paper III. Both runoffs are described by linear reservoirs:

\[
\begin{align*}
\text{slow} &= P_s \times lm_{old} \\
\text{fast} &= P_f \times lm_{old} \times (\text{melt} + \text{activerain})
\end{align*}
\]

\(15\)

\(16a\)
\[ \text{fast} = P_f \times \text{lm}_{\text{old}} \times (\text{melt} + \text{rain}) \]  
(16b)

\[ \text{runoff} = \text{slow} + \text{fast} \]  
(17a)

\[ \text{runoff} = \min\{(\text{slow} + \text{fast}), (\text{aw} - \text{evap})\} \]  
(17b)

where \( P_s \) (\text{month}^{-1}) and \( P_f \) (\text{mm}^{-1}) are tuneable parameters. Finally the land moisture storage is updated.

\[ \text{lm} = \text{lm}_{\text{old}} + (\text{active} + \text{melt} - \text{evappart} - \text{runoff}) \times \Delta t \]  
(18a)

\[ \text{lm} = \text{lm}_{\text{old}} + (\text{rain} + \text{melt} - \text{evap} - \text{runoff}) \times \Delta t \]  
(18b)

The first model version was written in Fortran 90, while the second was written in Matlab.

Model parameter tuning of WASMOD-M in gauged areas

Two different parameter estimation procedures have been applied, one in paper II and the other in paper III. Focus of both has been to identify acceptable parameter value sets.

Table 4. The complete set of parameters in WASMOD-M, and the ranges of the 5 parameters varied in the parameter tuning procedure. \( E_c \) is pre-processed.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Governs</th>
<th>Eq.</th>
<th>Fixed values (paper II)</th>
<th>Range (paper III)</th>
<th>Sampling interval (paper III)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_s ) (°C)</td>
<td>Snowfall</td>
<td>4, 6</td>
<td>1.5, 3.5</td>
<td>0–4</td>
<td>Uniform</td>
</tr>
<tr>
<td>( T_m ) (°C)</td>
<td>Snowmelt</td>
<td>4, 6</td>
<td>-1.5, -3.5</td>
<td>-4–0</td>
<td>Uniform</td>
</tr>
<tr>
<td>( A_e ) (-)</td>
<td>Actual evaporation</td>
<td>11a/14b</td>
<td>0.35, 0.55, 0.75</td>
<td>0–1</td>
<td>Uniform</td>
</tr>
<tr>
<td>( P_s ) (mon^{-1})</td>
<td>Slow runoff</td>
<td>15</td>
<td>( 5 \times 10^{-5}, 1 \times 10^{-4}, 0.0015, 0.004, 0.01, 0.025, 0.05, 0.14, 0.19, 0.25, 0.4, 0.6, 0.8, 0.9 )</td>
<td>( e^{-14} - e^{-0.2} )</td>
<td>Logarithmic</td>
</tr>
<tr>
<td>( P_f ) (mm^{-1})</td>
<td>Fast runoff</td>
<td>16</td>
<td>( 5 \times 10^{-6}, 1 \times 10^{-5}, 1 \times 10^{-4}, 5 \times 10^{-4}, 0.001, 0.0025, 0.005, 0.0095, 0.0145, 0.025 )</td>
<td>( e^{-18} - e^{-0.2} )</td>
<td>Logarithmic</td>
</tr>
<tr>
<td>( E_c ) (mm( \text{mon}^{-1} \text{ °C}^{-2} ))</td>
<td>Potential evaporation</td>
<td>3</td>
<td>0.0018</td>
<td>2.61 \times 10^{-4} – 0.999</td>
<td>Pre-processed</td>
</tr>
</tbody>
</table>

\( a) e^{-14} \approx 8.31 \times 10^{-7}, \quad b) e^{-18} \approx 1.52 \times 10^{-8} \)
The sets were used for regionalisation only in paper II, while a single parameter value sets was selected for a final simulation, but all acceptable sets were kept for all simulations in paper III. Monte Carlo-approaches were used. The model was first run with a combination of globally-uniform parameter-value sets, covering the range of “reasonable” values for each parameter (Table 4) in paper II. The most sensitive parameters were given the broadest range of possible values. The combination of the discrete values for all the parameters resulted in 1680 simulations. Only data from gauged areas were saved for further analysis. 15000 random parameter value combinations were tested for the whole upstream area of each basin in paper III. Parameter values were uniformly or logarithmically divided within the wide ranges and randomly combined.

Criteria

The simulated long-term average runoffs were compared with the measured values, through the volume error ($VE$) in paper II. $\text{Abs}(VE)$ varies from perfect fit at 0 to plus infinity:

$$VE = \frac{\sum_{\text{time}} \text{runoff}^{\text{sim}} - \sum_{\text{time}} \text{runoff}^{\text{obs}}}{\sum_{\text{time}} \text{runoff}^{\text{obs}}} \times 100\%$$

(19)

where $\text{runoff}^{\text{obs}}$ is the simulated runoff at every time step and $\text{runoff}^{\text{sim}}$ the corresponding simulated runoff. Parameter-value sets in paper II were considered “acceptable”, if the simulated basin runoff differed by less than ±20%, or 5 mm in dry areas, respectively, from the measured value. Several limits were used in paper III (Table 5). Limitations of the state variables (snowpack and land-moisture) were used in paper II. Land moisture was not allowed to exceed 1000 mm in any month while the snowpack limit was 1200 mm. The average land moisture and snowpack respectively during the first simulation year (1915) was not allowed to differ more than 200 mm and 100 mm respectively compared to the average of the last simulation year (2000). The removal of the state-variable limits was also investigated.

The measured time-series were considered in paper III, in addition to $VE$. The Nash coefficient ($NC$), and the limit of acceptability ($LA$) were used. Tuning against snow cover data was also used, with a snow/no-snow criteria ($SF$). The model was calibrated against monthly as well as annual average time series for each basin in paper III. Simulations were always made with monthly input and monthly timesteps, but the monthly output was also averaged to yearly time series. Evaluation criteria were calculated separately for the two output time series.

$NC$ varies from perfect fit at 1 to minus infinity:
\[
NC = 1 - \frac{\sum_{time} (\text{runoff}_{\text{obs}} - \text{runoff}_{\text{sim}})^2}{\sum_{time} (\text{runoff}_{\text{obs}} - \text{runoff}_{\text{obs}})^2}
\]  

(20)

where \(\bar{\cdot}\) denotes average.

The limit of acceptability (\(LA\)) criterion is presented by (Beven, 2006). It requires a modeller to predefine acceptable simulation errors, based on “effective observation error” of input data and discharge measurements. These limits can vary in time. Simulated runoff that falls within the acceptable limits at a given point in time is weighted with, e.g., a triangular or a trapezoidal function where a simulation close to the measured is given 100% weight whereas a simulation outside limits is given zero weight. The choice of predefined error limits was not obvious in the global case and subjective, wide limits were used. This was motivated since GRDC do not generally report rating-curve errors and the model-input data are uncertain. A symmetrical, triangular weighting function (with a zero minimum and a unit maximum) was used and \(LA\) was calculated as a time average.

We defined parameter-value sets to be “behavioral” in two ways. A first, relative definition was the selection of the best 1%, 3.3%, and 20% (150, 500, and 3000 of 15000) simulations for each criterion. A second definition used the absolute limits for each criterion (Table 5).

\(LA\) was defined by a range around the measured flow that simulated flows had to meet at least 95% of the time, i.e., less than 3 months in 5 years was allowed to fall outside of range. The initial range (\(LA_{75}\)) was given as \(\pm 75\%\) of the flow at each time step plus 3 mm to avoid high relative low-flow errors. If a sufficient number of simulations did not meet this criterion, we widened the range to \(\pm 99\%\) of the flow plus 3 mm (\(LA_{99}\)). If this was not enough, we widened the range until we got the requested number of simulations (\(LA_{\text{max}}\)).

