Caspian Environment Programme
Phase 2

Grid-based Model of the Caspian Sea Basin

Phase 2 Report

December 2001

This report has been prepared for the Caspian countries with funding from the European Union's Tacis Programme. The contents of the report are the result of a cooperation between the EU's experts and experts of the five Caspian countries.

The members of the consortium contracted to manage this work are listed below. Liabilities for the opinions expressed are limited by the contract between the Consortium and the European Commission. Any enquiries about the contents of the report should be addressed either to the relevant Caspian Regional Thematic Centre or a member of the Consortium:

ERM-Lahmeyer International GmbH
Konrad-Adenauer-Str 3
D – 63263 Neu-Isenburg (Frankfurt)
Germany
Tel. (+49-6102) 206-0
Fax (+49-6102) 206-202

DHI Water & Environment
Agern Allé 11
DK-2970 Horsholm
Denmark
Tel: (+45-45) 16 92 00
Fax: (+45-45) 16 92 92

GOPA Consultants
Hindenburgring 18,
D-61348 Bad Homburg
Germany
Tel: (+49-6172) 93 0-0
Fax: (+49-6172) 93 03 50

Caspian Environment Programme

Grid-based Model of the Caspian Sea Basin

Phase II Report

Emma Tate and Jeremy Meigh

Centre for Ecology & Hydrology

Wallingford

Oxfordshire OX10 8BB

UK

with the

State Hydrological Institute

St Petersburg

Russia

December 2001

Centre for Ecology & Hydrology

NATURAL ENVIRONMENT RESEARCH COUNCIL

Contents

Chapter 1. Introduction **1**

Chapter 2. Work carried out **3**

2.1. Verification of input data 3

2.2. Model development 3

2.2.1. Reservoir module 4

2.2.2. Glacier module 9

2.2.3. Water use module 13

2.3. Model calibration 14

2.3.1. Volga basin 16

2.3.2. Ural basin 17

2.3.3. South-West basin 17

2.4. Development of future scenarios 20

2.4.1. Climate change scenarios 20

2.4.2. Demand scenarios **Error! Bookmark not defined.**

Chapter 3. Results **28**

3.1. Model calibration 28

3.1.1. Volga basin 28

3.1.2. Ural basin 34

3.1.3. South-West basin 37

3.1.4. Summary of calibration results 39

3.2. Scenarios of future Caspian Sea inflows 42

Chapter 4. Conclusions and recommendations **46**

Chapter 5. References **48**

Annex A: Modelling Work in Phase I and Intermediate Phase **49**

A.1. Work carried out in Phase I 50

A.1.1. Model development 50

A.2. Work carried out in the Intermediate Phase 73

A.2.1. Verification of input data 73

A.2.2. Model development 73

A.2.3. Calibration procedure 78

A.3. References 85

Annex B: GWAVA Global Water Availability Assessment Model Version 3.3 User Notes 87

B.1. GWAVA Version 3.3 User Notes: General 88

B.1.1. Background 88

B.1.2. Objective of these User Notes 88

B.2. Data preparation 91

B.2.1. Parameters file [example: test_par_mountains.xls/.txt] 91

B.2.2. Sub-catchment calibration parameters file (optional) [example: subchge.txt] 97

B.2.3. Large reservoir operation data file (optional) [example: resinfo.txt] 98

B.2.4. Large reservoir area:storage data file (optional) [example: areastor.txt] 98

B.2.5. Climate baseline data file [example: testrnev.xls/.txt] 99

B.2.6. Precipitation and temperature anomaly file [example: test0505.txt] 100

B.2.7. General water demands file [example: testdems.xls/.txt] 101

B.2.8. Monthly water demands file (optional) [example: testmondems.xls/.txt] 106

B.2.9. Transfers between cells file (optional) [example: testtran.txt] 107

B.2.10. Climate change scenario file (optional) [example: test2050.xls/.txt] 108

B.2.11. Groundwater availability file (optional) [example: test_gw.xls/.txt] 109

B.2.12. Mountainous region data file (optional) [example: mount.dat] 109

B.3. Running the model 111

B.4. Model output 116

B.4.1. Output of time series of monthly flows 116

B.4.2. Output of mean monthly flows and demands 116

B.4.3. Output of annual data and water availability indices 117

B.5. Calibrating the model 119

B.5.1. Selection of suitable gauged flow series 119

B.5.2. Suggested calibration procedure 119

B.6. References 120

List of figures

Figure 1 Schematic of a normal reservoir operating curve 5

Figure 2 Flow diagram of the algorithm used to model large reservoirs 7

Figure 3 Modelled and observed inflows and outflows from (a) Kujbshevskoye and (b) Volgogradskoye reservoir. 8

Figure 4 Sensitivity tests for climate gradients with elevation on a glacier cell (43°N; 42°E) with one year of input data 12

Figure 5 Net abstraction (where variation is known) in the South-West part of the Caspian Sea basin during the baseline period 14

Figure 6 Map of Caspian Sea basin showing some calibration gauge sites 15

Figure 7 Map of South-West basin showing grid cells, main flow directions, and locations of main river basins and gauging stations 18

Figure 8 Comparison of baseline and 2050 monthly precipitation data for a) HadCM3, b) CGCM1 and c) Echam4 23

Figure 9 Comparison of baseline and 2050 monthly temperature data for a) HadCM3, b) CGCM1 and c) Echam4 23

Figure 10 Comparison of baseline and 2050 monthly potential evaporation data for a) HadCM3, b) CGCM1 and c) Echam4 24

Figure 11 Volga basin model calibration results 30

Figure 11 Volga basin model calibration results 32

Figure 12 Results of time-series modelling of the river Volga at Volgograd for the baseline period 33

Figure 13 Ural basin model calibration results 35

Figure 14 Results of time-series modelling of the river Ural inflows to the Caspian Sea, for the baseline period 36

Figure 15 South-West basin model calibration results 39

Figure 16 Results of time-series modelling of the river Kura inflows to the Caspian Sea, for the baseline period 41

Figure 17 Comparison of observed and scenario inflows 43

Figure 18 Mean monthly flows for 2050 scenarios – Volga basin at Volgograd 45

Figure 19 Mean monthly flows for 2050 scenarios –Ural basin inflows to Sea 45

Figure 20 Mean monthly flows for 2050 scenarios – South-West basin inflows to Sea 45

Figure A21: A map of the Caspian Sea region 51

Figure A23: Map of the precipitation factors used in the Volga and Ural sub-basins of the Caspian Sea basin for a) January, b) February and c) March 58

Figure A25: Map of the precipitation factors used in the Kavkaz region of the Caspian Sea basin for a) January, b) February and c) March 59

Figure A27: Caspian Sea basin cell linkages 66

Figure A29: Caspian Sea basin mean annual precipitation (mm) for baseline period 1961-90 67

Figure A31: Caspian Sea basin monthly mean temperature (°C) for baseline period 1961-90 68

Figure A33: Caspian Sea basin mean annual potential evaporation (mm) for baseline period 1961-90 68

Figure A35: Caspian Sea basin FAO soil texture classes 70

Figure A9 *Corrected Caspian Sea basin cell linkages* 76

Figure A10 *A map of the Caspian Sea basin showing gauges and sub-basins used for calibration* 79

Figure A11 *Results of the sensitivity analysis for two cells in the Caspian Sea basin for grid cell 21701445 (Volga basin) and 22601315 (Kura basin)* 84

List of tables

Table 1 Large reservoirs of the Caspian Sea basin 16

Table 2 Summary of observed flow series used for model calibration 19

Table 3 Summary of annual values and changes in meteorological variables for 2050 averaged over the Caspian Sea basin 25

Table 4 Coefficients used to estimate future demand scenarios from baseline (1961-90) demand data ..... **Error! Bookmark not defined.**

Table 5 Net abstractions of the Caspian Sea basin (Mm^3/year) 27

Table 6 Results of calibrating mean annual total volumes at gauges in the Volga basin 28

Table 7 Results of calibrating mean annual total volumes at gauges in the Ural basin 34

Table 8 Results of calibrating mean annual total volumes at gauges in the South-West basin 37

Table 9 Summary of inflows to the Caspian Sea 40

Table 10 Scenarios of future mean annual Caspian Sea inflows (Mm3/year) 42

Table 11 Percentage change in future mean annual Caspian Sea inflows for various scenarios (compared to observed value in calibration period, 1961-90) 44

Chapter 1. Introduction

This report is an output from the grid-based water resources modelling of the Caspian Sea basin carried out by the Centre for Ecology & Hydrology (UK) in association with the State Hydrological Institute (Russia). The grid-based modelling is an activity of the Caspian Centre for Water Level Fluctuations (CCWLF), based in Almaty, Kazakhstan. The work of the CCWLF is, in turn, a component of the overall Caspian Environment Programme funded by the European Union (EU-TACIS).

The general objectives of the project were:

- To develop and calibrate a grid-based model of the basin which is capable of reproducing observed river flows into the Caspian Sea with reasonable accuracy;
- To provide an assessment of the possible effects of scenarios of climate and water use change on the river flows into the Caspian Sea; and
- To enhance co-operation in the basin by developing a model which can be used by all the countries surrounding the Caspian Sea (Azerbaijan, Iran, Kazakhstan, Russia and Turkmenistan), and by providing training and installing the model in each country.

The grid-modelling work reported on here does not provide a complete examination of the potential future levels of the Caspian Sea since it only concentrates on one of the factors that affects sea levels, that is, the river flows into the sea. In order to examine the sea levels themselves, additional work is needed, including modelling the water balance of the sea itself and modelling surges in sea level. These activities have been carried out by other participating organisations as separate exercises, and they are reported on in other reports of the CCWLF.

The grid-modelling work has been carried out in a number of separate phases, which were as follows:

- In Phase I a realistic and consistent methodology for the work was devised. The approach was based on the GWAVA (Global Water Availability Assessment) model which had previously been developed by the Centre for Ecology & Hydrology as a global-scale approach to examining problems of water availability and demands on a gridded basis (Meigh *et al.*, 1998, 1999). To provide a preliminary grid model for application in the Caspian Sea basin, the GWAVA model was configured for the basin and the majority of the data needed to run it were assembled. Additions were made to the model so that it could compute the effects of snow accumulation and snow-melt on runoff, and so that it could take account of bulk transfers of water into or out of the basin. Preliminary calibrations were carried out, and some initial results were presented for present conditions as well as for climate change scenarios for the 2050s. This work was reported in Tate and Meigh (1999).
- The Intermediate Phase, between the end of Phase I and the start of Phase II, was intended to achieve some continuity of the project. It included refinement of model

input data, such as precipitation and evaporation values, and the collation of observed flows at key sites to improve model calibration. It was carried out in St Petersburg, Russia, from 10 to 21 April 2000 (Tate, 2000). A summary of the model development work carried out in both Phase I and the Intermediate Phase is included in Annex A; this also provides a description of the important features and main components of the modelling approach used.

- Phase II of the project ran until 6 December 2001. The objectives of this phase had been identified in earlier phases in order to complete the model development and provide results for use by other activities of the CCWLF. The principal activities were:
    - Verification of some of the input data, particularly with regard to the baseline climate data.
    - Further development of the model to improve the representation of water use, glaciers and large reservoirs within the basin, since they can have significant influence on the hydrographs of inflows to the Caspian Sea.
    - Calibration of the model for all significant inflows to the sea.
    - Development of climate change and water use scenarios, and their application to provide possible inflow sequences to the Caspian Sea. The flow sequences can be used in combination with sea water balance and surge models (as discussed above) to examine possible future sea levels.

These activities are described in more detail and the results are presented below. The work in this phase also included training participants from each of the five Caspian countries in the background methodologies and use of the model, and the installation of the model in each country is also planned.

Chapter 2. Work carried out

Following the work of Phase I and the Intermediate Phase, several key requirements for this phase were identified. Some modification of modules in the main program was seen to be necessary, particularly through:

- a more sophisticated reservoir operation module;
- a module to model flows derived from melting glaciers;
- more detailed treatment of water demand, to take account of monthly water use patterns, and changes in water use over the period of a run.

Verification of the input data was required (particularly for the climate data in the Caucasus mountains), and some additional, more detailed, input data were required from the basin countries to support the modelling of these new or refined components. Re-calibration of the newly refined model was carried out, then climate change scenarios and demand scenarios were selected to model the effects of the potential future climate and demand on inflows to the Caspian Sea.

The resulting grid-based model for the Caspian Sea basin is GWAVA Version 3.3. Detailed documentation on the requirements for setting up and running the model is given in Annex B.

2.1. Verification of input data

Following preliminary calibration runs carried out in Phase I and the Intermediate Phase, the modelled flow volumes in the Caucasus mountains region did not compare well with measured values. Further detailed verification of the input baseline climate data (precipitation, temperature and potential evaporation) for the region of the Caucasus mountains was carried out for the 1961 to 1990 baseline period. This was carried out by comparing monthly values at several sites with station measurements and maps of measured data.

On a half degree resolution, the baseline climate data were found to be generally of good quality. The Climatic Research Unit (CRU) baseline data were derived with relation to the mean elevation in each 0.5° grid cell (New *et al*., 1999), and this tends to mask the elevation effects on climate data in steeply mountainous regions. The solution to modelling these elevation effects was derived through the development of a glacier module, which implicitly required adjustments to the baseline climate data in mountainous regions.

2.2. Model development

Several important changes were made to the GWAVA model, in order to improve its performance in modelling the Caspian Sea basin. These were: development of an improved reservoir module to simulate important reservoirs; development of a glacier module to

simulate the effects of glaciers on runoff in upland parts of the catchment; and development of a more detailed model of water demand. These are described in detail below.

2.2.1. Reservoir module

In earlier versions of the model, reservoirs were simulated using a very simple water balance and a simple equation linking inflow to outflow. Recognising the important effects large reservoirs have on the flow regimes in the Caspian Sea basin, particularly along the Volga and Kura rivers, an improved method of simulating these large reservoirs was required, whilst still maintaining the global scale of the GWAVA model.

Typically, large reservoirs are controlled by patterns of monthly level or storage which the reservoir manager will aim for each month (see Figure 1). The reservoir manager must also take account of inflows to the reservoir, direct abstractions from the reservoir (*e.g*. for water supply or irrigation), and guaranteed releases from the reservoir. The relative levels of priority of these components will decide which should be curtailed in the event of insufficient water being available. In the Volga basin, guaranteed releases from the reservoirs have different levels of importance at different times of year:

- Spring: fisheries for caviar production in the Volga river delta;
- Autumn: navigation on the Volga river;
- Winter: hydroelectric power production.

The cascade of reservoirs on the Volga river have their water levels drawn down in winter to be able to accommodate spring flood flows produced by snowmelt in the upper catchment. The coefficient of variation (CV) of annual flows in the Volga river is around 0.18, which is very low and indicates that there is a consistent and reliable total flow volume from year to year. In reality, however, the timing of the peak flow can shift across a space of two months. This cascade of thirteen large reservoirs and several medium-sized reservoirs forms a complex system of flow regulation, with great control over the flows of the Volga river into the Caspian Sea.

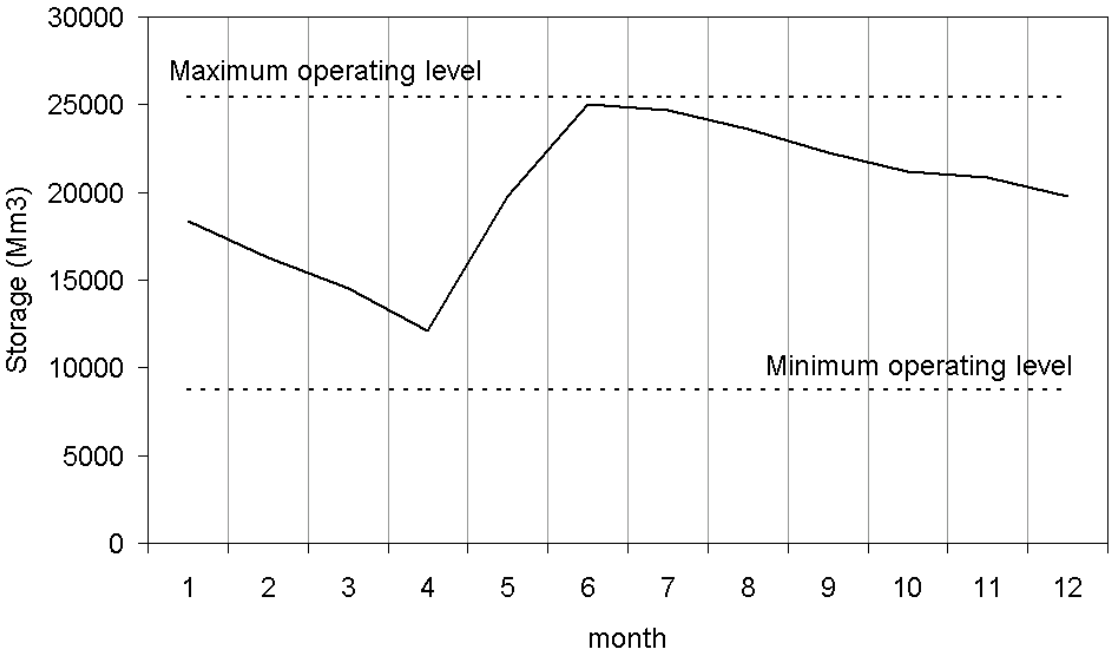
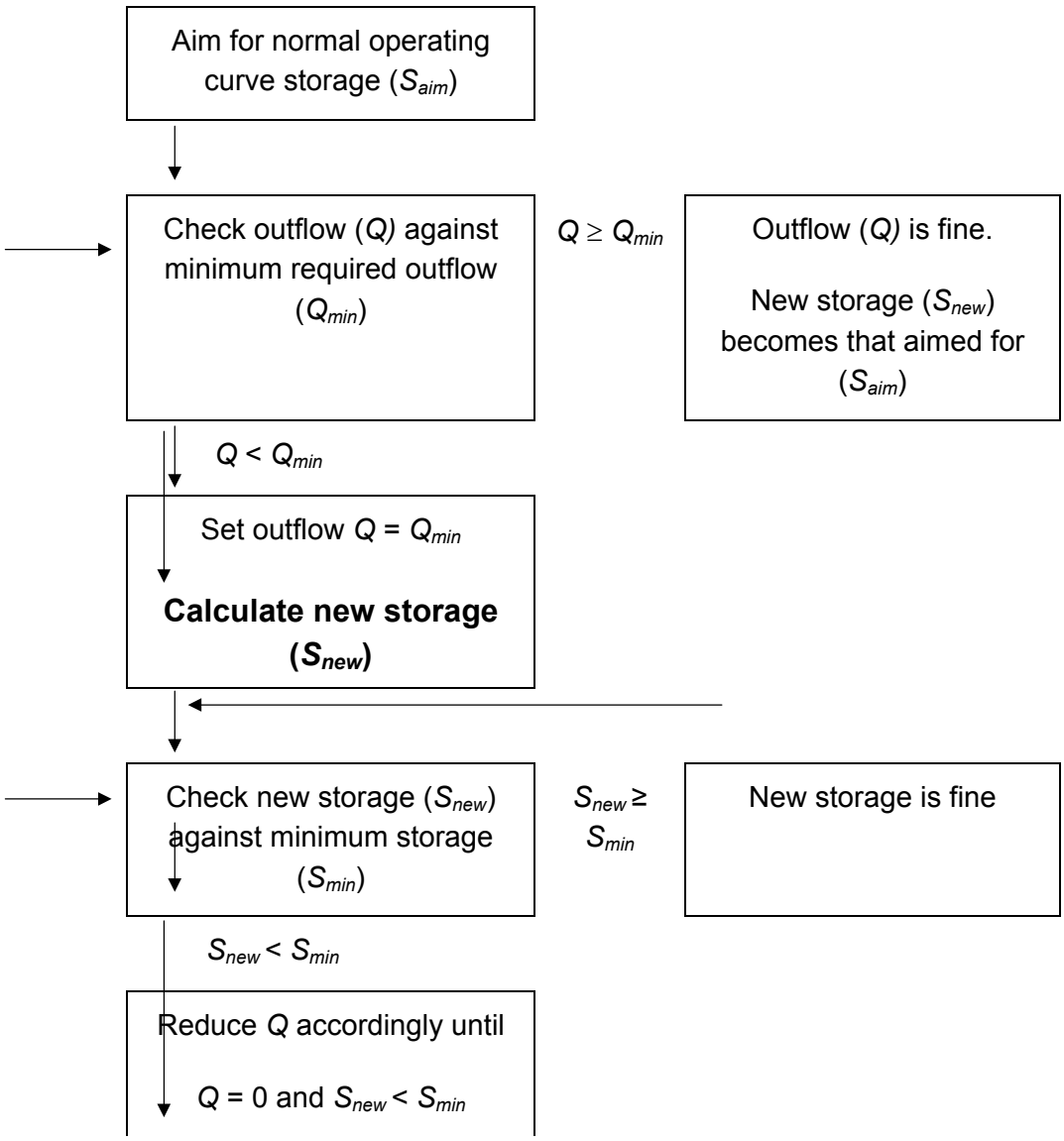


Figure 1 Schematic of a normal reservoir operating curve

The method of simulating such large reservoirs is represented schematically in Figure 2. This shows how the reservoir operation control curve is aimed for primarily in each month, outflows are checked to ensure that they satisfy the minimum required outflow, otherwise the actual storage in that month is reduced so that the minimum flow requirement can indeed be satisfied. Next, the resulting storage is checked against the minimum possible storage in the reservoir (the 'dead storage'; *i.e.* that storage which is below the outlets of the reservoir and cannot physically be accessed). In the event that this minimum storage is reached, the direct abstractions from the reservoir are necessarily curtailed. Should the reservoir volume reach its maximum possible, any spare inflow is simply spilled. This procedure is completed on a monthly time step, in line with the rest of the GWAVA model. Abstractions can be specified monthly and varying from year to year, rather than the mean annual total value used in the previous, simple reservoir module.



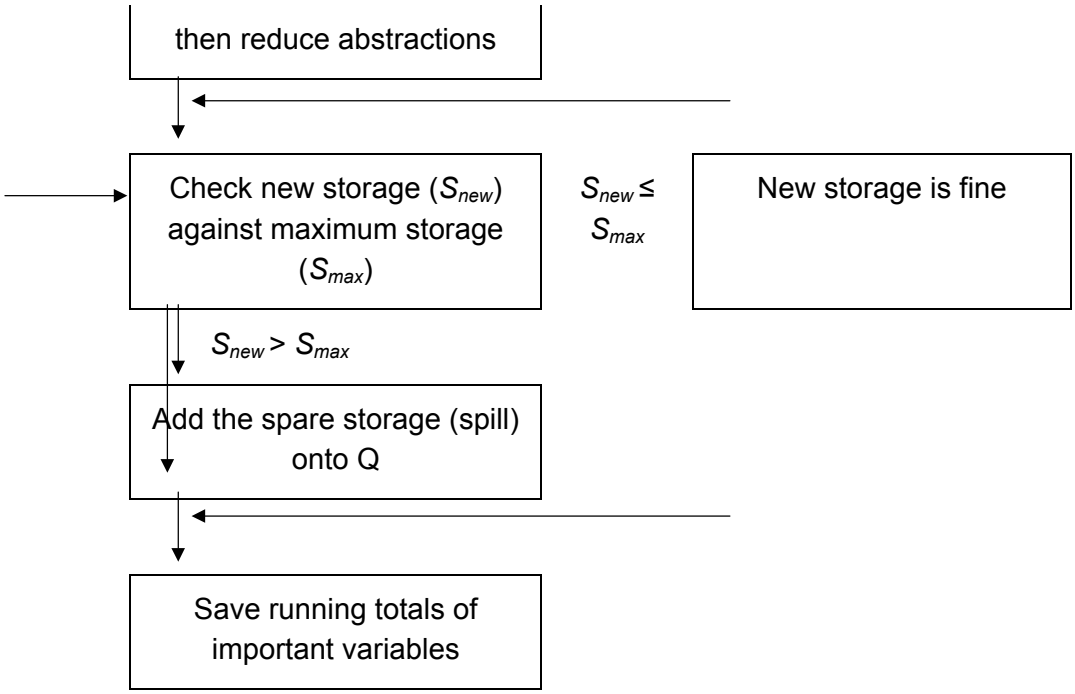
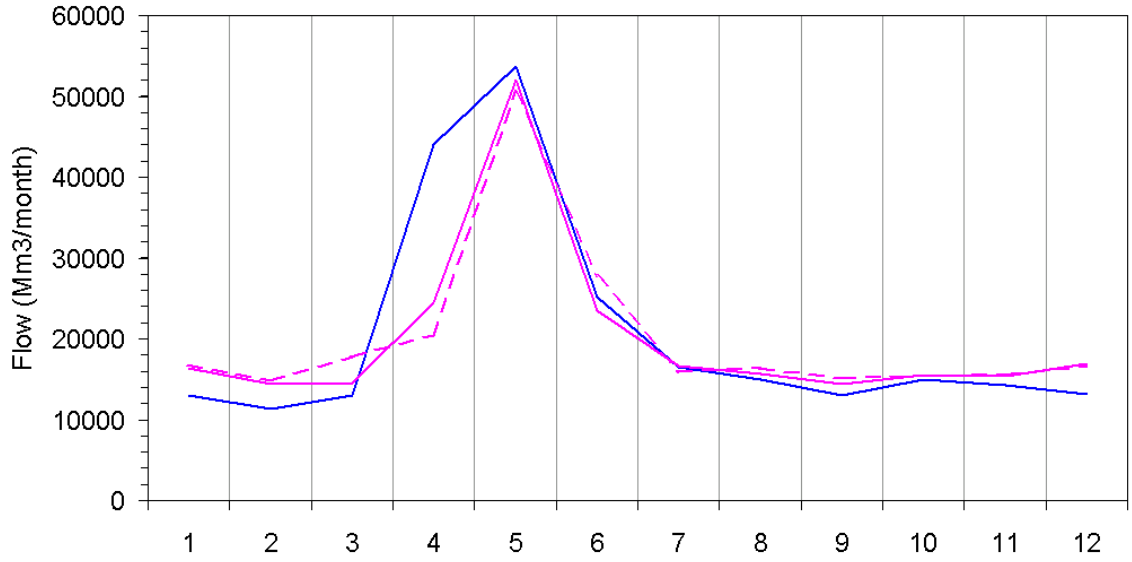
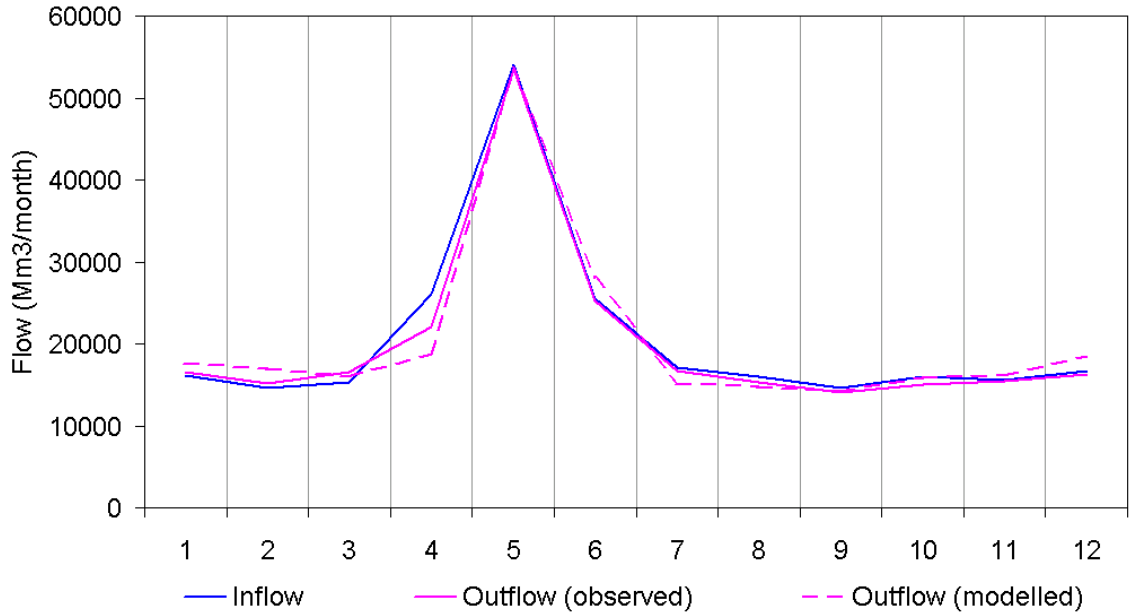


Figure 2 Flow diagram of the algorithm used to model large reservoirs

The algorithm was tested using observed inflow and outflow data for several large Volga reservoirs. The results of modelling several of the more important reservoirs show that the algorithm is effective in reproducing the observed behaviour (Figure 3).



(a)



(b)

Figure 3 Modelled and observed inflows and outflows from (a) Kujbshevskoye and (b) Volgogradskoye reservoir.

The State Hydrological Institute provided data on standard typical operating rules for each of the main Volga reservoirs. Operating rules represent the ideal operation pattern under average conditions; it is not clear how far actual operation varies from these patterns.

2.2.2. Glacier module

The need for development of a module to take account of glaciers was conceived when looking at the results from Phase I and the Intermediate Phase. Calibration of those rivers with a high proportion of their runoff originating from glacial melt-waters was less than satisfactory; this was especially true of the River Terek in the Caucasus region which derives 14% of its annual runoff from glacial melt, but more importantly derives around 30% of its *summer* runoff from glacial melt-waters. Two key areas for improvement were identified: modelling of *runoff volumes*, and modelling of *hydrograph shape*.

Modelled runoff volumes for the Terek were considerably lower than observed runoff volumes. This was probably due to elevation effects; the baseline climate was derived according to the mean elevation in each half degree cell. Significant variations in elevation can exist in the higher cells of the Caucasus mountains, but the elevation effects on climate were being masked through use of the mean elevation when generating climate data. In reality, precipitation tends to increase with elevation, and evaporation and temperature tend to decrease (up to an elevation limit of around 4000 metres – the ice limit – whence these trends are thought to reverse again).

The shape of the hydrograph was also considered to be in need of improvement. In the modelled situation, snowmelt was tending to occur only in the months of April and May, but in reality this snowmelt period is typically attenuated through to the mid summer months.

The modelling approach The approach taken to model the mountainous regions of the Caspian Sea basin, particularly the Caucasus region, is described as follows. For every cell lying in a mountainous region (*i.e.* every cell having a mean elevation over 2000 metres), extra information is required on the elevation distribution within the cell. These data are then represented as either (1) pairs of co-ordinates of the distribution function, or (2) parameters of the Pareto distribution fitting the elevation data. Either approach is suitable, but the Pareto distribution is simpler since it only requires three data values, whereas several more pairs of co-ordinates would normally be used. Additionally, information on the gradients of temperature and precipitation are also required for each cell, to adjust the baseline climate data accordingly. Initially these gradients were set at -10°C per kilometre above mean elevation (temperature gradient), +20% per kilometre above mean elevation (precipitation gradient), and –20% per kilometre above mean elevation (potential evaporation gradient).

Once the elevation distribution in a cell has been derived, the cell is divided into elevation (or climate) zones; within each zone the process of runoff formation is considered to be homogeneous. A modified version of the PDM rainfall-runoff model (the ‘quasiPDM’) is then applied to each climate zone to calculate its runoff formation. The quasiPDM is invoked for every cell classified as belonging to a mountainous region, instead of invoking the classic

PDM and snow-pack subroutines. The weighted sum of the component runoff values is then found, thereby representing the modelled runoff from the half-degree cell.

Details of the methodology Elevation data are taken from an external data file and an elevation distribution is created for the cell under consideration. The total elevation range is split into fixed elevation increments of 500 metres to define the computational zones.

The quasiPDM is based upon the classic PDM logic. However, within the quasiPDM, essentially there are three climate zones to consider for the different elevations, Z:

- Permanent glacier (Z > 4000 m);
- Grass (2000 m < Z < 4000 m);
- Grass plus forest (Z < 2000 m).

The quasiPDM requires temperature and elevation data in addition to those parameters normally passed to the PDM. The quasiPDM incorporates its own snowmelt module.

Each of the computational zones is processed individually. A check is made to see which of the climate zones it lies in (either wholly or partially). This process produces weights of the land coverage within each computational zone. The total runoff is then calculated as the sum of the runoff produced from each climate zone (the runoff of the grassy climate zone multiplied by the weighting of the grassy zone, etc), thus:

$$Runoff_{total} = W(1)*R_{grass} + W(2)*R_{forest+grass} + W(3)*R_{ice}$$

The model is then run on a daily basis. For the cell being processed, adjustments are made to the daily climate data according to the climatic gradients taken from the external data file (data passed from the main model). If the temperature is greater than the threshold melting temperature, then if there is part of the computational zone lying in the ice climate zone, its effective precipitation and direct runoff are calculated accordingly. The classic PDM is run for any grass and/or forest covered parts of the cell, including the interception model. The degree-day method of calculating the snow module is used to keep a track of the snow cover, just as it is in the classic PDM.

Elevation data are taken from the GED CD-ROM (NOAA-EPA, 1992) as ten-minute resolution modal, minimum and maximum elevations in metres. The data were created by Fleet Numerical Oceanography Center (FNOC) in 1992 as metres above sea level rounded to the nearest ten metres. The data were verified by comparison with a large-scale topographic map of a key area of the Caucasus mountainous region, and were found to be of good quality.

The sensitivities of the model's runoff formation to the gradients of temperature and precipitation in the sub-module were tested on an upstream glacier cell. Extremes of the temperature gradient of -6.5°C per km and -10°C per km were used; extremes of the precipitation gradient of +20% and +30% per km were used. This test was based on best

available knowledge of the extreme values the gradients are likely to take. Sensitivity of the model's evaporation to the temperature gradient was also tested.

Runoff formation was found to be much more sensitive to the gradient of temperature than to the gradient of precipitation (see Figure 4). In tests on eight upstream glacier cells in the Caucasus mountains, reducing the temperature gradient from 10 to 6.5 increased the annual runoff by an average of 33%. The gradients are centred on the mean elevation in the cell, since that is the elevation upon which the climate data are based. This means that the precipitation gradient for parts of the cell below the mean elevation will be negative and will reduce the actual precipitation values, whilst parts of the cell above the mean elevation will have a positive precipitation gradient and therefore higher precipitation. Sensitivity of evaporation to the temperature gradient was greatest in the winter months of November, December and March (see Figure 4).

This additional glacier modelling procedure was only invoked for those cells having a mean elevation over 2000 metres, a total of 59 cells in the whole basin.

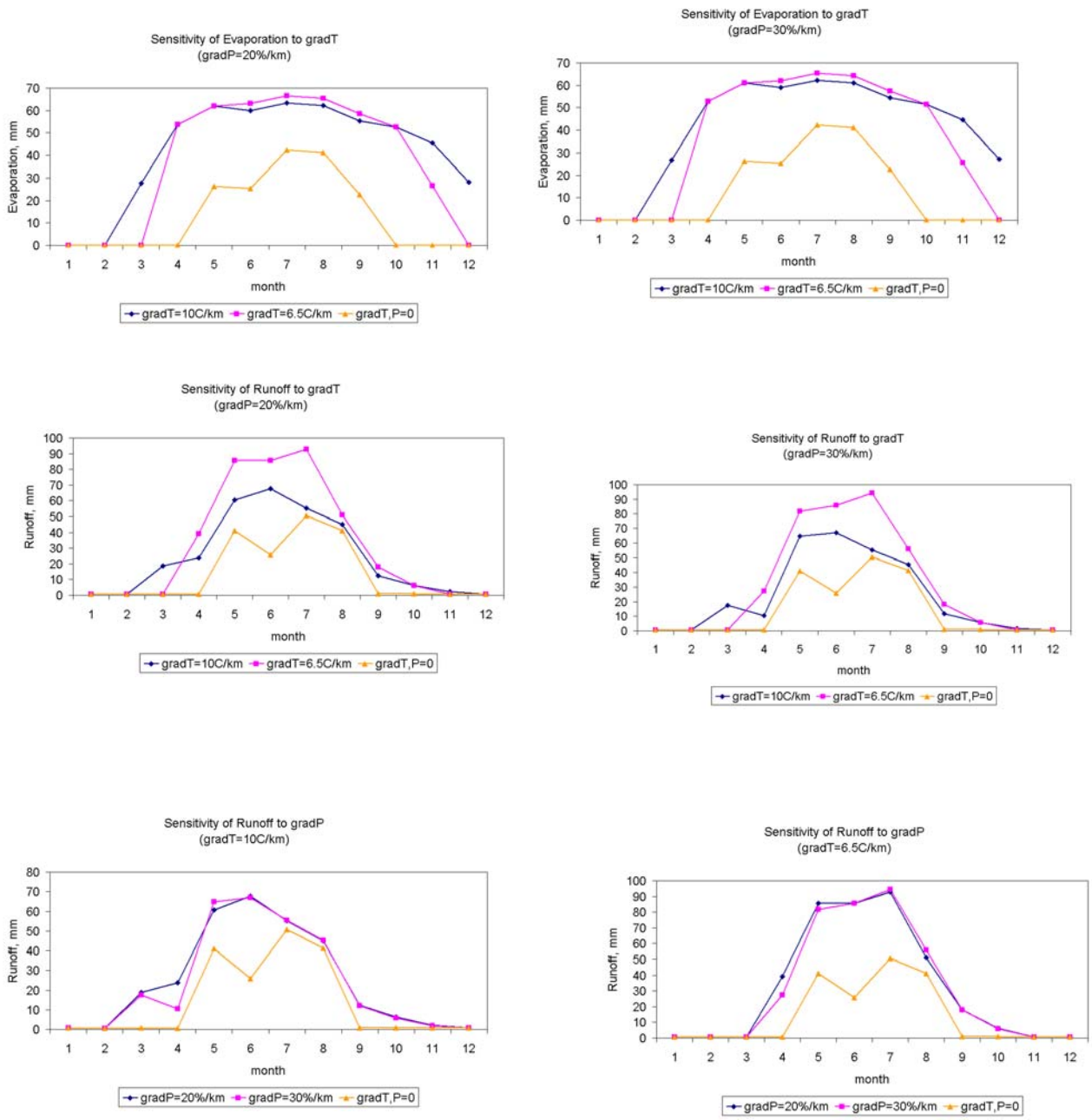


Figure 4 Sensitivity tests for climate gradients with elevation on a glacier cell (43°N; 42°E) with one year of input data

Note on the concept of soil freezing The sum of negative temperatures during a cold event (until the onset of melting) is the typical approach for calculating the depth of frozen soil. This depth of frozen soil is then used to reduce the infiltration capacity of the soil accordingly. The PDM rainfall-runoff model being used at the heart of the GWAVA model takes no account of infiltration capacity and thus this methodology could not be applied without making gross assumptions. In reality, the infiltration capacity depends on many factors including the soil

constitution and vegetation cover, and usually has a value equal to about 10% of the field capacity of the soil; it would require detailed calibration on a cell by cell basis, which is unsuitable here. Since the phenomenon of frozen soil only has an effect on runoff formation in severe winters, it was considered expedient to leave this modelling component for future interest. It would be of greater importance if the goal was to model extremes.