Snow calculation in WASMOD-M is independent of the land moisture storage and evaporation and could thus be calculated separately by using the input precipitation and temperature only. Snow calibration compared the existence of snow or not. The snowpack (\(sp\)), in mm, simulated by WASMOD-M is a representation of the snowpack at the last day of the month, available for melting next month. The observations give the percentage of the weeks within a month with spatial snow coverage more than 50%. The 100% and some of the 0% values were used as only these have information on the situation at the end of the month while the intermediate values are not connected with any information about when snow cover occurred during the weeks of the month. The months with 0% values were not used during Dec-Feb and when they were adjacent to a month with some snow above 50% spatial coverage. During the other months the 0% months were assumed to
represent true no-snow conditions although the 0% value only imply less than 50% spatial snow coverage during the whole month. The snow-fit criterion was based on measured and simulated snow periods:

\[
SF = \min \left( \frac{\text{snowmonths}_{\text{correct}}}{\text{snowmonths}_{\text{tot}}}, \frac{\text{nonsnowmonths}_{\text{correct}}}{\text{nonsnowmonths}_{\text{tot}}} \right)
\]  

(21)

where \(\text{snowmonths}_{\text{correct}}\) is the number of months with a simulated snowpack fitting months with measured snow and \(\text{nonsnowmonth}_{\text{correct}}\) the number of simulated months with no snowpack fitting measured snowfree months, \(\text{snowmonths}_{\text{tot}}\) is the total number of snowfree months and \(\text{nonsnowmonths}_{\text{tot}}\). The number of months is a total measure for all cells in the basin. \(SF\) varies from perfect fit at 1 to no fit at 0. Of the 344 northern hemisphere basins with snow measurements all had some cells with non-snow-classification, while 26 stations were lacking snow classification as none of the included measured cells had a snow cover for 100% of the time any month.

Inter-station areas

Of the 663 runoff stations, 208 represent single-station basins and 234 headwater sub-basins. These 442 basins are geographically independent, and were calibrated independently from each other. The remaining 221 stations are downstream stations, each of which is associated with one geographically independent inter-station sub-basin. Among the downstream stations, 90 are down-most and 13 of these have 10 or more upstream stations. The 1680 and 15000 simulations were applied with uniform parameter value combinations of the whole upstream area of each gauged station. An inter-station algorithm to tune the parameters of the inter-station areas was developed in paper II. The inter-station runoff was calculated as the difference between the upstream and the downstream average runoff. Such a simple procedure could not be applied to the time-series used in paper III because of routing delays. All 654 basins were instead treated separately in paper III.

The inter-station algorithm of paper II focused on improving the runoff at the downmost basins compared to what was achieved with the uniform parameter value combinations. It required that the runoff coefficient of the inter-station area was acceptable. Inter-station runoff was not considered if the sum of measured upstream and simulated inter-station runoff produced unacceptable results, \(i.e., VE\) larger than ±20%. The inter-station algorithm was also overruled if it increased the error at the downmost gauging station compared to a calibration for the whole basin against runoff at this station.

The single best parameter-value combination that was found for each sub-basin by the inter-station algorithm was used in the final, distributed parameter-value set.
Regionalisation of WASMOD-M’s parameter values to ungauged areas

Around 50% of the model’s land-surface cells are ungauged. The method of assigning parameter values to ungauged areas is called regionalisation. A simple regionalisation procedure was used in paper II, using proximity only. A window 17 cells high and 39 cells wide was applied around each gauged cell and cells belonging to gauged basins within the window were identified. Common parameter value combinations of these gauged stations were identified. The most commonly occurring parameter-value set among the gauged basins in this window was selected for each given cell, and if two or more parameter-value sets were equally common, the combination that gave the lowest error in the gauged areas was selected. Sub-areas that had been overwritten with downstream parameter-value sets donated those sets in the regionalisation. Basins with no parameter value combinations got regionalised parameter values. Cells belonging to basins where the 20%-error limit was not reached was not used in the regionalisation and were given their best available parameter-value set with an error above 20% for the final distributed parameter-value set.

Cells where no gauged basin or common parameter-value set was found within the moving window were given a single, default value. This default parameter-value set was subjectively derived from a map of acceptable parameter values for all gauged and regionalised cells. The map showed a few dominating parameter-value sets close to the areas with missing values. The most abundant of these was chosen as default.

Evaluation and validation of WASMOD-M

The within-year runoff patterns was used for model validation in paper II since parameter values were only tuned to the long-term average runoff. The long-term average within-year variations were plotted for all stations. A set of 15 gauging stations was especially studied since the result could be compared with results in at least two previous publications using other models. Total global and continental runoff with the final parameter value set was also compared with previously published estimates (Table 1). The availability of time-series allowed split-sample validation in paper III. Calibration was made against the first half of the available time-series and validation against the second half.
Routing of daily runoff produced by WASMOD

The results of paper II and III highlighted the need for river routing in WASMOD-M. It was also found that newly started studies of the precipitation heterogeneity at different scales needed a scale-independent routing method. A new routing algorithm using aggregated time-delay-histograms was developed and compared to the widely used cell-to-cell linear reservoir routing in paper IV.

Site and data description

The routing study was performed in the Dongjiang (East River) catchment, situated just north of Hong Kong in southern China (Figure 1). It is a tributary of the Pearl River. Runoff was simulated for the catchment area of the downmost gauging station Boluo. Its catchment area is 25,325 km². Climate data input was precipitation at 51 stations and daily air temperature, sunshine duration, relative humidity and wind-speed at the meteorological stations. Final input to this WASMOD version was precipitation, temperature and potential evaporation. The potential evaporation was calculated with the Penman equation for each climate station and then interpolated.

The global 1km flow network Hydro1k (USGS, 1999) was used. It has an equal area projection such that every cell is 1km². Several topographic related properties are included in Hydro1k and the slope was also used apart from the flow network and area. The Boluo discharge station was co-registered to Hydro1k by comparing coordinates and especially area. The best fit was achieved for a flow network of 25,886 cells (km²).

Resolutions

Model simulations were carried out for 5°–60' resolution, with a 1'-step. The input climate data was interpolated to the cells of Hydro1k. Kriging with linear variogram was used to interpolate precipitation and inverse distance weighting was used to interpolate the other datasets. The data were then averaged to each resolution. Uniformly distributed climate inputs were also used at each resolution with both routing methods, to identify the pure routing scale effect.

The linear reservoir routing requires an upscaled network. The network scaling algorithm (NSA) by Fekete et al. (2001) was used to generate a river network at each resolution consistent with the high-resolution Hydro1k. The proposed meandering factor to compensate for decreasing river-lengths with coarser resolutions was not used since this was implicitly accounted for in the calibration. Neither were the basin border cells considered to be either inside or outside of the catchment, but given the area that corresponds to the
actual catchment part of the cell. The runoff generating area thus remained the same for all resolutions.

The 30′ network STN-30p delineates the Dongjiang catchment to 10 cells (Figure 1).

Figure 1. The Dongjiang catchment (dark grey) in the STN-30p network. The surrounding basins (black outline) gauging stations (black dots) and their upstream sub-basins (light grey) are also shown.

Daily, distributed WASMOD

Dongjiang is a snow-free catchment and the snow algorithm was excluded from the model. The evaporation calculation was the same as in paper III. The fast and slow runoff calculations were changed to non-linear exponential formulations for this daily model version:

\[ SP = 1 - \exp\left(- \frac{lm_{old}}{c_1}\right) \]  \hspace{1cm} (22)

\[ fast = \left(P - evap\right) \times SP \]  \hspace{1cm} (23)

\[ slow = lm_{old} \times \left(1 - \exp\left(- \frac{lm_{old}}{c_2}\right)\right) \]  \hspace{1cm} (24)

where \( SP \) is the percentage of the catchment which is saturated, and \( c_1 \) and \( c_2 \) are runoff parameters. The three parameter values were varied in 250 simu-
lations using a similar approach as in paper III. The 250 runoff simulations for each resolution were saved to be used in the routing.

Linear reservoir routing

A linear reservoir approach is most commonly used in global cell-to-cell runoff routing and was implemented for comparison purposes. The linear reservoir (Eq. 1), which is a simplification of the Muskingum method without any influence of the inflow to the storage was combined with the water balance of each grid-cell

\[
\frac{dS}{dt} = Q_{in} + \text{runoff} - Q_{out} \tag{25}
\]

where \( t \) is time and \( Q_{in} \) is inflow to the cell from surrounding cells and \( \text{runoff} \) is the runoff generated in the cell itself, to

\[
k \frac{dQ_{out}}{dt} = Q_{in} + \text{runoff} - Q_{out} \tag{26}
\]

Linear reservoir routing is often applied with a finite difference approximation of Eq. 26 (e.g., Arora, et al. 1999) and was used here to determine the outflow of each cell. The travel time \( k \) between the grid-cells was set by the velocity and distance as in Eq. 1. The calibration of the velocity is described below.

Aggregated time-delay-histogram routing

The aggregated time-delay-histogram routing is a new routing method based on the time-delay histogram of TOPMODEL (Beven and Kirkby, 1979). It uses the knowledge from a high-resolution flow network at lower spatial resolutions. A parameter that is easy to aggregate in space at different resolutions is wanted. The method is based on pre-processing, where the time from each Hydro1k cell to a downstream cell is calculated and then aggregated to a lower resolution. One downstream destination Hydro1k cell, the one with the co-registered Boluo catchment station, was used in paper IV. The source and destination cells are selected among the high-resolution cells for the most accurate time and area representation.

Two parameters are calculated for each upstream high-resolution cells, the delay time \( t \), and the first-day percentage \( p \). The delay time is the total accumulated time it takes for the water from the cell to reach the destination cell. The distances between the cells along the flow paths are denoted \( d_1, d_2, \ldots \).
\[ t = \sum_{i=1}^{n} \frac{d_i}{v_{45} \sqrt{\tan(b_i)}} \] (27)

The network velocity \( v_{45} \) is calibrated separately and requires the possible runoffs simulated with WASMOD as input in the calibration. The runoff simulated by WASMOD was assumed to be evenly distributed in time during one day. This runoff will reach the sink during 24 hours of two consecutive days since no dispersion was assumed. The percentage of runoff that reaches the sink during the first day of arrival is denoted \( p \). The remaining 100\%-\( p \) arrives the following day. The delay time \( t \) is the time to the first arrival, and \( p \) describes how much of the day is left when this first runoff arrives.

The \( t \) and \( p \) of the Hydro1k cells in a low-resolution cell are combined to a time-delay histogram of the low-resolution cell instead of a coarse average. The time-delay histogram summarises the delays and first and second day percentages all Hydro1k cells in the low-resolution cell. An underlying assumption is that the runoff from the low-resolution cell is evenly distributed over all high-resolution cells. There is a possibility to include sub-grid variation too, but that would complicate the comparisons presented here.

The final routing is simple. The aggregated time-delay histogram is used to partition the generated runoff into the corresponding delayed days of time series at the destination cell with the percentage share of each day.

**Velocity calibration**

Both routing methods use one calibration parameter, the velocity, which is calibrated separately from the WASMOD simulations. The 250 runoff simulations of each resolution were combined with different velocities. Each tested velocity produced its own set of time-delay histograms in the time-delay-histogram method. Velocities and WASMOD parameters that gave the highest NC values compared to Boluo runoff were selected. Wide ranges of the velocities were initially tested to identify narrower ranges for the final calibration. For linear reservoir routing, the best velocity \( v \) was within \([0, 1]\) ms\(^{-1}\), and the time-delay-histogram routing velocity \( v_{45} \) was found to fall within \([7, 9]\) ms\(^{-1}\). These two velocities are not comparable as \( v_{45} \) gives the
velocity of a theoretical very steep river. A sampling interval of 0.1 ms$^{-1}$ was used to select the velocities for the final calibration.

**Benchmark**

The time-delay-histogram and the linear-reservoir routing are two very different methods and simulation differences might be complicated to explain. A combined method, that applies a time-delay method to the coarse resolution network, without the usage of the high-resolution data, was used to better differentiate between the resolution and methodological effect on the results. This benchmark routing was also applied at all scales and with the distributed and uniform climate input. The same velocity calibration procedure was applied.
Results

Tuning and criteria

The simulations with 1680 globally uniform parameter-value sets produced acceptable results for the whole upstream area that comprises the main-stream sub-basins of 485 of the 663 stations in paper \textbf{II}. The 178 non-accepted stations were divided into the following four categories: In the first category, 57 stations lacked input and validation data that produced acceptable runoff coefficients. The other three categories were the result of either data or model deficiencies: No acceptable parameter-value sets could be found for 29 stations, \textit{i.e.}, they exceeded the 20% $VE$ limit; 23 stations had unreasonable state variables for all parameter-value sets (\textit{e.g.}, the snowpack or the land moisture grew too large); 69 stations had no acceptable parameter-value set with acceptable state variables. After applying the inter-station algorithm, parameter values could be set for 531 of the total 663 stations. Since some of them were substituted by parameter values from downstream stations, the number of parameter-value sets was finally 429. The runoff was simulated with an absolute volume error less than 20% for the whole upstream area of 455 basins, and within 1% for 276 basins, with the final, distributed parameter value set.

Several different model evaluation criteria were calculated separately in paper \textbf{III}. The volume error was the easiest criteria to fulfil. All basins could be simulated with $VE$ less than $\pm 1\%$ as long as the basins had reasonable runoff coefficients. The improvement in paper \textbf{III} compared to paper \textbf{II} could be attributed to the larger number of simulations, 15000 instead of 1680, and the removal of the “direct loss” in the model algorithms. A clear majority of the stations had parameter-value sets that fulfilled the $LA_{75}$ criterion. The existence of snow or not was also well simulated. $NC$ was the strictest criterion, with only 479 basins that could be simulated with a performance of at least 0.5. This number is comparable to the 485 acceptable basins in paper \textbf{II}.

The number of basins that fall within a certain criteria limit holds for the individual criteria. The combinations of two or several criteria were analysed in paper \textbf{III}. 52\% of the basins could be simulated by parameter value sets that simultaneously fulfilled $NC \geq 0.5$, $VE \leq 1\%$, $SF \geq 0.75$, and $LA_{75}$. It was possible to find good runoff parameter-value sets concurrent with the best 1\% snow parameter-value sets for almost all snow-covered basins. It was
also possible to find common sets between the 1% best of NC and VE for 99% of all basins. NC and LA behave similarly in parameter space but common sets between the 1% best of them were only found for 83% of all basins. Common sets between the 1% best VE and LA were found for 52% of all basins.

Table 5. Number of basins with at least one parameter-value set resulting in simulated results fulfilling different evaluation-criteria limits for Nash (NC), volume error (VE), limit of acceptability (LA), and snow fit (SF) when evaluated against monthly and annual observations in paper III. There were 654 basins with runoff measurements and 344 with snow measurements. Percentages relate to these totals.