2.2.3. Water use module

In earlier phases of the work, water use was represented only as an annual value of abstractions less return flows which had to be constant through the whole period of a model run. In order to permit a more realistic representation of actual water use patterns, an additional component was added to allow water use quantities to vary in a monthly pattern and also for the monthly patterns to vary from year to year during the period of a run. The facility to do this for cells in which there are lakes or reservoirs is already included in the improved reservoir module (see Section 2.2.1); the additional module discussed here applies to cells where there is no lake or reservoir.

For cells where this type of data is available, the additional data files were created. It was found that the differences in the modelled flows were very minor. This was a result of the lack of detailed data for many of the larger abstractions which still had to be represented as constant annual amounts. Nevertheless, this additional module provides a useful facility which will become increasingly valuable when the Caspian countries make more detailed estimates of the patterns of large-scale water abstraction in the basin. Figure 5 indicates that net abstraction in the South-West region may have increased significantly over the baseline period (this represents net river abstractions where the variation over time is known, and excludes direct abstractions from reservoirs and net river abstractions for which only fixed data are available).

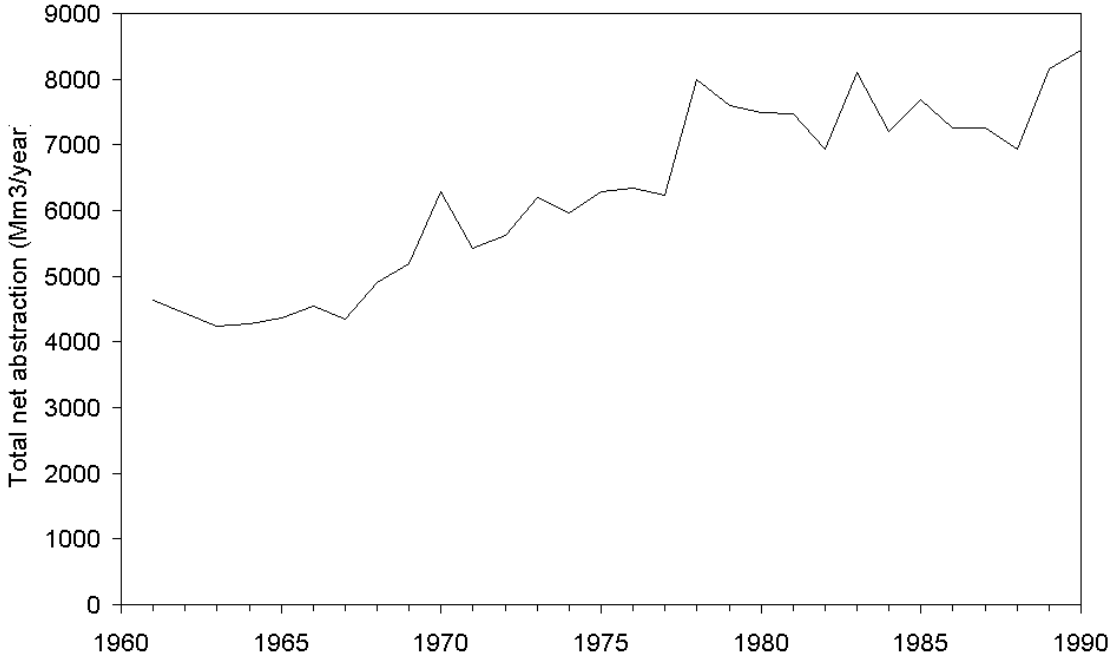


Figure 5 Net abstraction (where variation is known) in the South-West part of the Caspian Sea basin during the baseline period

2.3. Model calibration

The complete basin model was split into three manageable parts for the purposes of calibration: the Volga basin, the Ural basin, and the South-West basin. The calibration of each of the three regional models is described below, and the results are given in Section 3.1. The basin and some of the key gauging stations are shown in Figure 6. Additionally a more detailed gridded map of the South-Western part of the basin is shown in Figure 7, including the locations of the gauging stations used for calibration in that area.

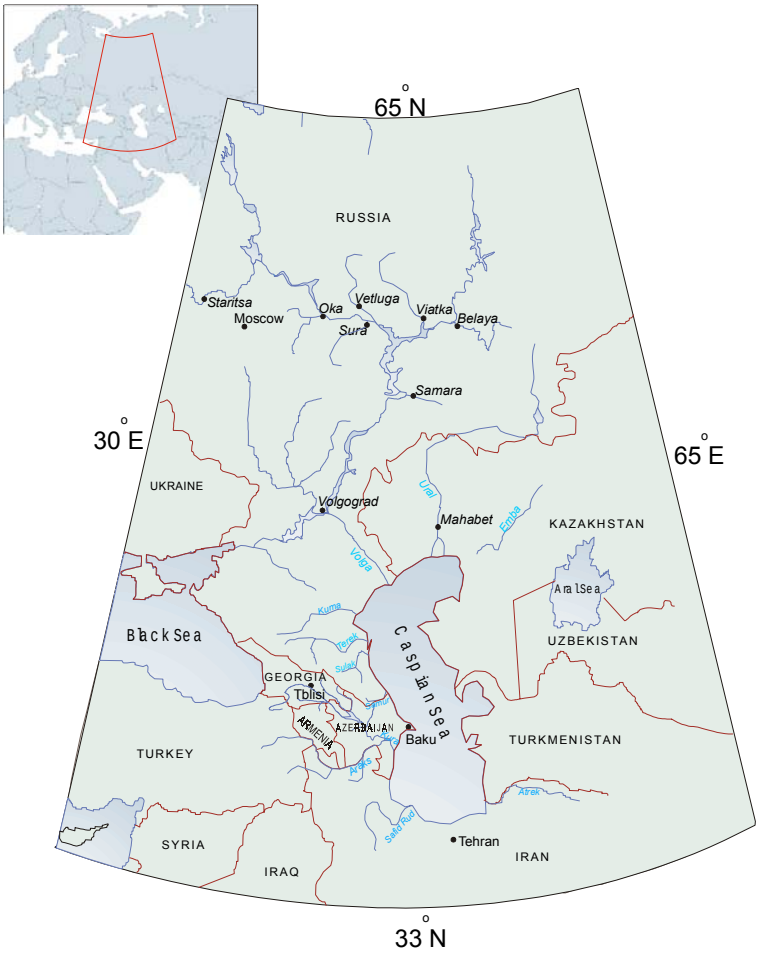


Figure 6 Map of Caspian Sea basin showing some calibration gauge sites

The large reservoirs simulated in the model are listed in Table 1. These reservoirs have significant impacts on the modelled hydrographs, and were an important consideration in the calibration procedure.

Reservoir name	River/model basin	Date of construction	Grid cell
Sheksninskoye/Belo ye	Volga	1964	59.0°N; 38.0°E
Rybinskoye	Volga	1964	58.0°N; 38.5°E
Kamskoye	Volga	1957	58.0°N; 56.0°E
Uglichskoye	Volga	1944	57.0°N; 38.0°E
Verkhnevolzhskoye	Volga	1948	56.5°N; 33.0°E
Ivankovskoye	Volga	1938	56.5°N; 37.0°E
Gorkovskoye	Volga	1958	56.5°N; 43.0°E
Votkinskoye	Volga	1965	56.5°N; 54.0°E

Cheboksarskoye	Volga	1989	56.0°N; 47.0°E
Nizhnekamskoye	Volga	1980	55.5°N; 52.0°E
Kujbyshevskoye	Volga	1958	53.5°N; 49.0°E
Saratovskoye	Volga	1969	52.0°N; 47.5°E
Volgogradskoye	Volga	1961	49.0°N; 44.5°E
Iriklinskoye	Ural	1967	51.5°N; 58.5°E
Mingechaurskoye	SW (Kura)	1954	40.5°N; 47.0°E
Araksinskoye	SW (Arax)	1973	38.5°N; 45.5°E

Table 1 Large reservoirs of the Caspian Sea basin

The observed flow series used for calibration are summarised in Table 2. Observed flow data were assembled for key locations in the basins, especially at the mouths of the major rivers flowing into the Caspian Sea. Some upstream gauges were used, to enable sub-catchment calibration to be performed. The sub-catchment calibration feature of the model was developed to cover calibration of the baseflow regression power (in mountainous and non-mountainous cells), in addition to the six parameters used for calibration in the Intermediate Phase (Vörösmarty parameter, surface water routing parameter, groundwater routing parameter, snowmelt factor, threshold temperature to split precipitation into snow or rain, threshold temperature for snowmelt). For more details of the sub-catchment calibration procedure, see the User Notes in Annex B.

In examining the quality of fit of the model, most weight was put on obtaining good agreement between the observed and modelled mean annual runoff volume, since it is long-term volumes of runoff which are crucial in determining the resulting levels of the Caspian Sea. The average monthly patterns of runoff were also examined to check that seasonal behaviour was being reproduced adequately. Finally, the time series of modelled flows were compared to the observed time series, to check that the model was simulating the variation in runoff from year to year.

2.3.1. Volga basin

The Volga basin is the largest of the three regional models, and produces the large majority of the inflow to the Caspian Sea (typically more than 80%), so the calibration of this model was particularly important. A total of eight gauged flow series was used, with flows at Volgograd representing the outflow of the Volga river into the large and complex Volga river delta, and thence into the Sea. Modelling of the Volga delta is not included here, but is considered as part of the modelling of the water balance of the Sea. The Volga river is highly regulated by a cascade of large reservoirs. The total number of large reservoirs being modelled by the newly developed reservoir module is thirteen. They are listed in Table 1. Because of the importance of the reservoirs, adjustment of the operation rules was also required as part of the calibration process.

2.3.2. Ural basin

The Ural basin is the smallest of the three regional models, and a total of two gauged flow series was used. The Ural river is regulated by a single large reservoir, detailed in Table 1.

2.3.3. South-West basin

A total of nine observed flow series was used for calibration; some of these series represent aggregates of small rivers. In one case (small rivers of the Iranian coast) the only available observations are at locations inland, and the observed series for these combined basins has been estimated by SHI based on the available information on river flows and abstractions in that area. There are two large reservoirs in this part of the Caspian Sea basin, detailed in Table 1.

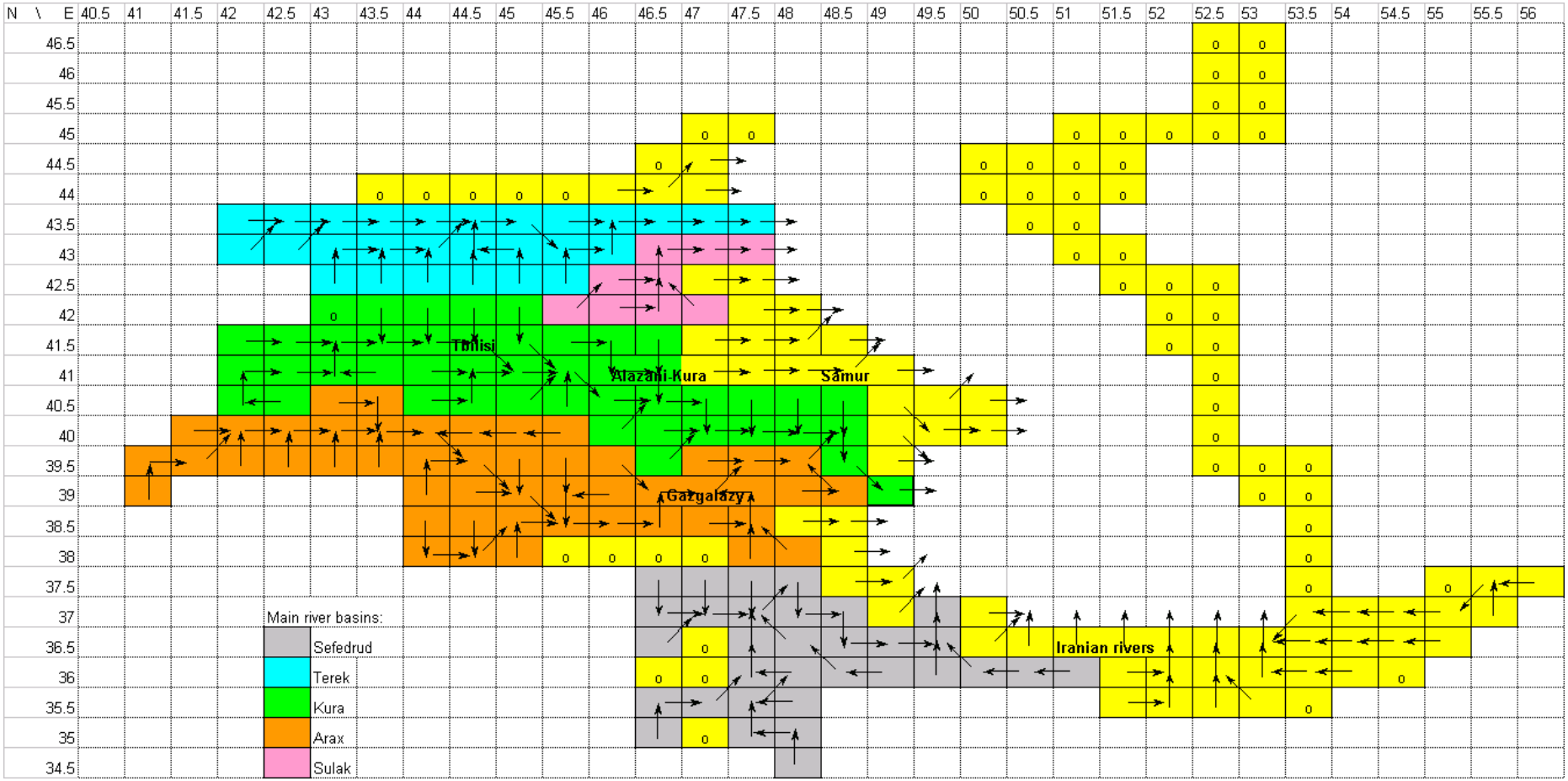


Figure 7 Map of South-West basin showing grid cells, main flow directions, and locations of main river basins and gauging stations

Model	Gauge	Data period	Observed flow (Mm^3)												
			Annual	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Volga	Oka	1961-90	39383	1935	1543	2076	10338	8644	2613	2191	1960	1815	2052	2191	2025
	Belaya	1961-90	26622	893	816	1079	5566	7633	2304	1708	1355	1215	1652	1400	1001
	Vyatka	1961-90	30529	877	645	716	5406	10342	3540	1886	1464	1237	1726	1461	1230
	Staritsa	1961-90	4782	154	162	343	1338	604	398	399	302	240	313	312	218
	Sura	1961-90	6072	204	176	382	2564	809	321	293	262	230	266	296	270
	Samara	1961-90	3334	86	70	174	1735	458	157	120	102	96	115	113	106
	Vetluga	1961-90	6750	143	96	110	1110	2609	744	430	293	232	367	362	256
	Volgograd	1961-90	242154	16507	15311	16577	22050	53540	25250	16666	15346	14011	15097	15498	16300
Ural	Ural to Sea	1961-90	6821	167	149	210	950	2209	1356	526	321	248	256	237	192
	Kushum	1961-85	8234	193	171	260	2092	2878	884	472	323	260	263	249	187
South-West	Terek to Sea	1961-90	6258	417	384	400	245	466	744	917	875	592	432	361	424
	Sulak to Sea	1961-90	4251	196	202	178	254	540	710	693	454	320	244	219	242
	Samur to Sea	1961-90	1496	52	41	55	130	223	330	277	182	71	62	41	33
	Kura to Sea	1961-90	15686	1486	1448	1586	1685	2022	1563	888	768	827	946	1101	1365
	Arax at Gazgalazy	1964-81	7448	552	519	658	1082	1388	875	537	415	286	331	354	452
	Alazani to Kura	1961-90	3481	187	173	260	417	537	472	307	220	225	243	226	212
	Small rivers of Iranian coast	1961-90	10144	504	616	997	1723	1988	1199	581	447	491	534	546	518
	Small rivers of West coast	1961-90	1031	54	62	111	111	96	98	74	58	87	123	93	65
	Kura at Tbilisi	1961-90	6178	230	232	468	1210	1430	868	430	258	236	271	277	269

Table 2 Summary of observed flow series used for model calibration

2.4. Development of future scenarios

Many researchers believe that global climate change is now occurring and is expected to continue in the future as a result of the emission of carbon dioxide and other greenhouse gases. These changes would be expected to lead to changes in precipitation, temperature and evaporation which could have profound effects on the water resources of the study region.

It is inevitable that changes in demand for water will occur, with growing populations, increasing levels of urbanisation and technological development; these will also have effects on water resources availability, and a range of plausible scenarios is studied here in combination with climate change.

2.4.1. Climate change scenarios

This section describes how the results of global climate models have been used to derive scenarios of precipitation, temperature and potential evaporation changes for the year 2050 to be applied as inputs to the water resources model.

The baseline period The baseline period is the popular 30-year 'normal' period as defined by the World Meteorological Organisation (WMO). The current WMO normal period is 1961-90. Use of a recommended baseline provides a standard reference to ensure comparability between impact studies. Other benefits of using this baseline are: the period ends in 1990, which is the common reference year used for climatic and non-climatic projections by the IPCC (Intergovernmental Panel on Climate Change); it represents the recent climate; most observed climatological data are most readily available for this period (IPCC-TGCIA, 1999).