<table>
<thead>
<tr>
<th>Criterion Type</th>
<th>Limit</th>
<th>Name</th>
<th>Monthly No</th>
<th>Monthly %</th>
<th>Annual No</th>
<th>Annual %</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC</td>
<td>≥ 0.8</td>
<td>NC0.8</td>
<td>157</td>
<td>24</td>
<td>103</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>≥ 0.5</td>
<td>NC0.5</td>
<td>479</td>
<td>73</td>
<td>363</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>≥ 0</td>
<td>NC0</td>
<td>642</td>
<td>98</td>
<td>560</td>
<td>86</td>
</tr>
<tr>
<td>VE</td>
<td>≤ 1%</td>
<td>VE1</td>
<td>632</td>
<td>97</td>
<td>630</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>≤ 20%</td>
<td>VE20</td>
<td>643</td>
<td>98</td>
<td>643</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>≤ 50%</td>
<td>VE50</td>
<td>651</td>
<td>100</td>
<td>651</td>
<td>100</td>
</tr>
<tr>
<td>LA</td>
<td>± 75% of obs.</td>
<td>LA75</td>
<td>580</td>
<td>89</td>
<td>650</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>± 99% of obs.</td>
<td>LA99</td>
<td>654</td>
<td>100</td>
<td>654</td>
<td>100</td>
</tr>
<tr>
<td>SF</td>
<td>≥ 95%</td>
<td>SF95</td>
<td>274</td>
<td>77</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>≥ 75%</td>
<td>SF75</td>
<td>321</td>
<td>91</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>≥ 50%</td>
<td>SF50</td>
<td>339</td>
<td>96</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Parameter-value behaviour

Being the criteria that is easiest to fulfil, VE, is also the least restrictive. It normally leaves most parameter spaces unconfined (Figure 2). The dynamics criteria NC and LA commonly confine the parameter spaces of the evaporation and runoff criteria. NC has a slight influence on the two snow parameters $T_s$ and $T_m$ but SF was needed to get a substantial constriction of them. The constriction was clear for $T_s$ but less so for $T_m$. Criteria and parameter behaviour differed between snow-covered and non-snow covered (arid and tropical) basins. VE easier confined the evaporation parameter $A_c$ in non-snow basins and the runoff parameters $P_s$ and $P_f$ in snow basins. $A_c$ was also better confined by LA and NC for the non-snow basins, while the runoff parameters were better confined for non-snow basins by the 1% best NC and LA values, but for snow-covered basins with the 20% best NC and LA values.
Figure 2. Runoff performance for River Sénégal at Bakel for combinations of values of the evaporation ($A_c$), slow-flow ($P_s$), and fast-flow ($P_f$) parameters. Calibration was made with the Nash criterion ($NC$) against observed monthly (mon) and annual (yr) observations and with the absolute value of the volume error (%) for the calibration time period (1904 – 1951). Values for the 500 (3.3%) best sets are shown as large dots. (Paper III)

Time aggregation

Equifinality of both runoff parameters, and also of the evaporation parameter, increased when successively calibrated against monthly, annual, and long-term average runoff (Figure 2). The trend was very clear when going from annual to long-term average aggregation but less clear when going from monthly to annual aggregation. Equifinality decreased in some cases from monthly to annual aggregation, especially for the evaporation parameter. It was also evident that $LA_{75}$ was too permissive in combination with annual data (Table 5).
Regionalisation

The 17×39 window reached the majority of the ungauged cells in paper II (Table 6). The regionalisation was successful since a majority of the regionalised cells got parameter values that were found among all or at least the majority of the surrounding basins.

Table 6. *Number of cells that fall into the different regionalisation categories in paper II.*

<table>
<thead>
<tr>
<th>Category</th>
<th>Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>No station found within the window =&gt; default value</td>
<td>6,337</td>
</tr>
<tr>
<td>Only one station (with accepted values) within the window</td>
<td>3,997</td>
</tr>
<tr>
<td>No common values was found between the stations (2 or more stations) =&gt; default value</td>
<td>1,674</td>
</tr>
<tr>
<td>A common value between up to 50% of the stations (4 or more)</td>
<td>2,917</td>
</tr>
<tr>
<td>A common value between more than 50% but less than 100% of the stations (3 or more stations)</td>
<td>10,272</td>
</tr>
<tr>
<td>Common value for all stations (2 or more stations)</td>
<td>7,349</td>
</tr>
<tr>
<td>Value set from own station</td>
<td>24,641</td>
</tr>
<tr>
<td>Value set from own station, not used in regionalisation</td>
<td>1,945</td>
</tr>
<tr>
<td>Total land cells</td>
<td>59,132</td>
</tr>
</tbody>
</table>

Routing

The performance of the aggregated time-delay histogram remained constant over all resolutions while the performance of the linear reservoir routing decreased with decreasing resolution and exhibited a zigzag pattern, with dramatically changing performances between small resolution differences (*Figure 3*). The combined benchmark simulation is somewhat better than the linear reservoir method but it has the same variability and performance decrease with decreasing resolution.

The optimum velocity was constant the over the different scales with the aggregated time-delay histogram, while it varied somewhat with the linear reservoir method. The best WASMOD parameter values were the same with all spatial resolutions with the aggregated time-delay-histogram method while a wide range of parameter values were selected as best for different resolutions with the linear routing algorithm.

The difference between the simulations to distributed and uniform climate forcings were minor compared to the large differences between the different routing models.
Figure 3. Best simulated runoff performance at Boluo in the Dongjiang basin with the NC criterion. TDH = aggregated time-delay histogram, LRR = linear reservoir routing, bench = benchmark run with time-delay calculated from the low resolution network. Both distributed (straight lines) and uniform (dotted) climate data are used for all three simulations. (Paper IV)

Validation
Within-year runoff patterns
The tuning against the long-term VE in paper II was compared with the within-year runoff dynamics. These were reasonably well captured for most of the 663 stations. There was a general tendency that simulated runoff varied more than the measured (Figure 4), while there were a few examples of the opposite. There was also a tendency for around 40% of the gauging stations that the simulated peak flows occurred earlier than the measured (Figure 4), while a delayed peak only occurred at 15% of the stations.
Figure 4. Long-term within-year variations in runoff and precipitation (grey bars) from 15 representative gauging stations at 6 continents averaged over different time periods (in units of 1000 m³ s⁻¹). Measured monthly runoff (white dots) and simulated with WASMOD-M (black dots) based on a parameter-value set tuned from 485 globally-distributed gauging stations. Note different ordinate scales. (paper II)

Global and continental water balances

Globally, nearly half of the 1680 parameter-value sets in paper II produced runoff within ±20% of the measured values integrated over all gauged areas. However, none of these globally-uniform sets could produce acceptable simulations of runoff from all continents. All acceptable globally-uniform parameter-value sets produced overestimated runoff from Europe and all sets caused state-variable problems at some point.

Continental and global runoff values with the distributed parameter-value set of WASMOD-M were in the same range as measurement compilations and comparable investigations (Table 1). Runoff was simulated better for individual continents with continentally-uniform parameter values, than with the globally distributed parameter-value set, but at the cost of compensating errors at the basin scale.

Split-sample validation

The access to time-series data in paper III provided the possibility of split-sample test. The best parameter-value sets identified during the calibration
period produced the among best validation period results for all criteria (Figure 5), although the absolute criteria value could differ.

![Figure 5](image)

**Figure 5.** Modeled runoff performance for validation (second half) versus calibration (first half of observation) periods for three evaluation criteria (Nash, NC; limit of acceptability, LA, and volume error, VE). Top row shows performance for Sénégal River at Bakel (observations 1904 –1989) and bottom row for Ob River at Salekhard (observations 1930 – 1999). Simulations based on the best 150 (1%) calibrated parameter-value sets are highlighted and the thin line gives the one-to-one relation. The best calibrated parameter-value sets for Ob River not fulfilling the 95% limit for LA_{75} during the calibration period are marked with crosses. (paper III)

Comparison with other models

The capacity to mimic observed runoff vary largely between present-day global water-balance models (Figure 6). The varying number of models per basin reflects the lack of easily available simulation data and model inter-comparisons. The non-calibrated models generally perform worse than the calibrated but also the latter can have large volume errors.
Figure 6. Volume errors (%) for WASMOD-M and 5 other global water-balance models at 13 major basins sorted after increasing specific runoff at their down-most discharge stations, from low (Sénégal) to high (Amazon). Published results (left-hand) of calibrated models are marked with filled symbols and non-calibrated models with open symbols. WASMOD-M results from paper III (right-hand) relate to runoff ranges simulated with the 150 (1%) best parameter-value sets based on volume error (VE) and the Nash criterion (NC). Reported runoff can come from different stations in the system, but with small area differences. Models are VIC with and without calibration (Nijssen, et al., 2001a), “K&PH” (Kleinen and Petschel-Held, 2007) and WASMOD-M “WN07“ (paper II). Additionally, with error percentage calculations from reported measured and simulated flow; WBM (Fekete et al. 1999b), WGHM before and after additional correction factors (Döll et al., 2003) and MacPDM with lowest absolute volume errors when several values are reported (Arnell, 1999; 2003; 2005) and with approximate values from figures for the European basins (Arnell, 1999).
Discussion

A major difference between global and catchment water balance modelling is that when studying a specific catchment only, there is normally a possibility to gather catchment-specific information that can be utilised in the modelling. Global-scale water balance modellers are referred to and dependent on global datasets and detailed information about each catchment is not available. The analysis must focus on the general picture and unexpected results for single basins cannot be explored in detail.

Data problems

Several types of data problems face the global water balance modeller. These errors were grouped into three classes in paper II: (i) geographical boundaries and positions, (ii) runoff- and climate-data quality, differing time periods and (iii) unknown or unavailable anthropogenic runoff influences.

Geographical boundaries and positions

Basin areas for given gauging stations can sometimes differ considerably (Renssen and Knoop, 2000), which causes problems in validation of digital river networks and results in different runoff estimates in model comparisons. One reason for the difference in arid areas is that digital flow networks normally calculate the whole upstream area, while station-based estimates can report values on the actively discharging area only (Fekete et al., 1999a). The area differences can introduce large differences in calculated runoff volumes (Arnell, 2005; Döll et al., 2003). Comparison of different runoff data-bases such as STN-30p/GRDC (Fekete et al., 1999b) and GHCDN (Dettinger and Diaz, 2000) reveal that reported gauging stations positions can differ as well.

Inconsistent runoff and climate data and differing time periods

The measured average runoff is higher than the precipitation in some basins (paper II; III; Fekete et al., 1999a; Nijsen et al., 2001a). Two main reasons are the sparse precipitation gauge network in mountainous areas where precipitation amounts are higher compared to lowlands, and the gauge under-
catch of precipitation (Adam and Lettenmaier, 2003; Adam et al., 2006; Legates and Willmott, 1990). Despite the runoff coefficient problems the precipitation correction is too high in some areas. The correction factors by Legates and Willmott (1990) used in this study are known to be too high for Europe (Arnell, 1999c; Döll et al., 2003). They also cause a double-correction in Russia since the stations that form the reference climatology for the CRU time series already are corrected for gauge undercatch (New et al., 1999). Some small Russian basins show too high runoff coefficients despite the double corrections. Another possible reason of too high runoff coefficients is glacier melt. The melting Alaskan glaciers studied by (Arendt et al., 2002) are however not situated in the same Alaskan basins where paper II and III present too high runoff coefficients.

The globally-uniform parameter $E_c$ in the potential evaporation calculation caused too high values in arid areas and too low in arctic areas in paper II. The usage of published potential evaporation datasets in paper III improved the calculations, but very high dryness indices (potential evaporation divided by precipitation) indicated that the potential evaporation was still too high in some arid areas, possibly in combination with too low precipitation. The simplest work-around would be to use the minimum instead of the maximum from the two reference datasets while a more detailed approach would be to use additional data, such as radiation, and utilise the knowledge about the effect of the different potential evaporation calculations (Vörösmarty et al., 1998; Arnell, 1999c; Kaspar, 2004).

Average runoff and precipitation are not constant in time and data from different time-periods can cause differences and complicate the calculation of inter-station runoff (paper II). Mismatch between measured upstream and downstream discharge are also caused by measurement errors (Fekete et al., 1999a), water withdrawal and river-bed seepage, and evaporation from open water bodies in warm climate (Döll et al., 2003). All global runoff estimates does not include information about the time period for the calculation (Table 1).

Unknown and unavailable anthropogenic runoff influences.

Most hydrologic models mimic natural variations of water-balance components (paper II). Regulation of rivers by dams, re-routing of water to other basins, and water extraction are major anthropogenic influences that considerably disrupt the natural water balance. Vörösmarty et al. (1997) andNilsson et al. (2005) quantify the influence of dams and the degree of regulation of large rivers and find that dams globally impact 83% of the global runoff, and represent a 700% increase in the standing stock of natural river water, with residence times for individual impoundments spanning from one day to several years. We are still far away from a trustworthy global database of time series accounting for anthropogenic influences of the global water cy-
cle. Hence, even a perfect fit between simulated and measured runoff needs to be evaluated critically and should not be overrated.

Parameter behaviour

It was found in paper II that globally uniform parameter-value sets could produce good global runoff estimates, but at the cost of compensating errors at the continental scale, and that continentally uniform parameter-value sets could produce good continental runoff estimates, but at the cost of compensating errors at the basin scale. Similarly, Arnell (2005) found large discrepancies for individual basins simulated with MacroPDM, although the total runoff fitted well. Optimal parameter-value sets can differ between upstream and downstream sub-basins in the same system (paper II; III; Döll, et al., 2003) but the tuning against inter-station runoff is complicated, especially for time-series, but also for average values. Inter-station sub-basin cover 63% of the gauged areas in “UNH/GRDC Composite Runoff Fields” and inter-station tuning is thus an important, but rarely discussed task. Some problems to tackle are mis-matching time-periods of upstream and downstream gauges, non-modelled losses such as evaporation and water withdrawal, and the question if bad simulated upstream runoff should be corrected or not to provide good inflow to the downstream area.

All parameters are assigned uniform values over large areas in WASMOD-M, as opposite to the other global water balance models where some parameters vary with, e.g., the properties of vegetation and soil. These differences complicate cell-by-cell model comparison (paper II) and make model estimations from WASMOD-M less reliable for, e.g., a country. Further distribution of the parameter values is a future model development goal.

The assumption in paper II that WASMOD-M is overparameterised when tuned only against VE was confirmed in paper III. The model evaluation criteria NC and LA focusing on the runoff dynamics confined the evaporation- and runoff-parameter spaces to a much higher degree than VE. The evaporation and slow-flow parameters $A_c$ and $P_s$ could not be simultaneously well confined by any criteria. $A_c$ was more confined by all runoff criteria for non-snow than snow basins, and the opposite was partly true for $P_s$ and $P_f$. These climatic differences between parameters need further attention in future studies, and the low number of parameters might still need reduction.