The selected GCM scenarios The Intergovernmental Panel on Climate Change (IPCC) strongly recommends using the results of several GCMs in an impact assessment, due to the known uncertainties of GCMs. The scenarios used in this study are HadCM3, Echam4, CGCM1, selected because they are adequate to represent the range of results that are available and include results of one of the latest GCMs. Also, these are some of the few model outputs that provide us with the necessary variables to calculate potential evaporation. The three models are:

- HadCM3 – UK Hadley Centre for Climate Prediction and Research
- CGCM1 – Canadian Centre for Climate Modelling and Analysis
- ECHAM4 – German Climate Research Centre

Each of the GCMs uses a number of atmospheric and ocean layers coupled together to simulate the climate. The HadCM3 model comprises 19 atmospheric levels and 20 ocean levels, and has a spatial resolution of 2.5° by 3.75° (latitude by longitude). HadCM3 has a much improved sea surface temperature (SST) and sea ice climatology when compared to earlier Hadley Centre models without flux adjustments, and is one of the most recent developments in climate modelling. The model has been run for simulation periods of more

than a thousand years. The CGCM1 model has 10 vertical atmospheric levels, and 29 ocean layers. Simulations go back to 1850. An increase of CO_2 at a rate of 1% per year compounded is assumed from 1990, and its climate sensitivity is about 3.5°C. The spatial resolution is approximately 3.7° by 3.7°. The ECHAM4 GCM was originally an atmospheric model which was later parameterised to be used for climate simulations. Historic forcing from 1860 to 1990 is used, followed by a 1% annual increase in radiative forcing to 2099. It has 19 atmospheric layers and is coupled to an ocean model. The spatial resolution is approximately 2.8° by 2.8°, and the climate sensitivity is about 2.6°C. More details of the models can be found on the IPCC-DDC web site (http://ipcc-ddc.cru.uea.ac.uk/) and the UK Hadley Centre website (http://www.met-office.gov.uk/research/hadleycentre/). Results from the different GCM experiments are provided as 30-year averages of monthly absolute changes between the baseline period 1961-1990 and a future time horizon. The time horizon used was the decade of the 2050s (2040-2069). The same meteorological variables as in the baseline climatology were extracted and potential evaporation was calculated in the same way. The mean monthly changes were then expressed in terms of percentage change (precipitation and evaporation), or absolute change (temperature), compared to the baseline values. This provides a means of downscaling the GCM results from the spatial resolution used in the experiment to the 0.5° resolution used for the rainfall-runoff modelling. The same rainfall anomaly series as derived for the baseline period was used for the future scenarios since the available GCM results do not provide reliable information on the year-to-year variability of precipitation.

The selected emissions scenarios A main selection criterion, when choosing climate change scenario results for use in modelling studies, is to use SRES emissions scenarios as much as possible (SRES is the IPCC Special Report on Emissions Scenarios). The SRES scenarios cover a wide range of the main driving forces of future emissions, from demographic to technological and economic developments. The IPCC recommends that the new scenarios be used not only in the IPCC's future assessments of climate change, its impacts, and adaptation and mitigation options, but also as the basis for analyses by the wider research and policy community of climate change and other environmental problems. They recommend that a range of SRES scenarios with a variety of assumptions regarding driving forces be used in any analysis.

The selected GCM results were derived using emissions scenarios IS92a and A2a. The IPCC-DDC (Data Distribution Centre) provides results from a range of Global Climate Model (GCM) experiments carried out by a number of climate modelling centres. These experiments are based on a set of emission scenarios of net greenhouse gases and aerosols for the next 100 years. The emission scenarios have been designed to give pertinent, plausible and alternative futures. They are deduced from population rise and economic growth projections, combined with projections in the type of energy sources to be used.

A set of four different scenarios define the IS92 emissions scenarios recommended by the IPCC Second Assessment Report (IPCC, 1995). The IS92a emission scenario used in this study corresponds to an increase in the concentration of greenhouse gas in the atmosphere equivalent to an increase in the CO_2 concentration of 1% per annum. It assumes a mix of conventional and renewable energy sources being used, and follows the 1991 World Bank

population projection, but with economic growth assumptions generally below the World Bank's ones. It has been widely adopted as a standard scenario for use in impact assessments.

Since their inception in 1994, new developments have called for improvements to the IS92 emissions scenarios, and it is one of this new set of emissions scenarios that is used in the HadCM3 results employed here (A2a). These scenarios include the latest information on economic restructuring throughout the world, different trends in and rates of technological change, and reflect current emission commitments. The full set of SRES new emissions scenarios is broadly described as follows (from the IPCC Data Distribution Centre):

- A1 – A future world of rapid economic growth, low population growth and rapid introduction of new and more efficient technology;
- A2 – A heterogeneous world, where regional cultural identities are strengthened. There is high population growth and less concern for rapid economic development;
- B1 – A convergent world, with emphasis on global solutions to environmental and social sustainability;
- B2 – A world with emphasis on local solutions to economic, social and environmental sustainability.

The SRES A2 storyline is considered by the IPCC Data Distribution Centre as the best approximation for the 1% forced GCM results and the SRES B1 storyline as the best approximation for the ~0.5% forced experiments.

The full text of the IPCC SRES report can be found at http://www.grida.no/climate/ipcc/emission/ index.htm.

Changes of means and variability Outputs of GCMs are applied here as monthly adjustments to the baseline climate, assuming no change in climatic variability between the baseline and future climate. Thus the pattern of inter-annual variability of climate remains unchanged. However, changes in climatic variability may be as important as changes in the mean climate for a region. Unfortunately, GCM estimates of future climatic variability are subject to great uncertainties. Having said that, there is some evidence to suggest that precipitation variability may change, with increases in precipitation intensity, giving more extreme rainfall events (IPCC-TGCIA, 1999).

Precipitation scenarios Using the above approach, percentage changes in monthly precipitation were derived for each cell for the year 2050. While, on an annual basis, the precipitation changes seem reasonable, there are problems in applying the GCM results on a monthly basis. This comes about because the baseline precipitations generated by the GCMs are different to the baseline precipitation values used in this study. The model is very sensitive to the precipitation inputs, and some effort has been taken to derive a plausible set of baseline precipitation data. However, an examination of the baseline precipitation data from the GCMs shows that this is considerably less satisfactory, as might be expected from the results of models which run on coarse grids. Thus, it was necessary to retain the original

baseline precipitation data set, and apply the scenario changes to these values.

Maps of the GCM precipitation outputs used in the grid model for the region can be seen in Figure 8, along with temperature outputs (Figure 9) and potential evaporation outputs (Figure 10). Note that these plots can be misleading since they represent average annual values rather than the monthly data used in the model.

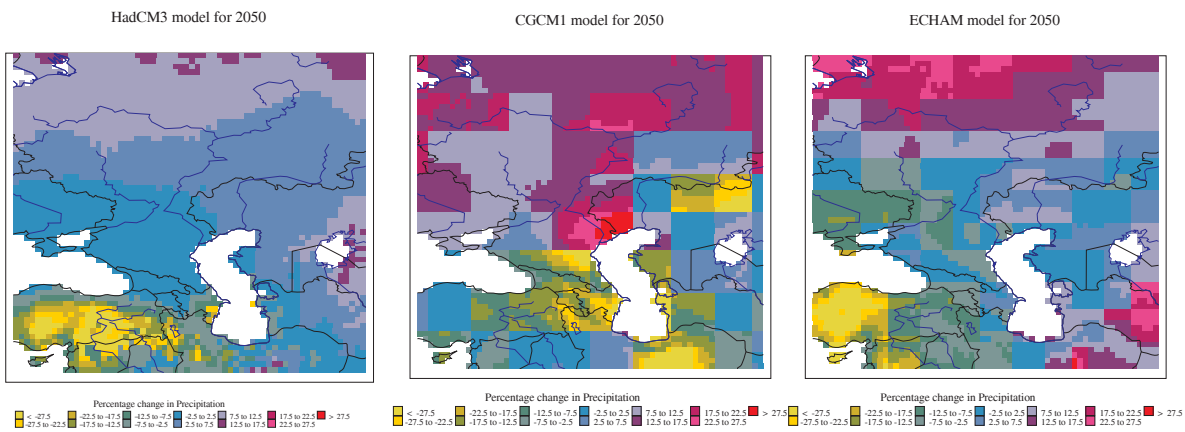


Figure 8 Comparison of baseline and 2050 monthly precipitation data for a) HadCM3, b) CGCM1 and c) Echam4

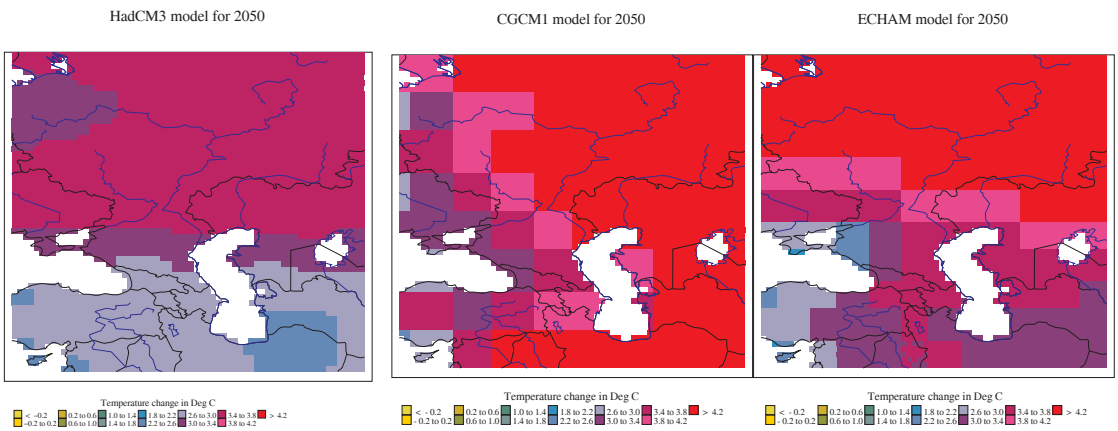


Figure 9 Comparison of baseline and 2050 monthly temperature data for a) HadCM3, b) CGCM1 and c) Echam4

Evaporation scenarios To obtain evaporation scenarios for 2050, the procedure described above was followed to derive monthly values for each cell for each of the following variables: temperature (°C); specific humidity (g/kg); incoming solar radiation (W/m^2); wind speed (m/s);

and total cloud amount. A summary of the annual values and changes is given in Table 3. Of these variables, temperature is the most significant driving force for potential evaporation, and the annual changes in temperature are plotted in Figure 9 (a, b, and c) for each of HadCM3, CGCM1 and ECHAM4 respectively.

The second most important driving variable for evaporation is humidity. These data were used in the Penman method to derive estimates of monthly potential evaporation for both the baseline condition and the scenarios. Because the input data are different to those used in the baseline estimates of evaporation, a different formulation of the Penman equation had to be used. In particular, the specific humidity values had to be converted to values of vapour pressure, and the total cloud amount figures had to be used in combination with the incoming solar radiation data in calculating the radiation terms in the method. The results of the calculations were the percentage change in PE for each cell. In a similar manner to the precipitation scenarios, the changes in PE were applied to the previously derived evaporation data set. The PE data and changes are summarised on an annual basis in Table 3 and Figure 10 (a, b, and c) for each of HadCM3, CGCM1 and ECHAM4 respectively.

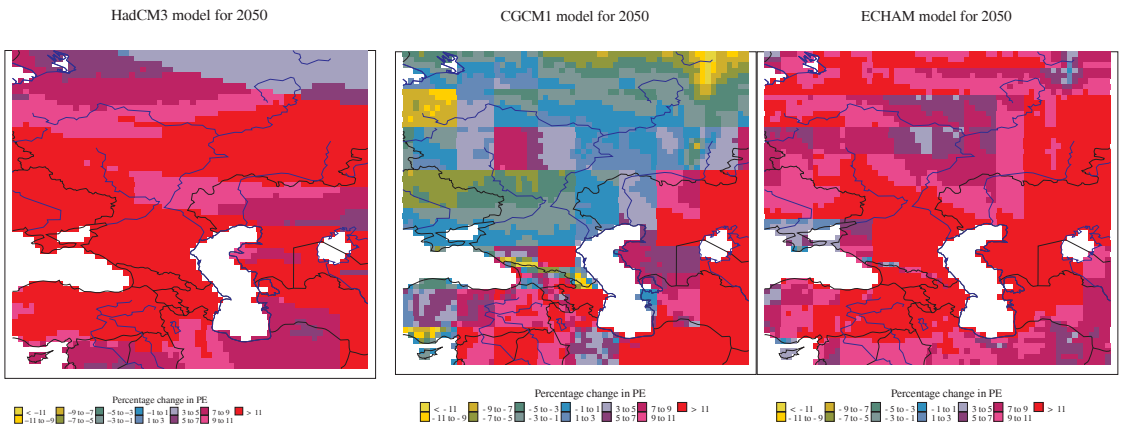


Figure 10 Comparison of baseline and 2050 monthly potential evaporation data for a) HadCM3, b) CGCM1 and c) Echam4

		Baseline value (1961-90)	Type of change *	GCM scenario (2050)		
				HadCM3	CGCM1	ECHAM4
Annual total precipitation (mm)	Min.	115.3	(Ave. annual %)	-10.1	-32.2	-13.1
	Ave.	516.1		2.5	7.7	7.6
	Max.	1106.0		14.0	33.1	23.8
Annual mean temperature (^{o}C)	Min.	-4.5	(Ave. annual abs)	2.5	3.1	3.1
	Ave.	5.2		3.5	4.5	4.3
	Max.	20.3		3.8	6.7	5.1
Annual total potential evaporation (mm)	Min.	485.4	(Ave. annual %)	8.1	-17.7	0.6
	Ave.	852.2		18.6	1.6	9.9
	Max.	1588.2		35.5	30.5	23.8

* Values presented for changes are the minima, averages and maxima calculated from the average annual change in the variable for each cell. Changes are either percentages (%) or absolute values (abs), as indicated.

Table 3 Summary of annual values and changes in meteorological variables for 2050 averaged over the Caspian Sea basin

Discussion of GCMs It can be seen above, that in some aspects the three GCMs give quite different results. It is the expected changes in rainfall that are most important for a water resources study such as this, and these are also the GCM outputs which differ most between models. The HadCM3 model predicts only a 2.5% increase in rainfall across the basin, while Echam4 and CGCM1 predict much larger increases of around 7.5% to 8%. It can also be noted that there is a difference from earlier phases of the project where the HadCM2 model was used (HadCM3 results were not available at that time). Whereas HadCM2 rainfall changes were roughly of the same magnitude as Echam4 and CGCM1 (lying approximately between the two), the more recent HadCM3 predicts much smaller changes. These differences in results of HadCM3 compared to HadCM2 are supported by a worldwide comparison of the two models which shows for many parts of the world a complete reversal of the expected patterns of future rainfall (IPCC, 2001b). Similarly, the divergence of rainfall results for the different GCMs has been noted for much of the world, with the Caspian region being one where a wide range of different predictions are available (IPCC, 2001a).

It should also be noted that HadCM3 differs in that a newer emissions scenario (A2a) has been used, while the other models use the older scenario (IS92a) as updated results are not yet available for them. However, these emissions scenarios are broadly similar, and the differences are small in relation to the inherent difference between the models themselves.

HadCM3 is a newer model which is more sophisticated in a number of respects, and on this basis, it might be expected that the results would be better or more reliable than the older

models. However, all the models are considered to be state-of-the-art approaches, and there is no objective way to determine which is better. Therefore, following IPCC recommendations, it is considered that the only reasonable approach is to use the outputs of a range of GCMs, treating each as equally valid. This approach is widely followed and accepted for climate impact studies. From the GCM results for rainfall, we can conclude, that with the present state of knowledge in global climate modelling, future rainfall over the Caspian basin (as elsewhere) is highly uncertain. What the models do agree on much more closely is that there will be a substantial rise in temperature, with average increases across the basin of between 3.5 and 4.5°C according to the three models. This change will have some clear consequences for water resources, including a reduced proportion of precipitation falling as snow.

Using these factors the absolute quantities of net demands (abstractions less return flows) are given in Table 5, and are plotted for the whole basin in Figure 1. The demands for future scenarios have been applied in the model at the same locations as used for the existing demands.

River basin	Year	Demand scenario	Time horizon	
	2000		2030	2050
		High	15840	18221
Volga	7592	Medium	12470	15184
		Low	7920	7592
		High	2383	2913
Ural	1923	Medium	2251	2648
		Low	1854	1854
		High	6817	8426
Kura	5074	Medium	6120	7165
		Low	5074	5074
		High	5971	6872
Terek	5070	Medium	5746	6422
		Low	5070	5070
		High	1305	1595
Sulak	1015	Medium	1233	1450
		Low	1015	1015
		High	4250	4397
Arax	4104	Medium	4192	4280
		Low	4104	4104
		High	6600	7200
Iran	6000	Medium	6300	6600

		Low	6000	6000
		High	42510	49624
Total	30777	Medium	37987	43749
		Low	30708	30708

Table 4 Net abstractions of the Caspian Sea basin (Mm^3/year)

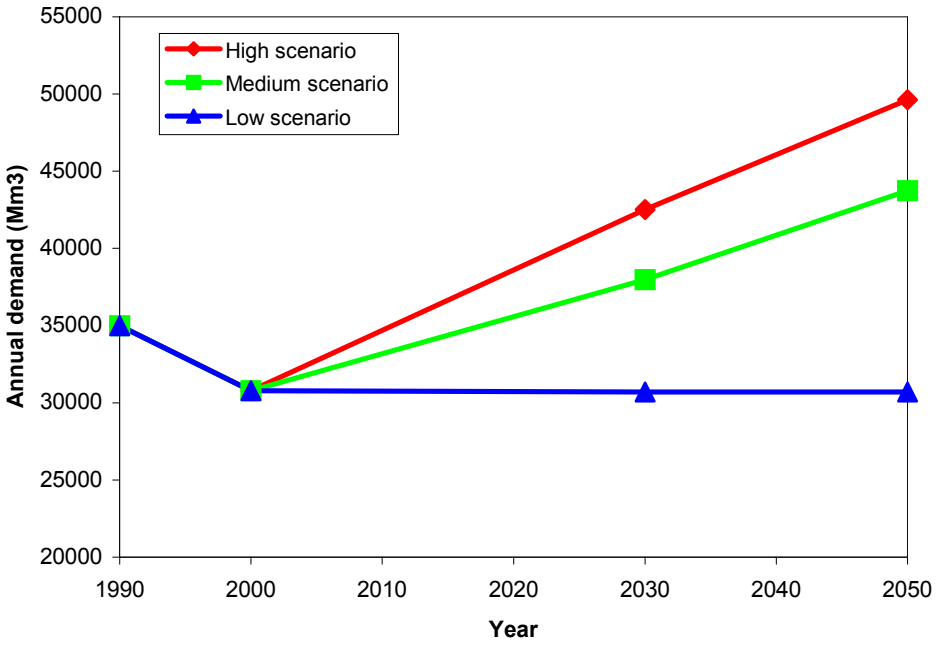


Figure 11. Summary of water demand scenarios for the whole Caspian Sea basin

Chapter 3. Results

The results of the grid-modelling of the Caspian Sea basin are presented in two sub-sections: the first shows the results of the calibration of the baseline model for current conditions; the second describes the application of several scenarios of future climate and demand, and their effects on Caspian Sea inflows.

3.1. Model calibration

The model calibration was divided into three parts, as described in Section. Each site was calibrated in three stages: first, mean annual total volume; second, mean monthly hydrograph; third, time-series results. The results of the model calibration are presented below.

3.1.1. Volga basin

Results of the modelling of baseline flows at each of the Volga catchment calibration gauges can be found in Table 5 and the comparison of the mean monthly hydrographs for the modelling period can be seen in Figure 12 (a-h). The time-series output, comparing observed and modelled flows for the 1961-90 baseline period can be seen in Figure 13. 'Modelled' results refer to the best model calibration run in each case.

Gauge site	Mean annual total flow (Mm^3)		
	Observed	Modelled	Difference
Oka	39383	41870	+6.3 %
Belaya	26622	22652	-14.9 %
Vyatka	30529	28462	-6.8 %
Staritsa	4782	5987	+25.2 %
Sura	6072	10945	+80.3 %
Samara	3334	7682	+130.4 %
Vetluga	6750	6863	+1.7 %
Volgograd	242154	238310	-1.6 %

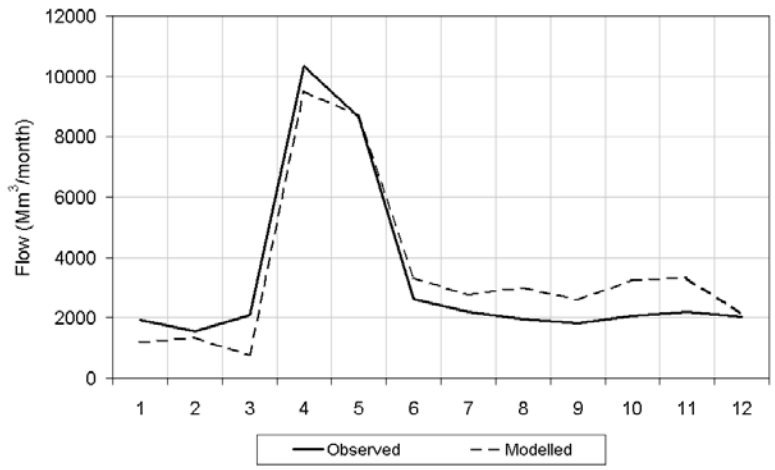
Table 5 Results of calibrating mean annual total volumes at gauges in the Volga basin

The key result here is for the Volga at Volgograd, which is downstream of all the other gauged locations, and represents the Volga inflows to the Caspian Sea. Compared to the other basins (Ural and South-West), the Volga is by far the major inflow to the Caspian Sea, contributing 83% of the total. Therefore the most important calibration result is the modelling of the annual volume at Volgograd. It can be seen that this is very closely matched, diverging from the observed by less than 2%. The pattern of monthly flows at Volgograd (Figure 12h) is also very well reproduced. It should be noted that one of the major factors controlling the flow in the Volga is the operation of the cascade of large reservoirs. This means that the observed flow hydrograph is largely a result of man-made factors, and, by reliably modelling the

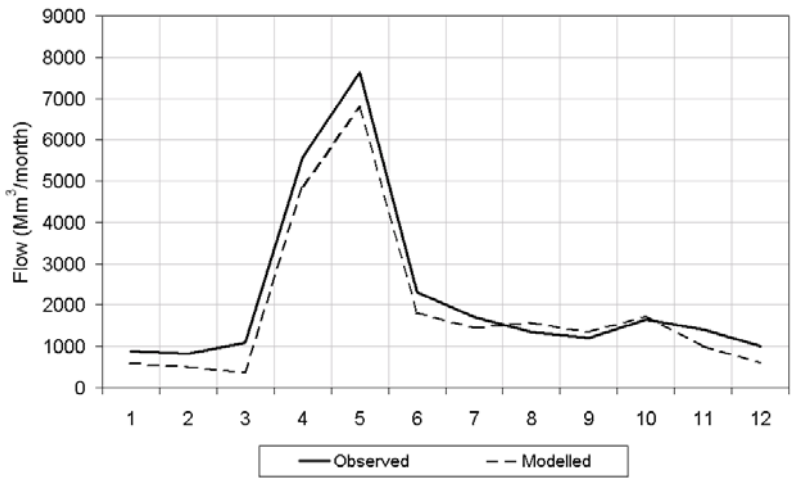
operation of the reservoirs (see Section 2.2.1) in combination with realistic runoff generation across the basin, accurate simulation of the observed flows has been achieved. The time series (Figure 13) shows that the year-to-year variability at Volgograd is also reasonably well simulated, but the modelled flows tend to be somewhat less variable than the observed. This may be partly a result of the input data or the model itself not truly reflecting the real variability. But, in addition, reservoir operation has been maintained constant from year to year in the model, whereas in reality there must have been some variation in procedures. This factor means that the modelled flows would tend to be less variable than the observed.

As already mentioned, the key calibration point is Volgograd. Results for the other locations upstream show that in the main the model is able to reproduce the distribution of runoff across the basin reasonably well. The larger sub-catchments (Oka, Belaya and Vyatka (Figure 12 a-c)) are well calibrated. Some of the smaller catchments (Figure 12d-g) are not so good, but these contribute only a small proportion of the total runoff, so this is not of any great significance.

(a) Oka



(b) Belaya



(c) Vyatka

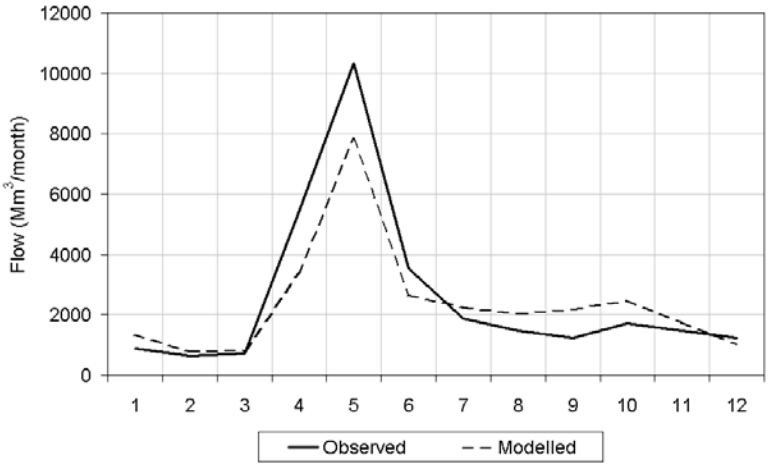
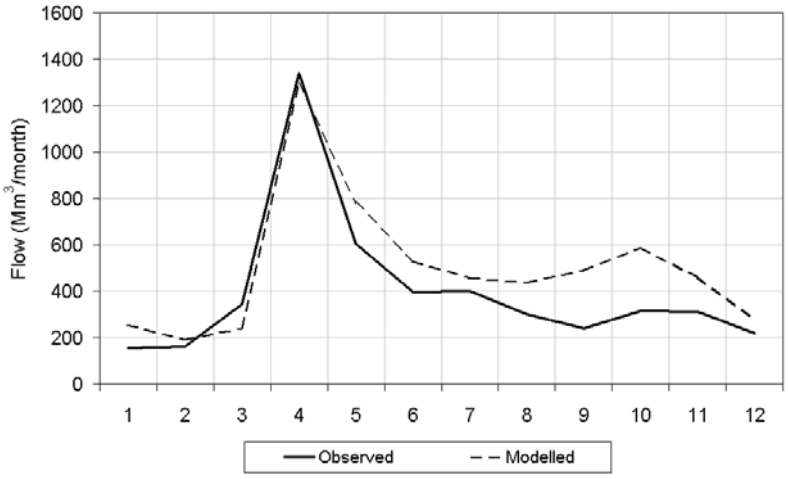
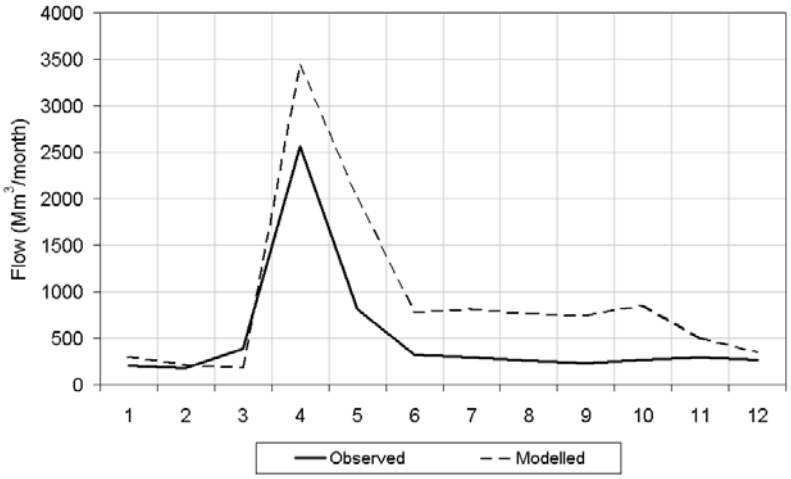


Figure 12 Volga basin model calibration results

(d) Staritsa



(e) Sura



(f) Samara

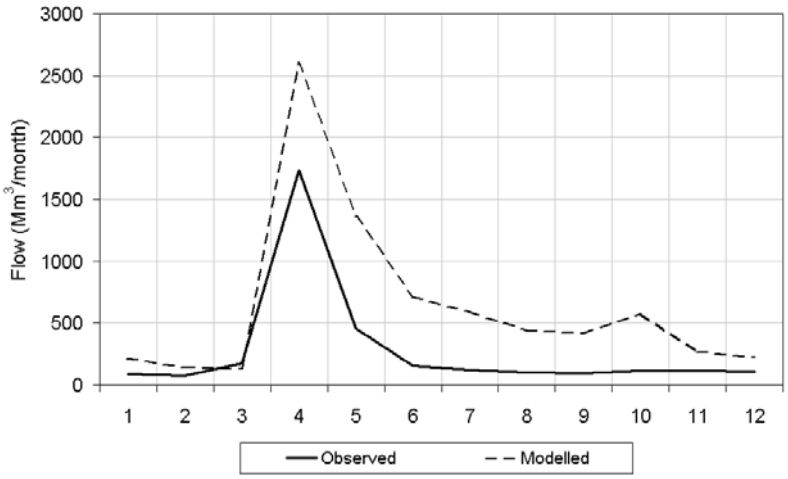
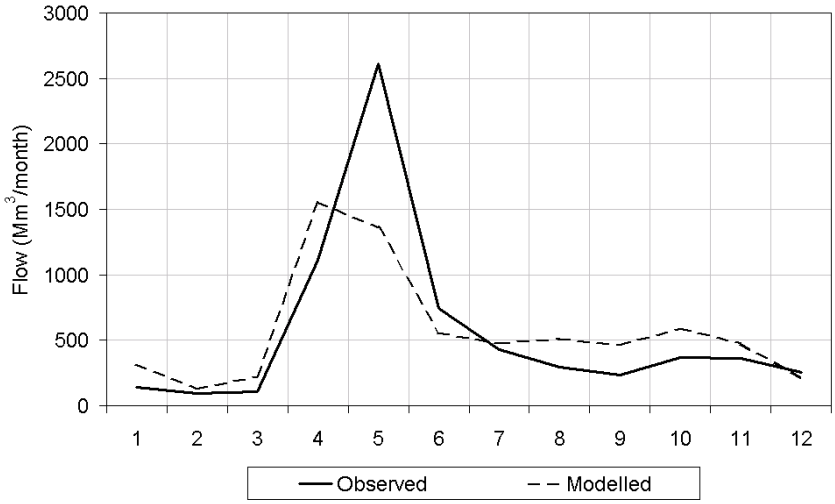


Figure 12 Volga basin model calibration results

(g) Vetluga



(h) Volgograd

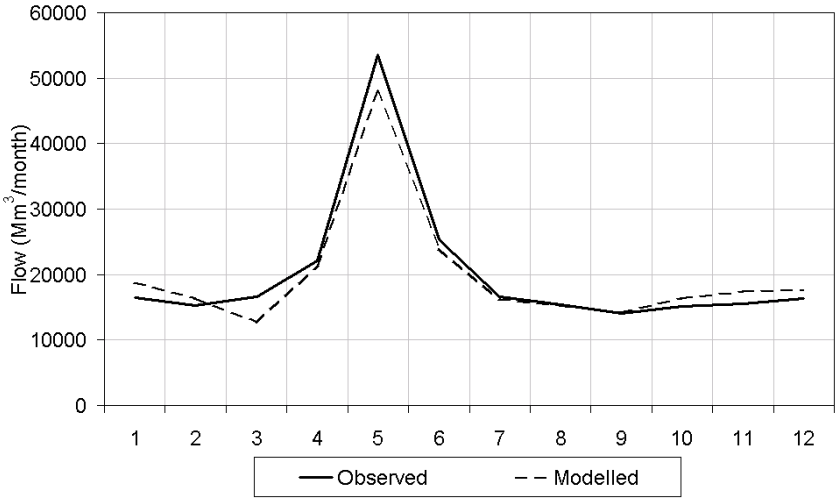


Figure 12. Volga basin model calibration results

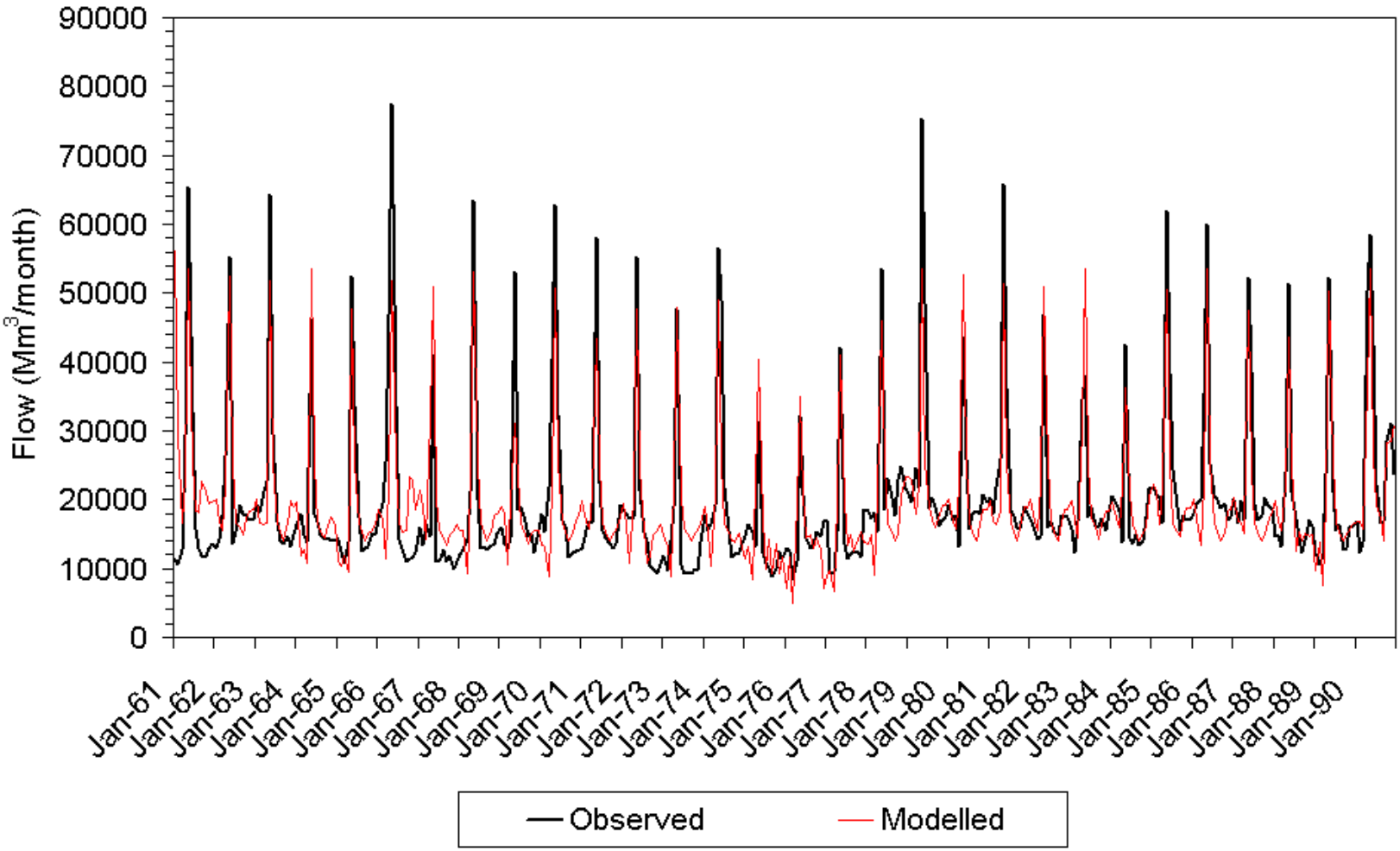


Figure 13. Results of time-series modelling of the river Volga at Volgograd for the baseline period

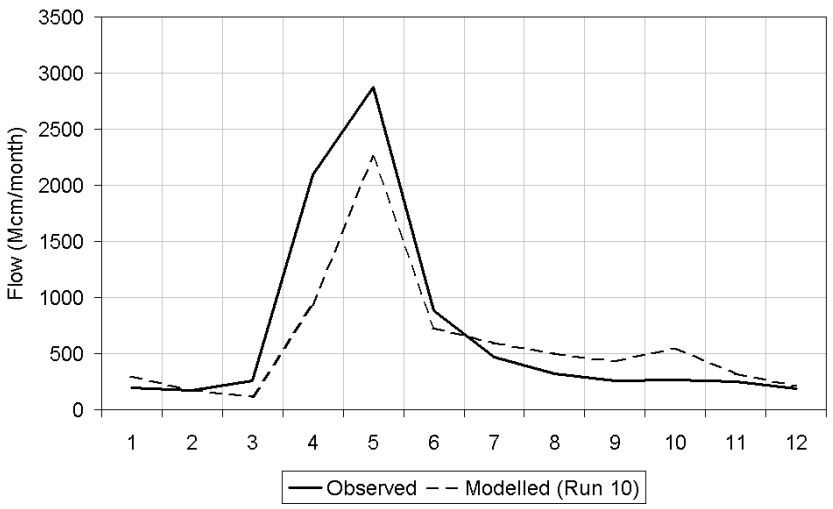
3.1.2. Ural basin

Results of the modelling of baseline flows at each of the Ural catchment calibration gauges can be seen in Table 6. Comparisons of the mean monthly hydrographs are shown in Figure 14, and the time-series of the Ural flows can be seen in Figure 15. 'Modelled' results refer to the best model calibration run in each case.