The evaluation against snow-covered data through SF in paper III showed that the snow parameters $T_s$ and $T_m$ were confined by SF. All equifinality could not be removed and future model versions might use a fixed $T_m$ value. This finding might be challenged with snow-water equivalent data or snow cover data that can be used also at the start and end of the snow season. Evaporation data would also be very useful for model evaluation and tuning.
The number of basins with $VE$ values within 20% simulated by the final parameter value set in paper II increases to 455 from the used 429. This is however lower than the 485 basins which were well modelled with the uniform parameter-value-sets. The decrease is caused by the overwriting by downstream sets and the usage of the upstream measured runoff in the regionalisation algorithm. The latter approach was chosen to avoid compensating errors downstreams, and because usage of the whole range of upstream accepted parameter value sets combined with the downstream sets would give too many combinations.

Time aggregation and criteria relationships

Generally, the equifinality of the runoff and evaporation parameter values increased when the model was tuned successively against monthly, annual and long-term average runoff (paper III). The most common exception was the evaporation parameter $A_c$, that on average was slightly more confined by the annual $NC$ and $LA$ values than by the monthly. Generally the monthly and annual $NC$ and $LA$ criteria confined the parameter values to about the same parameter-value space, although the selection of the 1% best parameter values could differ.

$NC$ and $LA$ values are not comparable between the time aggregations. $NC$ is normally lower for annual time-series because of their lower variability. The majority of the stations had short time-series which made the annual $NC$ values uncertain. Usage of whole time-series, instead of only the first period might partly solve this problem.

Just as $NC$ and $LA$ are not comparable between time aggregations, they are not fully comparable between basins either. $NC$ depends on the observed runoff variability and high values can be achieved also with poor models (Schaefli and Gupta, 2007). The improvement over a benchmark, calculated from, e.g., scaled precipitation is worth exploring.

In many senses $NC$ and $LA$ behaved similarly, but the best $NC$ values fitted much better than the best $LA$ values with the best $VE$ values. The $LA$ criterion needs further investigations before it can be fully utilised on the global scale. It is likely that narrower limits easier would identify good parameter value sets of the well-simulated basins. Further studies are also needed of the not so well modelled stations, both regarding the inclusion of additional processes to the model, and the criteria behaviour.

The larger equifinality of the $VE$ criterion also resulted in a wider range of possible runoff dynamics, while the constriction of the snow parameters though $SF$ only marginally influenced the runoff. Also Dunn and Colohan (1999) and Udnæs et al. (2007) find that internal model consistency can improve with snow data in addition to the runoff data, but that the total runoff is only marginally affected.
The $VE$ criterion was combined with limits for the state-variables in paper II. The usage of criteria focusing on the dynamics minimised these problems in paper III and no state-variable limits were used. The few stations that still experienced problems coincided with low $NC$ values and warm-up times much longer than for other stations. Preparation of a distributed initial state-variable set, in combination with continued usage of criteria focusing on the dynamics, might eliminate the problem.

Validation

The model performance tuned against $VE$ only was compared against within-year dynamics in paper II whereas the access to time-series data enabled validation against an independent time-period in paper III. Comparison with other global runoff estimates and the performance of other models in individual basins were also used for comparison in papers II and III. Comparison against other models and estimates can, strictly speaking, not be called validation in a Popperian sense. When simulations agree with those of earlier models, it does not prove that the simulations are right since all earlier models might be wrong.

Within-year dynamics and split-sample validation

The present version of WASMOD-M does not account for, e.g., time delays along river networks, evaporation losses from dams and rivers, flood inundation, anthropogenic water abstraction, river regulation, river routing, subgrid variability, glacier dynamics, permafrost and capillary rise. Future model development may include additional processes and input data if validation data and consistency checks can substantiate them (paper II). The missing routing (both natural and anthropogenic though dams) is a main reason why WASMOD-M often simulated peak runoff too early (paper II). The within-year dynamics however agreed reasonably well with the measured. Alterations by humans, which are more common during the validation period, as well as climatic variations, affected the split-sample test (paper III). The best parameter of the calibration period values were the best, or among the best also for the validation period. For some basins was the whole set of 15000 parameter value combinations improved for the validation period, and for some basins were all impaired.

Model shortcomings were not always easy to explain, but some reasons were non-modelled lakes, that gave $NC$ values below 0 and too high runoff coefficients that gave too high $VE$ values and extreme values of the evaporation parameter $A_e$ (paper III). Larger basins were normally easier to model since different errors could cancel each other.
Comparison with other models and estimates

A total global runoff of 38,605 km$^3$yr$^{-1}$ was simulated in paper II (Table 1). This was in the same range as previous model- and measurement-based estimates. It was higher than the estimate by Döll et al. (2003), which can be expected as they do not apply correction factors to the CRU precipitation, but lower than the estimate Gerten et al. (2004) who also do not correct the CRU precipitation. There is a tendency that measurement-based estimates are higher than the model-based, although the opposite should be expected. The runoff from Europe approached earlier estimates when uncorrected precipitation was used, and the runoff from South America was well simulated if the state variable were allowed to exceed their limits.

All global water balance models perform “reasonably well”, but every model still has large errors in some areas, and these do not necessary coincide (Figure 6). The development of global water balance models would benefit from model intercomparisons, with, e.g., similar input data and river networks. Gerten et al. (2004) and Hanasaki et al. (2007a) compare the current results of some of the models.

WASMOD-M is a simple global water balance-model, and the simplicity introduces limitations in the possible uses of modelled data. WASMOD-M, together with Macro-PDM and WBM, only simulate natural runoff, without any major human alterations. As humans affect a major part of the global river runoff (Vörösmarty et al., 2004; Oki and Kanae, 2006) this is too simplistic in most river basins. It was shown in paper III, that WASMOD-M, despite being very parsimonious, still suffer from equifinality. Except for the model by (Kleinen and Petschel-Held, 2007) WASMOD-M use the fewest number of parameters, and it can be questioned if the other water balance models are not overparameterised as well. This question might be weakened by the fact that WASMOD-M has the highest number of calibrated model parameters. Parameter values are commonly estimated from physiographic datasets instead of being calibrated (Arnell, 1999c; Döll et al., 2003; Vörösmarty et al., 1998). Theoretically this is better than calibration, but the uncertainties introduced by these datasets (Hannerz and Lotsch, 2006; Peel et al., 2007), must be taken into account.

An interesting question, which has no clear answer, is whether additional correction factors should be applied when a model fails. When a correct upstream input is wanted for a downstream station, corrections are applied by Döll et al. (2003) and Fekete et al. (1999a). Arnell (2003) who only evaluate runoff against the downmost station do not use any correction. The overwriting of the upstream best parameter-value sets by the downstream best, as used in paper II and by Nijssen (2001a), can also be seen as a correction.
Routing

The new aggregated time-delay-histogram routing is not dependent on the flow network scale and can be used for further studies of the impact of precipitation heterogeneities at different resolutions. The widely used linear reservoir routing can however introduce much larger errors than the uncertainties caused by low-resolution climate data. The variability of the linear river routing and the benchmark simulation is most likely caused by the large changes of the flow networks between the different resolutions and the relocation of the gauging station (paper IV). Decreasing model performance with decreasing resolution is also noted by (Guo et al., 2004), but while the performance of aggregated time-delay-histogram remained constant with decreasing resolution their proposed multiple-outlet approach of the low-resolution network also decreased in performance, although as not as much as the standard single-cell outlet network. The zigzag pattern can however not be seen in their result. One possible reason is that they test much fewer resolutions.

The variable identification of the best WASMOD-M parameter values in the calibration procedure with the linear reservoir routing implies that non-optimum runoff parameters are selected to compensate for errors in the routing procedure.

The work in paper II and III highlight the need for routing in WASMOD-M. The aggregated time-delay-histogram method has potential to be applied globally. One advantage is that lakes are implicitly represented with their almost zero slope. One drawback is the calibration, but as the velocity parameter was constant over the scales in the Dongjiang catchment it can be speculated that it will remain constant also in other basins. The preparation of the time-delay histograms will be computationally demanding, but once this data is prepared, the routing itself much quicker than cell-to-cell routing.

It should be stressed that the time-delay-histogram method can be applied for any destination cell, not just the basin outlet. Maps of accumulated discharge, as provided from cell-to-cell routing, can be produced by applying the method with a destination cell in each low-resolution cell. The high-resolution destination cell should be the cell with the largest upstream area within the low-resolution cell. The time-delay-histogram method can be methodologically classified as between source-to-sink and cell-to-cell routing.
Model calculations of global runoff differ a lot between different models. The errors of simulated river runoff can be large and do not always point in the same direction. Global water balance models perform “reasonably well”, and do not show as good performances as can be found when investigating a single well-monitored basin. Data access and quality is a main limitation. Precipitation and river regulation are major sources of uncertainty.

It was shown that the new simple global water balance model, WASMOD-M, could simulate the global runoff “reasonably well”. A simple regionalisation procedure that utilised proximity only contributed to calculate a global runoff estimate in line with earlier estimations. Already tuning against the volume error produced reasonable within-year patterns produced in most basins, but the volume error was not enough to confine the parameter space. The parameter space and the simulated hydrograph could be better confined by, e.g., the Nash criterion. Calibration against snow-cover data could confine the snow parameters better, although some equifinality still persisted. Thus, even the simple WASMOD-M showed sign of being overparameterised. At the same time it was shown that global uniform parameter value sets produce reasonable global runoff, but at the cost of compensating errors at the continental scale.

A new aggregated time-delay-histogram routing method, that utilised high-resolution flow network information in low-resolution calculations performed equally well over all scales in the Dongjiang catchment, while the standard linear reservoir routing decreased in performance with decreasing resolution. This algorithm was developed for later improvement of WASMOD-M. The somewhat ambiguous runoff evaluations against monthly and annual observations will be easier to analyse when the routing is implemented.

The development of global water balance models would benefit from better precipitation data, both in terms of coverage and gauge corrections. Larger spatial and temporal coverage and less error in runoff data would also be very useful, together with estimation of other properties, such as evaporation and snow, for further model consistency checks. A global data-base of dam-operation schemes is probably not possible, because of company secrets, but would be very useful for the research community. Model intercomparisons and model ensembles that use the same standardised inputs, would also con-
tribute. Estimates of global runoff should be followed by information about time period and included flows.

Possibilities for future work

Further work with WASMOD -M could continue along three paths, testing, development, and application. These are no separate paths and numbers eventually presented from the application path should be given together with uncertainty estimates. Model testing will point out important parts of the model to improve, and new model versions should be thoroughly tested as well.

Testing could be continued in different directions. One example is climatic split-sample tests to improve trust in model applicability for climate-impact studies. When improved regionalisation methods have been incorporated they should be tested with cross-validation techniques. The simulations results from paper III could be further analysed in a proxy-basin context. A forthcoming model version with the new routing algorithm could be expected to behave differently when evaluated against monthly versus annual observations so the time-aggregation tests should be repeated. Since the study of model efficiency is not an absolute science, development of benchmarks along the suggestions of Schaefl and Gupta (2007) should be tested, not only against WASMOD-M simulations but also against simple statistical models to verify that model performance is not just the effect of pure luck.

The input data used has well-known shortcomings and the model simulations will improve with better data, especially for precipitation. Usage of several driving datasets will provide another uncertainty assessment. Additional data on, e.g., evaporation and snow-water equivalent will provide both testing and further calibration possibilities.

Global water balance models, as being applied everywhere, might be useful for exploration of advantages and disadvantages of more general models (Andréassian et al., 2006) compared to models with adoptions to the conditions and data availability at each sites (Beven, 2007). The learning from the model applications is central to both approaches.

Model development could first of all be continued with regionalisation and routing. Regionalisation could encompass methods based on independent physiographic datasets to get reasonably distributed parameter-value datasets both within and between basins. Distributed parameter-value sets are expected to produce better inter-station runoff. The model algorithms could be reformulated to avoid calibrated parameters. The evaporation formulation could be revisited on the basis of physically reasonable limits in various parts of the world. Routing could be included that account for forthcoming databases of dams (Reidy Liermann, 2007). Rivers like Colorado, Huang He, Nile, and Amu Darya and Syr Darya (Gleick, 2003) will not be
well simulated unless water losses withdrawal and open water evaporation are modelled. Long-term effects of melting glaciers in the world, especially in Asia, could be studied through inclusion of these processes.

Investigations of many large-scale water-resource problems such as water conflicts between nations, trans-boundary water management, virtual water trade, population growth in relation to water scarcity, changed land use and climate change are possible applications after model improvements.
Acknowledgements

Welcome to the most read part of every thesis! It is an important part since there are so many that have their share in this work. I wouldn’t have made it without all of you!

First of all, my supervisors Sven Halldin and Chong-yu Xu: I’ve been fortunate to have very kind supervisors that also have been available almost all of the time. Your doors were always open and quick e-mail answers came from Chong-yu in Oslo as well as in Hongkong. Sven, every text suggestion that passed you improved and I appreciate that help a lot. Remaining errors in this thesis are my own. You also managed to find the general picture among all the details that I provided to you with. But when it comes to administrative peculiarities we were commonly interested! Thanks also to Bitte for selecting your vacation in such a way that Sven was available to check my texts almost anytime during the last week!

I’m happy to have had Lars-Christer Lundin as my third supervisor. As you did not participate in the daily work you could see everything with some distance and always had good suggestions at the formal meetings where we gathered.

Lebing Gong, already the day I got to know that I would get you as fellow global modeller I started to think about what to write about you here as I knew that your addition would speed-up and improve my work a lot – and it certainly has. You’re smart, quick and ambitious. You are also very kind – I do not know how many times you told me not to worry this last autumn. As the first author of the routing paper you also got very much pressure on you this last month because of my thesis deadline, together with the other tasks you had to fulfil.

Keith Beven, I appreciate that you have taken your time in reading and correcting my papers. Although we didn’t meet so much during your year here in Uppsala I’ve noticed how you really have inspired my modelling colleagues.

Visualisation is very important for understanding model data. Andreas G and Dan helped me with GIS program tricks.

The global modeller is dependent on datasets gathered and made available. I’m very grateful to all those individuals and groups that freely provide their data. The data used in this thesis come from: Tim Mitchell and David Viner at Climate Research Unit (CRU) at University of East Anglia, Willmott, Matsuura and collaborators at University of Delaware, GRID-Tsukuba
at National Institute for Environmental Studies and Chung-Hyun Ahn and Ryutaro Tateishi at Chiba University, Water Systems Research Group at University of New Hampshire, Thomas de Couet at Global Runoff Data Centre (GRDC), Michael Dettinger and Henry Diaz though Byron Bodo and the Data Support Section at the National Center for Atmospheric Research, International Satellite Land-Surface Climatology Project (ISLSCP) and the National Snow and Ice Data Center (NSIDC) at University of Colorado, Boulder.

My calculations have sometimes been computer demanding and several LUVA co-workers have let me use their computers. Some of the computations in paper III and IV were performed on UPPMAX resources under Project p2006015.

The major part of the funding for this project came from Swedish Research Council on contracts 629-2002-287 and 621-2002-4352. Additional funding came from the Department of Earth Sciences, the National Graduate School in Scientific Computing and Liljewalchs traveling scholarship. The 3 start-up months after a parental leave decided by our faculty is well worth exporting to other universities.