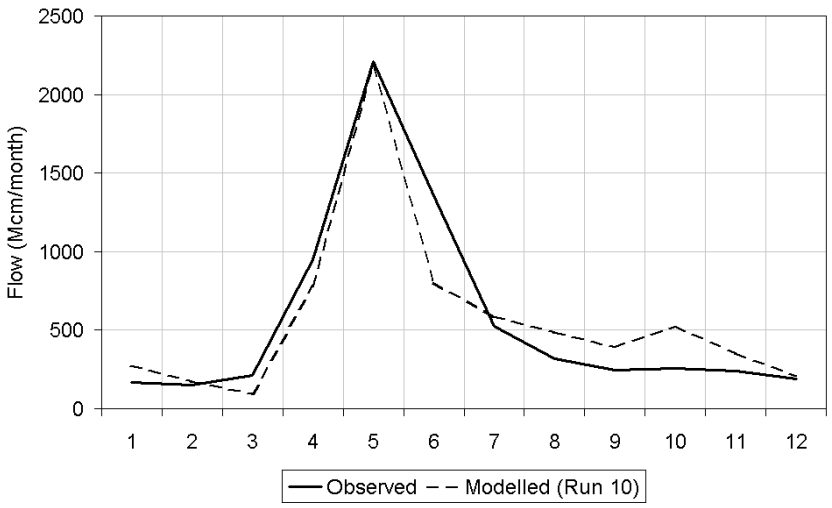
Gauge site	**Mean annual total flow (Mm^3)**		
	Observed	**Modelled**	**Difference**
Kushum	8234	7112	-13.6 %
Ural to Sea	6821	6849	+0.4 %

Table 6. Results of calibrating mean annual total volumes at gauges in the Ural basin

The important result is the Ural inflow to the Sea, while the other location at Kushum is upstream of this. It can be seen from Table 7 that the runoff is generated upstream of Kushum, after which it declines due to losses while flowing through an arid area. The annual volume of flow for the basin is extremely well matched (with a mean annual flow of about 6,800 million m^3 this is less than 3% of the Volga flow). The monthly pattern and time series are also reasonably good.



(a) Kushum



(b) Ural inflow to Caspian Sea

Figure 14. Ural basin model calibration results

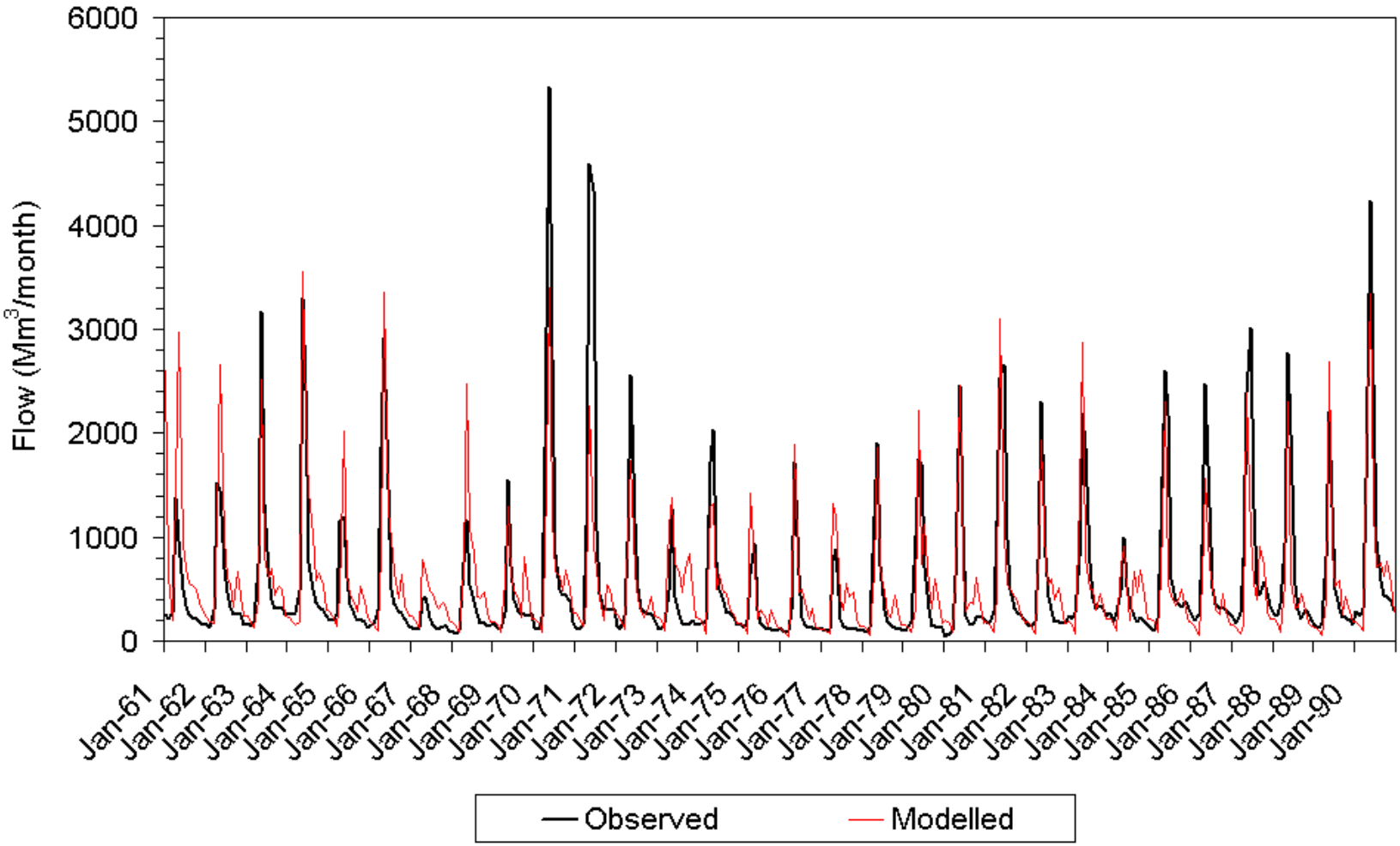


Figure 15. Results of time-series modelling of the river Ural inflows to the Caspian Sea, for the baseline period

3.1.3. South-West basin

Results of the modelling of baseline flows at each of the South-West catchment calibration gauges can be seen in Table 7. Comparisons of the mean monthly hydrographs at key points in the Kura basin are shown in Figure 16 (a-d), and the time-series of the river Kura flows can be seen in Figure 17. 'Modelled' results refer to the best model calibration run in each case.

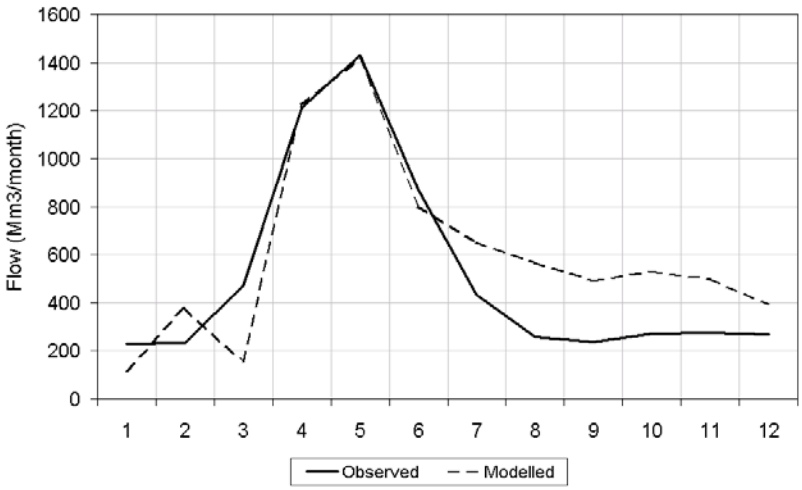
Gauge site	Mean annual total flow (Mm^3)		
	Observed	Modelled	Difference
Terek to Sea	6258	8427	+35 %
Sulak to Sea	4251	6401	+50 %
Samur to Sea	1496	2989	+100 %
Kura to Sea	15686	14527	-7 %
Arax at Gazgalazy	7448	7134	-4 %
Alazani at Kura	3481	3270	-6 %
Small rivers of Iranian coast	10144	8392	-17 %
Small rivers of West coast	1031	3400	+230 %
Kura at Tbilisi	6178	7222	+17 %
South-West regional total	38866	44136	+13 %

Table 7. Results of calibrating mean annual total volumes at gauges in the South-West basin

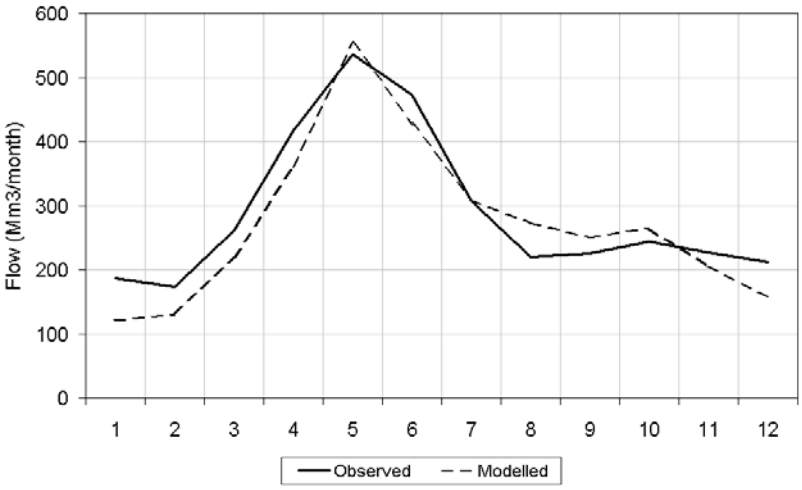
This part of the Caspian Sea basin is made up of a few moderately large catchments flowing into the Sea (the Terek, Sulak, Samur and Kura; by far the most important of these is the Kura); plus many more smaller catchments which have been grouped as "small rivers of Iranian coast" and "small rivers of West coast". The remaining locations for which calibration data have been examined are sub-catchments of the Kura basin (Arax at Gazgalazy, Alazani at Kura, and Kura at Tbilisi). The total observed flow of all the catchments of this part of the basin (that is all the inflows besides the Volga and Ural) is 38,866 million m^3, which is 16 % of the Volga flow.

Generally the larger catchments are reasonably well modelled, with adequate matches to the mean annual flow volumes, and reasonable simulation of the monthly patterns. Some of the small catchments are less good, but as before these errors are not significant in relation to the total inflow to the Sea. Although there are many more large dams in the Volga, some of

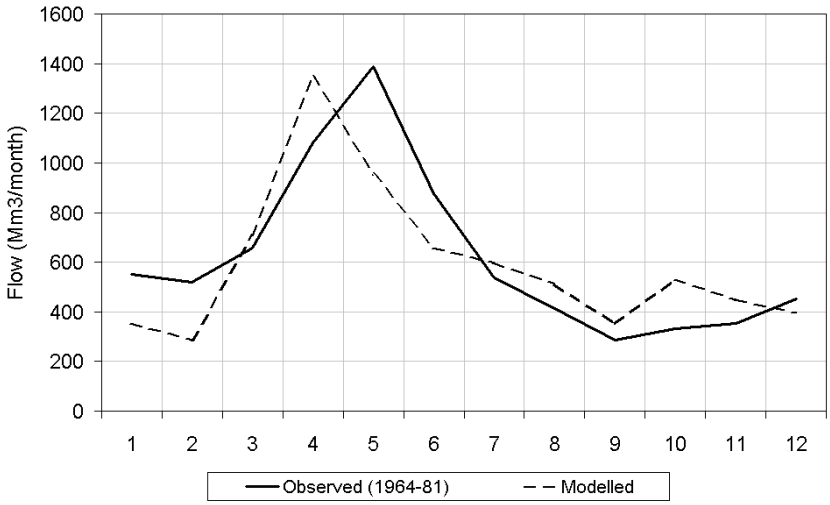
the catchments in the South-West region, especially the Kura basin, are in fact more heavily controlled and have larger abstractions in proportion to their annual runoff. Unfortunately the quality of the available data on these abstractions was not good. In most cases and for most of the largest abstractions, only annual quantities were available. When this was the case, it was necessary to assume in the model that the abstraction quantities were uniform through the year, although that might not be so, especially in the case of irrigation. More detailed information on the monthly patterns of abstractions were only available for some of the smaller abstraction points. The results of this inadequate information can be seen in the Kura basin. For upstream locations such as the Kura at Tbilisi and Alazani inflow to Kura (Figure 16a-b) the hydrographs are reasonably good, but for the Kura inflow to the Sea (Figure 16d) which is downstream of the major abstractions, the hydrograph is much less well modelled, although the annual volume is good. It is likely that this is largely a result of the insufficiently detailed abstraction data.



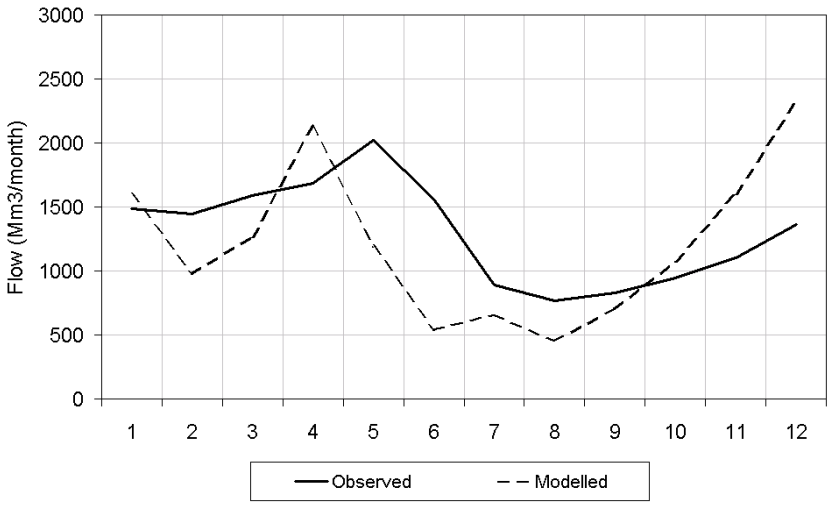
(a) Kura at Tbilisi



(b) Alazani inflow to Kura



(c) Arax at Gazgalazy



(d) Kura inflow to Caspian Sea

Figure 16. South-West basin model calibration results

3.1.4. Summary of calibration results

In summary, it can be concluded that, overall, the inflows to the Caspian Sea are accurately modelled, with good matching of the annual flow volume, which is the most important factor. The seasonal pattern of flow and the year-to-year variability are also reproduced adequately. For some of the smaller sub-catchments in the basin, the calibration is less good, as discussed in more detail above, but these are insignificant in relation to the total flows into the Sea.

A summary of the modelled and observed inflows to the Caspian Sea, for the baseline period, is presented in Table 8. This shows that for each of the three sub-basins, the modelled flow volumes are good, combining to give a very good modelled mean annual total inflow to the Caspian Sea (less than 1% different from the observed inflow over the same period).

Basin	**Mean annual total flow (Mm^3)**		
	Observed	**Modelled**	**Difference**
Volga	242154	238310	-1.6 %
Ural	6821	6849	+0.4 %
South-West	38866	44136	+13.0 %
Caspian Sea (total)	287841	289295	+0.5 %

Table 8. Summary of inflows to the Caspian Sea

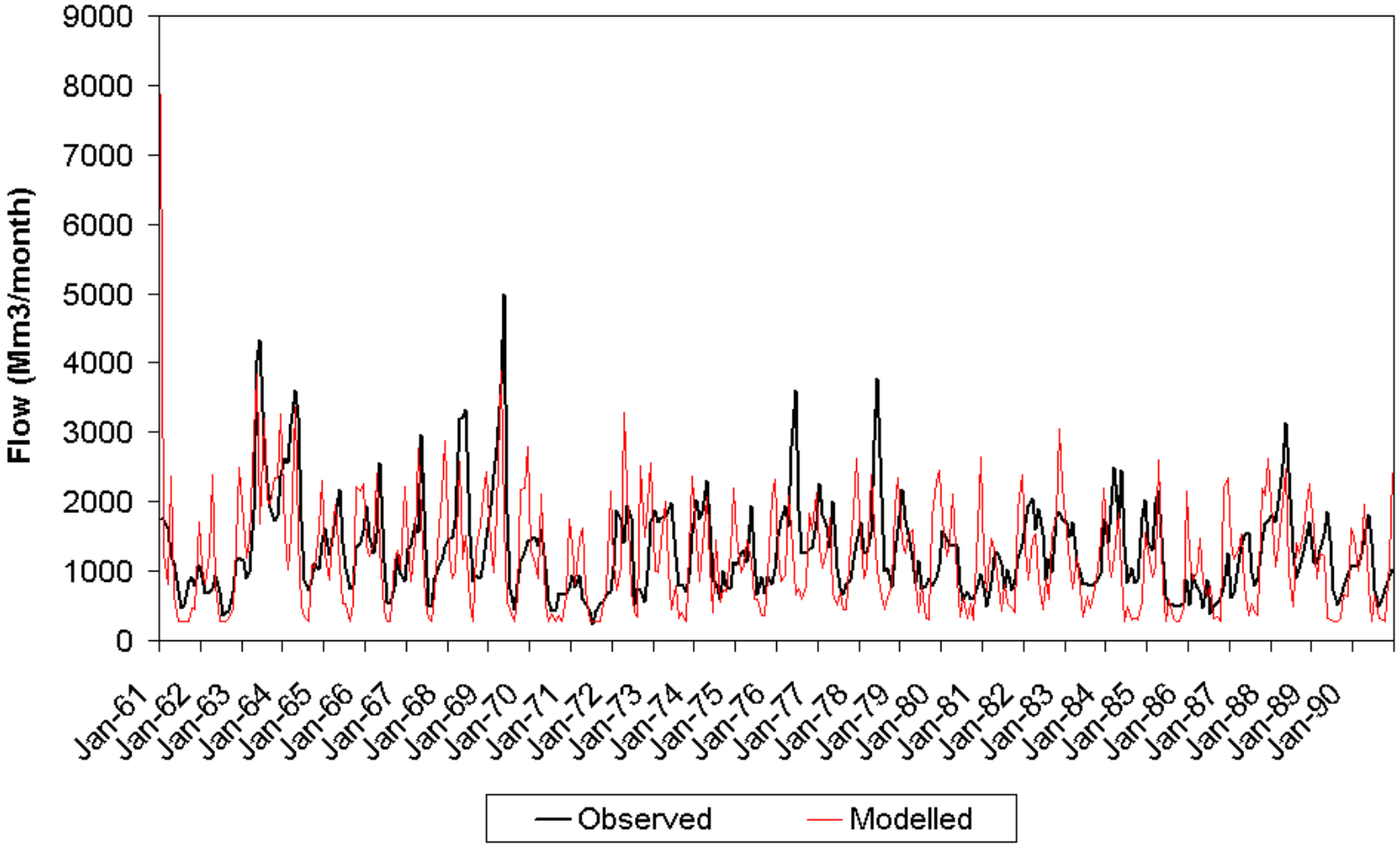


Figure 17. Results of time-series modelling of the river Kura inflows to the Caspian Sea, for the baseline period

3.2. Scenarios of future Caspian Sea inflows

Possible future inflow series for the Caspian Sea were generated using the calibrated GWAVA model with each of the three climate models and each of the three water demand scenarios described above. The flow series were generated for the 30-year period (2040-69) which is taken to provide a set of scenarios representing conditions in the decade of the 2050s.