When I started in the hydrology group, it was me, Auli, and “It’s raining men, hallelujah!” for the rest, and the lunches even more so (Fredrik – the vacation in Paris, Fritjof, Johan Ö, Klas, Magnus, Niclas, Tony, Andreas G and Micke). I miss you guys! Tomas Nord has been a reliable lunch companion over the years. You, later with the aid of Johan T, also keep our computer system running and free from viruses, worms, trojans etc. so we do not have to worry. It is also always nice to stop at your open door for a chat.

In this world of men I noticed after a while that I had to giggle sometimes and that was best made with the meteorologists (Anna S, Cicci, Karin T – the company, Xiaoli – the game evenings). Hanna Z, I’m happy that you joined the hydrology group (and started your own process of smoothing the educational background among the Ph.D. students). Hanna was later followed by Zairis and Ida (the limit of acceptability inspiration) so if Julia (belonging to all old divisions!) is counted in we are five women and four men (José Luis, Lebing, Martin, Michael) as long as I’m not a doctor!

From Erik K to Elias, the list of nice young researchers that has passed the “LUVA(L) Challenge” group is long: Andreas B, Angelica, Anna L, Aziz, Dan, Håkan, Jenny, Karin J, Sara-Šofia, Toddy (die Gespräche), Anna R, Andreas K (the evening talks), Babatunde, Björn C, Erik S (the practical thesis information), Igor, Kim, Maria, Mathias, Wasana, Abdulhadi, Anne-Marie, Björn S, Hanna R, Jim, Karin A, Marcus, Flora, Fredrik H, Louise, Markus, Stefan. From other parts of the department I’m happy for the cooperation and friendship with Camilla P, Chris A and Mattias L. Thanks also to all colleagues that have checked that I’m doing all right during these last stressful months (Allan, Cicci, Hélène, ...). Our very quick printing office also has an important share in that it was possible for me to finish this work.
I have learnt a lot from teaching (including the complexity of reverse parking in snow) in the co-operation with other teachers (Allan, Rajinder, Roger, ...) and contacts with students. The UVC years were also fun. It has been inspiring to meet other researchers at conferences and courses (the birthday and football in Umeå).

The soil science project during my undergraduate studies, under the guidance of Petra Bergkvist and Nick Jarvis at SLU was very important for my wish to do a Ph.D. I’m also very happy for the year I spent with Per-Erik Jansson and co-workers at KTH, in cooperation with the Uppsala limnologists, where I got both practical measuring experience and some SVAT knowledge. Further back, the advanced environmental science course by Jack-Sture Börjesson and Sten Willstrand at Nils Ericsonsgymnasiet stands out in my memory – not only because of the fika! The projects during that course I carried out with Kajsa, which leads me over to all my friends among the “theatre people”, whom I met both at lunches, courses and fully outside of work. This text is becoming lengthy but at least I want to mention Sara J, Kneppers, KTF and Uppsala baptistförsamling too, and my (lunches with Tove) and Daniels relatives.

My parents and parents-in-law have a large share in this work. They have travelled several hours during many weekends of this autumn to come and help us at home such that I could focus on working. Thanks for all the love and support! Gustav and Anders, it is quite fun that we end up doing programming all of us, isn’t it? Maybe it is something in the genes from Rosa, Arne, Britt-Marie and Conne?

Daniel, my love, you’re the best! You answered computer questions, discussed practical, theoretical and philosophical research question with me and took care of Ture day and night while I was working. Without you and your hugs this had not been possible. (For those of you that share my interest in details: I’m married Widén Nilsson but I publish as Widén-Nilsson.)

Ture – dina leenden är de finaste jag någonsin sett. Nu skall vi leta efter roliga lekplatser i Bern tillsammans!
”Vatten som rör sig i en flod antingen lockas, jagas eller rör sig av sig självt. Om det lockas, eller begäres, vem är den begärande? Om det jagas, vem är den som jagar det? Om det rör sig av sig självt visar det prov på förnuft, vilket är omöjligt hos kroppar som ständig förändrar form, ty i sådana kroppar finns inte något förstånd.”


“Bara en vattenpuss
ville jag vara
och spegla himlen”


Det finns en uppsjö av utvärderingsmått som anger hur bra en viss simuleringsmodell är jämfört med mätdata. Utvärderingsmåttan används både för att analysera modellens beteende och för att välja ut de parametervärden som ger bäst resultat (kalibrering).

Fem av de sex parametrarna i WASMOD-M behöver kalibreras, medan den sjätte bestäms i förväg. Parametrarna är begreppsmässiga och kan därmed inte mätas utan måste kalibreras. Två parametervärden används för snöberäkningen, en för avdunstningen och två för avrinningen. Den sjätte,
icke-kalibrerade parametern, används före själva modellberäkningarna för att räkna ut den potentiella avdunstningen, som förutom nederbörd och temperatur driver modellen.


Flödesdata fanns att tillgå från ungefär 50% av jordens landområden, eller 72% om man endast räknar de områden som producerar avrinning (Fekete m.fl., 2002). Kalibrerade parametrar behöver alltså överföras till de områden där kalibrering inte kan göras. Detta kallas regionalisering. En mycket enkel kalibreringsmetod utvecklades, som söker efter parametervärdeskombinationer i närliggande områden till en ruta som saknar parametervärden.

De viktigaste slutsatserna i avhandlingen är:

- Modellberäkningar av flöden i floder varierar kraftigt mellan modeller. Felen i de olika modellerna är stora och inte alltid entydiga. Även WASMOD-M, som är en mycket enkel modell, visade indikationer på att vara överparameteriserad och särskilt snöberäkningen skulle kunna förenklas. Modellutvecklingen skulle främjas av en studie som jämför de olika modellerna, när de använder samma data. Osäkerhet i data är en stor bidragande orsak till modellosäkerheten. Osäkerheterna finns i alla utnyttjade data, men felen i uppmätt nederbörd är viktigast eftersom den främst av allt styr den simulerade avrinningen.

- Volymfelet är det vanligaste måttet på en modells förmåga att simulera flödet förbi en viss mät punkt. När WASMOD-M kalibreras endast med hjälp av volymfelet gick det att även få en ganska bra inomårsvariation, men parametervärdena kunde inte bestämmas entydigt. Med ett mått som
även tar hänsyn till inomårsvariationen blir både parametervärdena och den möjliga simulerade avrinningen begränsade till ett snävare intervall.

- Genom att använda sig av andra data än vattenföring, i det här fallet snö, kan antalet möjliga parametervärdeskombinationer begränsas. De förbättrade parametervärdena behöver dock inte ha så stor effekt på den simulerade avrinningen. Tillgång till globala avdunstningsdata skulle vara mycket användbart för vidare studier.


Det tar tid att bygga upp vattenbalansmodeller och det finns många möjliga vidareutvecklingar av modellen. Det är också viktigt att fortsätta att kritiskt testa modellen på fler sätt än vad som hittills gjorts. När modellen börjar tillämpas för att t.ex. studera gränsöverskridande vattenresurser, virtuell vattenhandel, befolkningsstillväxt och vattenresurser samt effekterna av förändrad markanvändning och förändrat klimat är det viktigt att även informera om osäkerheterna i beräkningarna.
References


GRDC. 2007, The Global Runoff Data Centre, 56002 Koblenz, Germany


Mearns, L.O., F. Giorgi, L. McDaniel, and C. Shields. 1995. Analysis of Variability and Diurnal Range of Daily Temperature in a Nested Regional Climate Model -
Comparison with Observations and Doubled Co2 Results. Climate Dynamics 11:193-209.


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”.)