The results are summarised in Table 97 and Table 108 and in Figure 18 to Figure 21. Table 9 shows the mean annual flows for each of the major divisions of the basin as well as the total for the whole Caspian Sea for each combination of climate and demand scenarios. The mean annual flow scenarios for the whole basin are illustrated graphically in Figure 18. Table 10 shows the results for the whole Caspian inflow as percentage changes compared to the observed value in the calibration period, 1961-90*. Figure 19 to Figure 21 show the mean monthly flows for each of the scenarios for each of the major sub-basins.

Basin	Demand scenario		
	Low	Medium	High
Climate model: HadCM3			
Volga	232938	223366	219800
Ural	6115	5810	5709
South-West	44491	38475	35600
Total	283544	267651	261109
Climate model: Echam4			
Volga	316257	306961	302991
Ural	9183	8878	8777
South-West	49970	43674	40487
Total	375410	359513	352255
Climate model: CGCM1			
Volga	344613	336523	333217
Ural	13252	12948	12847
South-West	50038	43698	40459
Total	407903	393169	386523

Table 9. Scenarios of future mean annual Caspian Sea inflows (Mm3/year)

The different scenarios produce a wide range of results which are in line with the differences in GCM predictions of rainfall discussed in Section 2.4.1. The resulting flows are more

* Observed inflow to the Caspian Sea for 1961-90 was 287,841 million m^3, and the modelled flow for the same period (289,295 million m^3) is extremely similar, differing by only 0.5%.

variable between GCMs than are the rainfalls. This is as expected because runoff is typically sensitive to small changes in rainfall inputs. The lowest change in rainfall (+2.5% with HadCM3) leads to a decrease of 1.5% in flow even with the low demand scenario (which is less than the demand in the calibration period), illustrating the impact that increased evaporation will have, causing higher water losses throughout the basin.

The plots of monthly distributions of flow show clearly the impact of increasing temperature, with earlier snowmelt bringing the peak flow season forward from May to April across the whole basin. For the Volga basin, the plots also show the impact of the high level of control of flow by the cascade of reservoirs, with the result that flows are very similar from May through to September for all scenarios, but differ greatly at other times of year. The reservoir operation has been kept the same while running the scenarios, but in reality it might be expected that operation would be changed in response to the changing flow regimes. However there is insufficient information to predict how this might happen.

Generally, the plots show that the different demand scenarios are of little significance compared to the potential impacts of climate changes. Only in the South-West basin are demands sufficiently large in relation to resources for the potential changes to make a noticeable difference to the future scenarios.

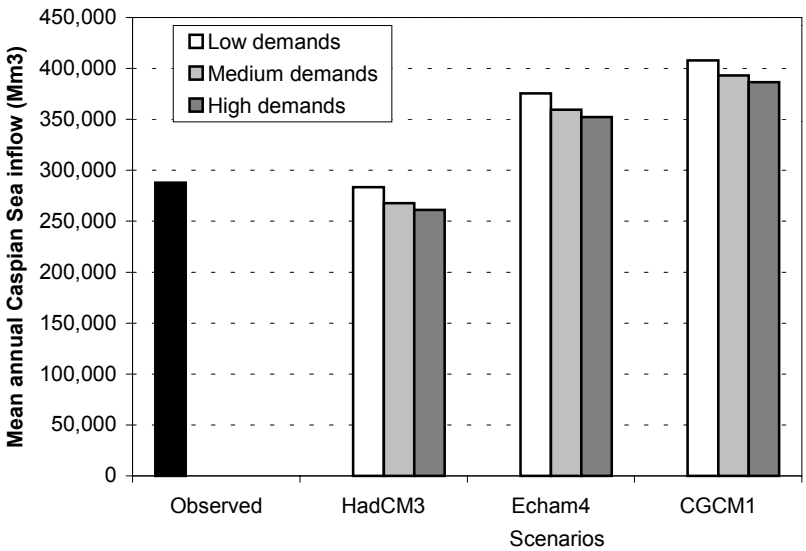


Figure 18. Comparison of observed and scenario inflows

Climate model	**Demand scenario**		
	Low	**Medium**	**High**
HadCM3	-1.5 %	-7.0 %	-9.3%
Echam4	+30.4 %	+24.9 %	+22.4%
CGCM1	+41.7 %	+36.6 %	+34.3%

Table 10. Percentage change in future mean annual Caspian Sea inflows for various scenarios (compared to observed value in calibration period, 1961-90)

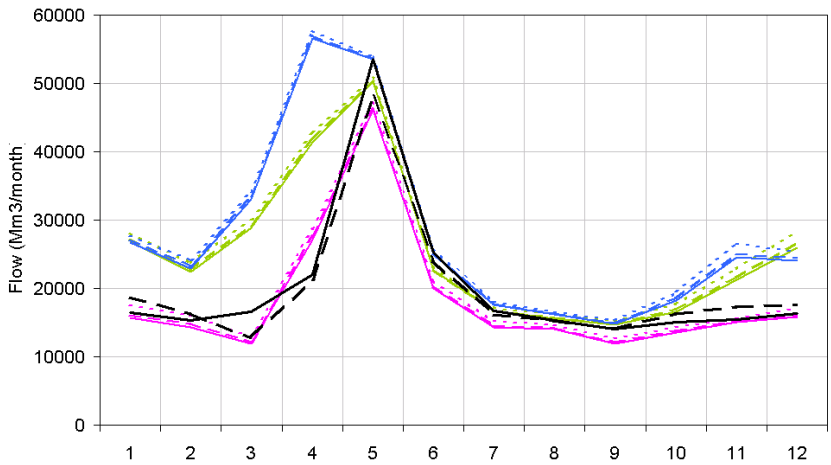


Figure 19. Mean monthly flows for 2050 scenarios – Volga basin at Volgograd

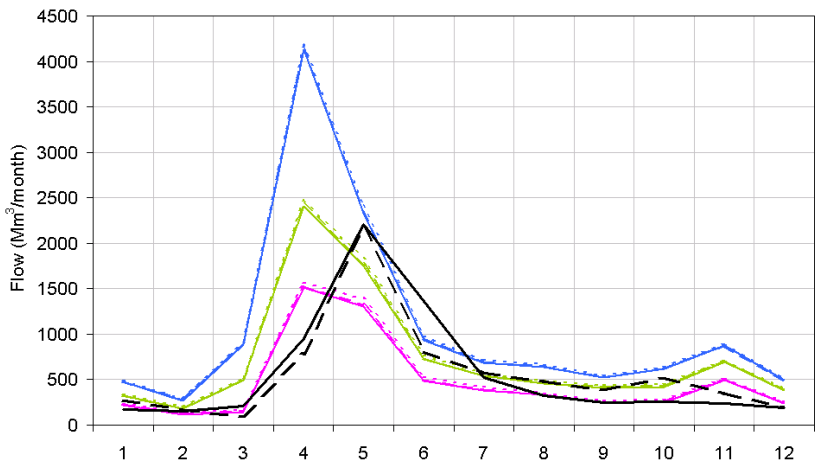


Figure 20. Mean monthly flows for 2050 scenarios –Ural basin inflows to Sea

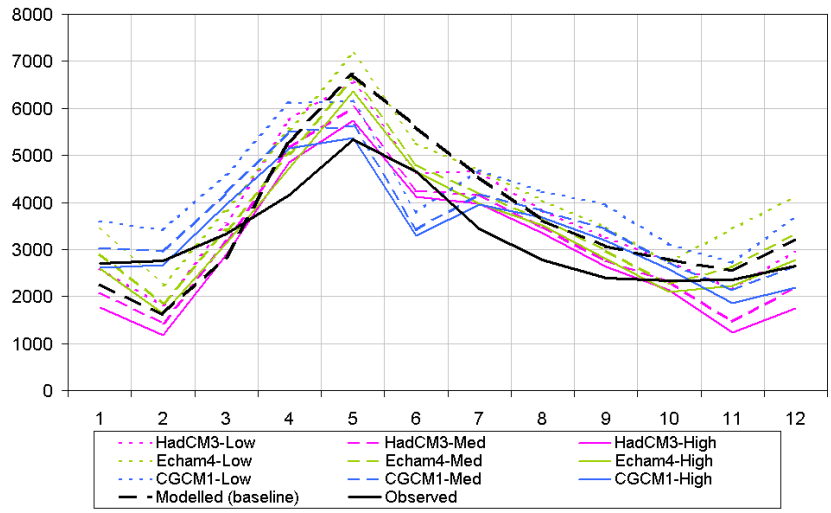


Figure 21. Mean monthly flows for 2050 scenarios – South-West basin inflows to Sea

Chapter 4. Conclusions and recommendations

Overall, the inflows to the Caspian Sea are accurately modelled, with good matching of the most important factor: the annual flow volume. The seasonal pattern of flow and the year-to-year variability are also reproduced adequately. For some of the smaller sub-catchments in the basin, the calibration is less good, but these are insignificant in relation to the total flows into the Sea. Compared to previous phases of the study, the calibration results are much improved, and this is primarily a result of the improved reservoir operation and glacier modules which have been included. One source of remaining inadequacies in model performance is likely to be the lack of detailed abstraction data that reflect seasonal patterns of water use; this is only of real significance in the South-Western parts of the basin.

The scenarios of future inflows show that there is a very high degree of uncertainty of future flows into the Caspian Sea, with a range from a decrease of 9% to an increase of 42% for the most extreme combinations of demand and climate possibilities that were examined. The greatest part of this uncertainty is due to the difference in rainfall generated by the different climate models. As discussed above (Section 2.4.1), it is not possible to take any one GCM result as better or more reliable than the others, and the only reasonable approach is to treat the three GCMs as equally valid. Thus, we can conclude, that with the present state of knowledge in global climate modelling, future inflows to the Caspian Sea are highly uncertain. Climate change is clearly the dominant factor compared to demand changes, and (except in the South-West region) expected changes in water use have small impacts in relation to the range of results produced by the climate uncertainty.

The climate models agree that there will be a substantial rise in temperature, and this impact is clearly shown in the results, with the peak runoff expected to occur in April rather than May as at present, and increased evaporation losses (which reduce to some degree the effects of increased rainfall).

The modelling work carried out in this study has included contributions from experts in all the basin countries, and the model, its input data and the results are being provided to all five countries. This is one step towards increasing co-operation in water resources across the basin.

There are a number of activities and approaches which could be taken to provide further improvements in modelling the Caspian Sea basin, these include:

- The need to obtain improved water abstraction data. Inadequacies in these data have led to problems in model calibration in the South-West of the basin. An intensive study to obtain or estimate past water abstractions, including the seasonal pattern and changes from year to year, is needed.
- The climate model inputs are derived from very large grid cells (of the order of 3 by 3 degrees), but regional GCMs are now starting to become available which model the specifics of a system in greater detail. The possibility of applying regional climate

models to the Caspian basin should be investigated. This may lead to some resolution of the wide divergence of the currently available GCM results.

- Since the Caspian Sea and its basin is a dynamic system and the response of sea levels to changes in input takes many years to reach equilibrium, a dynamic modelling system which combines all elements of the system (catchment, delta and water balance models) to model its response continuously from present conditions into the future will provide more realistic and meaningful capabilities.
- Improved calibration of the model could be obtained by considering a larger number of basin subdivisions. This would allow each part of the basin to be calibrated separately to reflect local conditions in more detail, but this would be a major exercise.
- The use of a smaller grid size, especially in the mountainous and southern parts of the basin where the 0.5 degree grid cell is inadequate to represent the real processes, could provide improved performance (although this has been partly solved by the glacier model used for high mountain cells).
- More involvement of scientists from the basin should be included in future phases of the work. This could help especially by bringing specialist knowledge on water use and hydrological problems for particular parts of the basin.
- Improvements of the rainfall-runoff model to include soil freezing (see note at the end of Section 2.2.2) should be considered.
- In the future, it may be necessary to look at a number of factors in more detail. These include: more detailed scenarios of demands that are tied to specific economic predictions; examination of how land use change could affect hydrological processes; and consideration of how reservoir operation might be changed in response to changing flows.

Chapter 5. References

INCO-COPERNICUS (2000) *Hydrological and hydrodynamical studies and predictions of the Caspian Sea water level rise: Impact of climate factors and man's activities*. The CASSEAS Project. EU INCO-COPERNICUS Contract No. IC15-CT97-0101.

IPCC (1995) *Climate Change 1995: The Science of Climate Change*. Second Assessment Report. J.T. Houghton, L.G. Meira Filho, B.A. Callender, N. Harris, A. Kattenberg and K. Maskell (Eds.). Cambridge University Press, UK. 572pp.

IPCC (2001a) *Climate change 2001: The Scientific Basis*. Working Group I contribution to the IPCC Third Assessment Report. WMO/UNEP.

IPCC (2001b) *Climate change 2001: Impacts, Adaptation and Vulnerability*. Working Group II contribution to the IPCC Third Assessment Report. WMO/UNEP.

IPCC-TGCIA (1999) *Guidelines on the use of scenario data for climate impact and adaptation assessment.* Version 1. Prepared by Carter, T.R., M. Hulme, and M. Lal. Intergovernmental Panel on Climate Change, Task Group on Scenarios for Climate Impact Assessment, 69pp.

Meigh J. R., McKenzie A. A., Austin B. N., Bradford R. B. and Reynard N. S. (1998) *Assessment of global water resources – Phase II, Estimates of present and future water availability for Eastern and Southern Africa*. Institute of Hydrology / Department for International Development, 104pp.

Meigh J. R., McKenzie A. A. and Sene K. J. (1999) A grid-based approach to water scarcity estimates for eastern and southern Africa. *Water Resources Management*, **13**, 85-115.

New, M., Hulme, M. and Jones, P. (1999) Representing twentieth-century space-time climate variability. Part I: Development of a 1961-90 mean monthly terrestrial climatology. *Journal of Climate*, **12**, 829-856.

NOAA-EPA Global Ecosystems Database Project (1992) *Global Ecosystems Database Version 1.0*. User's Guide, Documentation, Reprints, and Digital Data on CD-ROM. USDOC/NOAA National Geophysical Data Center, Boulder, Colorado, USA.

State Water Inventory (1987) *Long term data on the regime of surface water resources*. Volume I, Issue 26, Leningrad Hydrometeorological Data, Russian Federation.

Tate, E. L. and Meigh, J. R. (1999) *Grid-based model of the Caspian Sea basin – Phase I report*. Report to EU-TACIS. Institute of Hydrology, Wallingford, UK.

Tate, E. L. (2000) *Grid-based model of the Caspian Sea basin – Intermediate phase report*. Report to EU-TACIS. Centre for Ecology and Hydrology, Wallingford, UK.

Annex A:
Modelling Work in Phase I
and
Intermediate Phase

A.1. Work carried out in Phase I

A.1.1. Model development

The modelling approach developed for the Caspian Sea basin is based on work aimed at examining water resources availability on a world-wide scale (Meigh *et al*., 1998). Previous studies had generally been rather simplistic, and the new approach was designed to provide a more realistic and consistent methodology. It incorporated the following:

- A consistent methodology which can be applied across all countries and regions.
- Use of the drainage basin as the logical unit for examining resources. This overcomes problems of transfers between countries sharing the same basin.
- Seasonal and year-to-year variability in the surface water flows are important when looking at water availability, and the project included consideration of these factors to assess the water which is actually available for use. Variability in water demands can also be significant.

The approach was adopted for use in the Caspian Sea basin, with the ultimate aim of examining the probable effects of variability in the basin's water resources on the level of the Sea.

The overall approach to estimating water resources availability is to examine surface flows, groundwater yield and water demands on a gridded basis. By using a grid to cover the study area, the regional variability of supply and demand can be demonstrated.

The grid size chosen for the study is 0.5° by 0.5°. This choice is a compromise between that needed to represent spatial variability and the availability of suitable data. It is believed that the 0.5° grid allows spatial variability to be represented to a reasonable level; a coarser grid would begin to lose adequate representation of variability, while a finer resolution would require excessively large amounts of data and greatly increased computational effort. Another reason for choosing a 0.5° grid is that several global data sets are now available at this resolution, making the application of the approach on a global scale possible. Although some types of data are available at finer resolution, most are not, and it would be an onerous task to generate the additional gridded data sets that would be required from the raw information. The region studied is shown in Figure A22.

Estimation of surface flows Based on the 0.5° grid discussed above, surface runoff is generated for each cell by using a rainfall-runoff model and the flows are then routed through the linked cells to estimate total runoff for each. The effects of lakes and wetlands, water abstractions, return flows, artificial water transfers, flow routing and transmission losses are also taken into account. The basic time step of the model is monthly, and a 30-year time series of flows is generated. This period is actually run twice, the first thirty years providing a 'warm-up' to ensure that model behaviour is stable, and the second period of 30 years being used to generate the final flow sequences.

The preliminary stage, before generation of flows, is to determine the main direction of flow for each cell. It was assumed that all the flow from one cell would flow into one of the adjoining cells. Based on these directions, the order in which the cells must be processed is determined so that the flows from upstream cells have always been calculated before processing the cell which they flow into. The processing order also takes into account the artificial transfers between cells.

In an earlier study, the possibility of allowing more than one flow direction to be assigned to each cell was examined in an attempt to improve the representation of catchment boundaries within the gridded system. However, it was found this would lead to difficulties in defining the order of processing the cells, and it was judged that the greatly increased complexity of the model would not be worthwhile for the relatively minor improvements in results which would be expected. For the larger catchments, which are of greatest interest in a global-scale study such as this, representation of the catchments by the grid cells is reasonably accurate. For improved representation of smaller catchments, finer definition of the input data would be needed so that the whole modelling process could operate on a smaller grid size. With the current level of data availability, a smaller grid size is not realistic at present.

Figure A22: A map of the Caspian Sea region

Having defined the flow directions and processing order, the main steps in the procedure for estimating surface runoff are then as described in the following sections. They are also illustrated in a flow diagram in Meigh *et al*. (1998).

Runoff is generated for each cell using a rainfall-runoff model; the model chosen was the probability-distributed model (PDM) developed by Moore (1985). This is a conceptual model, based on physical processes, which takes precipitation and potential evaporation (PE) as inputs. The available input data were monthly mean precipitation, evaporation and temperature values for each 0.5° by 0.5° cell, and anomaly series for the period 1961 to 1990.

Generation of local runoff The PDM controls the conversion of precipitation to runoff. Three additional elements were added to this basic model. The first is an interception model to take account of precipitation that is intercepted by the forest canopy in areas where there is significant forest cover. The second addition is an element to make the model response more realistic in arid areas where there is no flow in the dry season. The third addition enables the PDM to be applicable under snow-fall conditions. The PDM itself and these additional elements are described below. The choice of this model follows work at the Institute of Hydrology and the University of Southampton examining the effects of climate change on river flows in the Britain (Arnell and Reynard, 1996; Arnell, 1996b), Europe (Arnell, 1996a), and Southern Africa (Reynard, 1996). A number of models were investigated and the PDM was chosen for all these studies. It was used on a catchment basis for the work in Britain, and on a 0.5° gridded basis elsewhere. The use of the model over a wide range of environments and climates in these studies indicates that it should also be adequate for the study region in the Caspian Sea basin. The PDM is based on physical processes, but without requiring a large number of parameters, and thus it is suitable for investigating the changes in runoff expected due to climate change.

The PDM rainfall-runoff model The PDM works by partitioning the precipitation input in two parts: one part leads to direct runoff which is a rapid response to precipitation, while the second part leads to the baseflow component of runoff, a much slower response to the precipitation input. These two runoff elements are summed to give the total runoff from the catchment, which is taken to be the grid cell in this study.

The model is shown schematically in Meigh *et al.* (1998). The precipitation input enters a soil moisture store for which a maximum capacity is defined. No direct runoff occurs until the capacity of this store has been satisfied. For any time interval (in this study a daily interval was used), precipitation is added to the soil moisture store, and it is depleted by evaporation and drainage to groundwater recharge. Actual evaporation is assumed to occur at the potential rate until field capacity is reached, below which the ratio of actual evaporation to potential evaporation declines linearly to zero. If the soil storage at the end of the day is less than the maximum capacity, no direct runoff occurs, but groundwater recharge takes place at a rate dependent on the current soil moisture storage. Direct runoff only occurs when the maximum soil moisture storage capacity is exceeded. However, in a natural catchment, not all parts would have the same soil storage properties: steep, upper slopes tend to have smaller capacities than the deeper, cultivated soils of the lower valleys. The PDM allows for

this by having an array of soil stores which have different capacities. The frequency of occurrence of stores of a given capacity is governed by a probability distribution, defined by the model parameters. This distribution of soil stores accords well with modern hydrological practice whereby runoff processes are thought to be realistically described by the concept of variable contributing areas that respond to changing soil moisture conditions. Thus, under dry antecedent conditions, direct (or surface) runoff is generated only from those parts of the catchment having low soil storage capacities, and precipitation over the remainder of the catchment goes into soil storage. After a period of rain, more and more of the catchment becomes saturated and the proportion of the catchment generating direct surface runoff increases.

The direct runoff is next routed through a store in order to distribute the runoff volume in time. This is done by means of a cascade of two linear reservoirs. A linear reservoir has the property that the outflow, q_s, is a linear function of the surface storage, S_s:

$$q_s = s_{rout} S_s$$

where s_{rout} is the surface routing parameter. The outflow from the first linear reservoir provides the input to the second which has the same routing parameter, and the outflow from this is the final estimate of direct runoff from the model.

That part of the precipitation which is neither lost to evaporation nor goes to direct runoff is retained in the soil moisture store. When the contents of this store exceed the field capacity (which is defined as another of the model parameters), this is treated as drainage to groundwater recharge and enters the groundwater store. The outflow from this store provides the baseflow component of runoff. This is computed by means of a non-linear reservoir which has the form:

$$q_b = g_{rout} S_b^3$$

where q_b is the baseflow, S_b is the groundwater storage, and g_{rout} is the baseflow routing parameter. The surface and baseflow components are then summed to calculate the total runoff from the catchment, $q_t = q_s + q_b$.

The PDM runs on a daily time step, but the input precipitation and evaporation data are only available as monthly values. To overcome this, the average numbers of rain-days for each month in each cell were used to estimate daily precipitation, with the monthly total divided uniformly over the rain-days which are randomly distributed within the month. A number of other methods of deriving daily precipitation have been examined in an earlier study, but the results were not strongly sensitive to this choice, and so the simple approach of allocating the same amount to each rain-day was chosen. The monthly evaporation values were distributed uniformly over each day in the month.

From the daily precipitation input, the PDM generates daily flows and these are then summed to obtain monthly flows; from this point onwards the overall model run is on a monthly basis only.

Parameters of the PDM The parameters of the PDM, as described above, are:

c_{max} – maximum capacity of the soil moisture store (saturation capacity),

f_c – field capacity of the soil,

b – parameter describing the frequency of occurrence of the soil moisture stores,

s_{rout} – surface routing parameter,

g_{rout} – baseflow routing parameter.

It is not practicable to calibrate the model for each grid cell with reference to observed river flow data. Therefore, in order to allow it to be applied over the very large areas which are needed in this study, the parameters where determined from the physical characteristics of the cells. The two most sensitive parameters are c_{max} and f_c, and these were related to estimates of the percentage forest cover and the soil type. The model results are less sensitive to the remaining three parameters and these were held constant throughout the basin. The advantage of this approach to rainfall-runoff modelling is that calibration for individual cells is not necessary as the parameters are derived solely from observed physical characteristics. This means that flows can be generated with a reasonable degree of confidence for each of the very large number of cells needed to provide coverage of large parts of the earth. Conversely, its disadvantage is that the available physical characteristics may not be sufficient to reflect the true variability of hydrological processes across the region. Further detail on the selection of model parameters is given in Meigh *et al*. (1998).

Additional model elements As mentioned above there were three additional elements added to the PDM. The first was a model to account for **interception losses** in areas with significant amounts of forest cover. Following Calder (1990), the intercepted precipitation, p_i, was calculated as:

$$p_i = \gamma (1 - e^{-\delta p})$$

where p is the input precipitation, and γ and δ are the model parameters, which were set at $\gamma = 2$, $\delta = 0.5$ for all cells. For each cell, the proportion covered by forest was estimated, and the interception model was only applied in modelling the runoff from this proportion of the cell. Also, in modelling areas covered with forest, the potential evaporation was assumed to be 10% greater than in other areas, in order to account for the higher rates of evaporation which are typically found in forest areas compared to grass cover.

The second additional model element was in the **groundwater** part of the PDM. It was found that the model tends to generate small (but not negligible) flows even in long periods without precipitation. While this is not a problem in humid regions, in arid regions it was found to produce unrealistic flows during the dry season. Unless there is a spring source providing a continual baseflow, river flows in such regions tend to drop to zero for much of the dry season. Experimentation with the parameters of the PDM showed that this problem could not be resolved with the standard model. Therefore, an additional element was

added, consisting of a small drainage from the groundwater store (illustrated in Meigh *et al*. (1998)). The amount of this drainage would be expected to vary depending on the soil type, with higher rates in coarser grained soils, and on the climate, with the drainage being more significant in drier regions. A model of the following form was selected:

$$d = C_d (1 - P/P_d) S_b \qquad \text{for } P < P_d$$

where d is the rate of drainage from the groundwater store, C_d is the drainage parameter, P is the mean annual precipitation for the cell, P_d is the limiting value of mean annual precipitation, above which the model has no effect, and S_b is the groundwater storage. P_d was set as a constant for the whole region, while C_d was varied according to soil type.

The third additional element was a **snowmelt module**. Snowmelt is a major component of the hydrograph of the Volga River, which itself provides around 80% of the total inflow to the Caspian Sea. The effect of snowmelt is to provide a delay, sometimes of several months, between precipitation and the resulting melt-waters which contribute to runoff. The Institute of Hydrology has developed several variants of the PDM for modelling snowmelt, primarily for application to small mountain catchments in temperate maritime climates, *e.g*. the uplands of Scotland. Such models, *e.g*. PACK (Moore *et al*., 1999), typically include the following components:

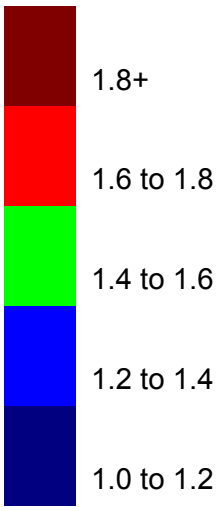
1. A model for partitioning precipitation into rainfall and snow, typically according to the air temperature relative to some user defined 'freezing' temperature;
2. A snowmelt module, at the simplest level assuming melt is proportional to air temperature above freezing, although more complex models can include the effects of wind, heating from precipitation and solar radiation, sometimes with an energy budget calculation;
3. A one-dimensional snow-pack storage model, typically with separate 'stores' for the dry part of the snow-pack (snow which has not yet melted) and the wet part (snow which has melted but which is still in the pack). For the wet store, two drainage rate constants are also specified; allowing faster drainage to occur once a critical storage is exceeded;
4. A component allowing for partial coverage of snow in the region of interest;
5. Allowance for the influence of elevation on temperature via elevation zones or hypsometric curves.

For the Caspian Sea basin application, the available data could not at present support such a detailed level of modelling; in particular items 4 and 5, although the grid-based approach implicitly includes some allowance for these effects in the measured data for the baseline period, *e.g*. the interpolated mean monthly temperature for a grid cell will be based on temperatures measured at nearby meteorological stations at a similar elevation. The State Hydrological Institute (SHI), which has already developed its own snowmelt models for the region, also suggests that the snowmelt module should include an allowance for soil freezing. The following paragraphs describe the main modifications to made to allow for snow and snowmelt in estimating surface runoff to the Caspian Sea.

In applying a temperature-based approach for modelling snowmelt, and for partitioning precipitation into rainfall and snow, monthly mean temperatures need to be disaggregated into daily mean values. In a previous study for western and central Europe using the UK's Climate Research Unit (CRU) data set (Arnell, 1999), daily temperatures were generated by fitting a sine curve to the monthly average values for each cell, then applying a normally distributed random component about this curve, with a user-defined standard deviation of 2°C. For the Caspian Sea application, a similar approach was used; however, since baseline monthly values for 1961-1990 were used, rather than long term averages, the daily values were obtained by linear interpolation between monthly values, assuming that the monthly value applies in the middle of the month, *e.g.* 15th Jan, 14th Feb...., and then applying the random component. For this first attempt, the same standard deviation of 2°C was assumed as that used by Arnell.

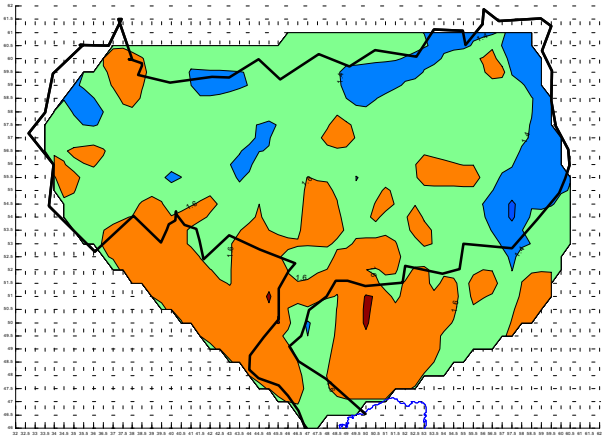
The mean global baseline precipitation data do not account for the significant undercatch in cold regions, so a grid of correction factors was calculated by SHI and applied to the raw precipitation data. This was achieved by taking data from the 132 meteorological stations in the Volga basin, 32 stations in the Ural basin, and 40 stations in the south-western basin region. The correction factors are more significant in the northern part of the Caspian Sea basin, where it is colder and therefore receives more precipitation falling as snow. Each meteorological station has factors to correct for the effects of wind on precipitation falling as snow (Hydrometeorological Publishing House, 1968). These factors were interpolated to the 0.5° grid for the whole Caspian Sea basin, using the Green Function distance method of weighting values, and the factors applied to the raw precipitation data. Maps of the factors for different parts of the basin, where such factors are relevant, are shown in Figure A23 and Figure A24. These are examples for the months of January, February and March.

The key for the precipitation factors in Figure A23 is as follows:

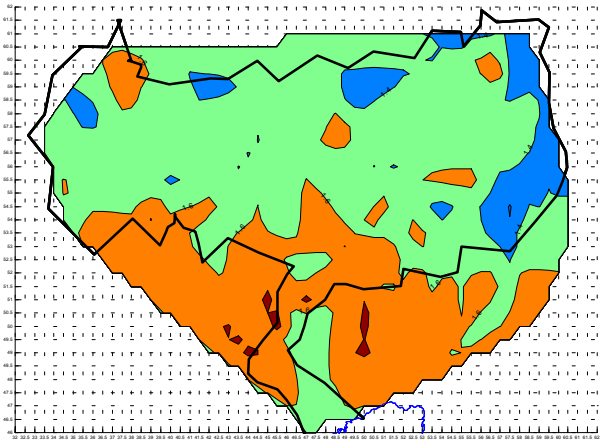


The key for the precipitation factors in Figure A24 is as follows:

a)



b)



c)

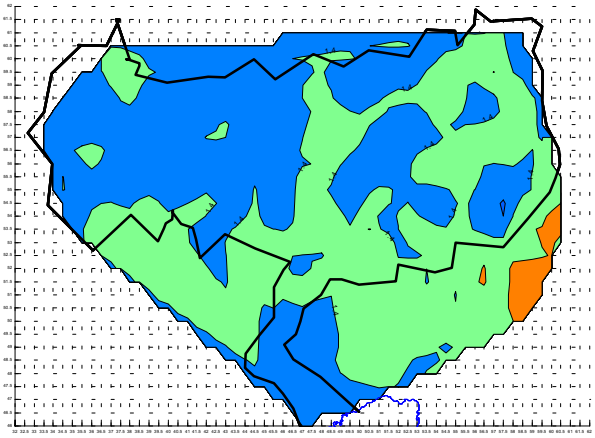
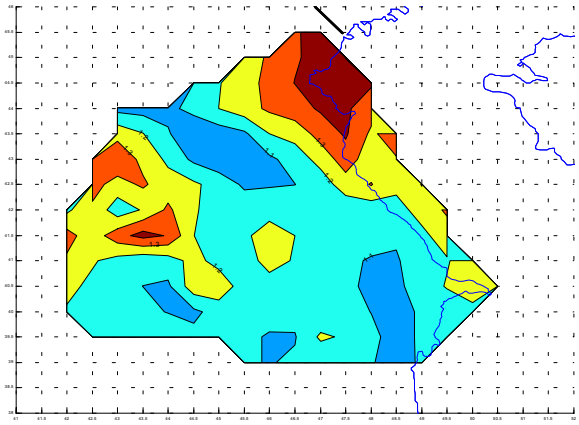
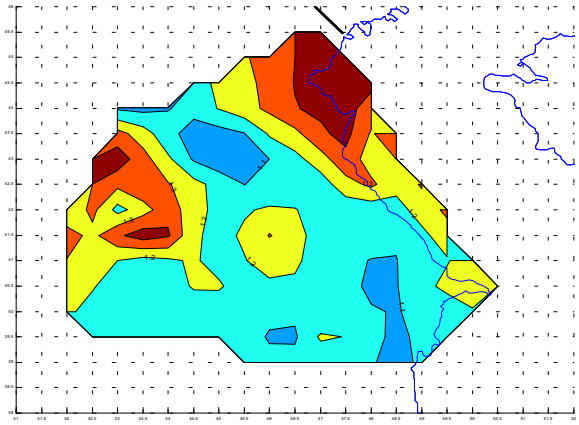


Figure A23: Map of the precipitation factors used in the Volga and Ural sub-basins of the Caspian Sea basin for a) January, b) February and c) March

a)



b)



c)

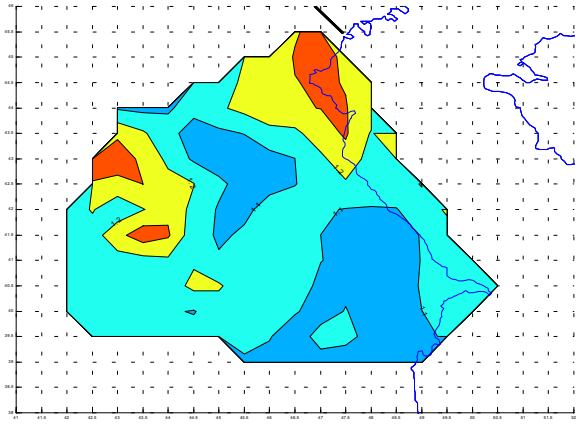


Figure A24: Map of the precipitation factors used in the Kavkaz region of the Caspian Sea basin for a) January, b) February and c) March

The partitioning method used is the usual one of assuming that, above a threshold temperature Ts, all precipitation falls as rainfall, and below that temperature all precipitation

falls as snow. To start with, a typical threshold temperature for the UK of 1°C was used (Moore *et al*., 1999).

The simplest possible melt formulation was used, *i.e*. a temperature excess model. This is normally thought to be most suitable for humid, temperate conditions (warm, moist and windy), when melting occurs primarily from turbulent heat exchange, rather than other mechanisms such as heating from direct radiation or the heat input from falling rainfall. However, the temperature excess approach has also been used successfully in Switzerland and Sweden (*e.g*. Braun and Renner, 1992), and is used by SHI in their operational models. The model simply assumes that, above a threshold melting temperature Tm, the rate of snowmelt is proportional to the temperature excess (T-Tm). The melt factor recommended by SHI for preliminary modelling is 2mm/day/°C with a melting temperature of 0°C.

Having defined the snow inputs to the snow-pack, and the melt within the pack, a simple snow storage model is required to keep account of the current depth of snow-pack in the 'wet' and 'dry' state. For this first attempt, direct evaporation from the snow surface was neglected. The drainage from the pack then forms the input to the PDM. For estimating the drainage, Moore *et al*. (1999) suggest a simple linear outflow relationship, with outflow proportional to storage depth, and with two rate coefficients, allowing slow and rapid drainage. Typical values for these coefficients are 0.15 and 0.85 per day. Rapid drainage is assumed to occur when the storage exceeds a critical fraction of the total depth, say 0.1. Drainage is assumed to stop when the air temperature falls below a critical value, taken to be 0°C. Normally, any new precipitation is added to the wet store but, in the case that the pack only contains wet snow, any new precipitation is added directly to the drainage from this store to avoid the situation of the pack persisting for an unrealistic length of time.

For the Caspian Sea basin application, an additional complication is that infiltration in the basin is reduced and, in some cases, stopped when the depth of frozen soil exceeds certain limits (*e.g*. for sandy, loamy soils, there is a reduction for depths in the range 0.3 to 0.6m, with infiltration stopped at greater depths). Incorporating this effect would require a modification to the PDM itself in the form of a conceptual store for the frozen soil, such that the baseflow component of the model has no input when the depth of frozen soil exceeds a threshold, and is reduced for values less than that threshold.

For calculating the depth of frozen soil, SHI assume that this is proportional to the integrated value of the air temperature from the time the temperature first drops below zero, less an allowance for the insulating effect of any snow-pack above the soil surface. It is assumed that, once melting starts, the snow-pack melts before the soil, and that a fraction of the melt (*e.g*. 20%) infiltrates into the soil helping the soil to melt also.

Incorporating this method into the grid model could be attempted at a later stage of the project. The challenge will be to adjust the PDM to correctly allow for reductions or breaks in infiltration during times when the soil is frozen. Advice from the PDM developers will be required for this step. For the limited time-frame of Phase I, the effects of frozen soil have been ignored, but may be included at a later date if the calibration results suggest that this is required (Sene, 1999).

Summing runoff from upstream cells Having generated the local flows using the PDM as described above, flows are routed between cells and summed to provide the total runoff in each cell. At this stage, water which is consumed in the cells is taken into account, and a loss in flow between the cells can also be included.

The summation is as follows:

$$Q_S = Q_L + \Sigma (1-L)(Q_U - Q_C + Q_R)' - Q_T$$

Where:

Q_S = total runoff for the current cell;

Q_L = locally generated runoff for the current cell;

Q_U = total runoff for an adjoining upstream cell (this was Q_S when that cell was itself the current cell);

Q_C = water consumed in an adjoining upstream cell;

Q_R = return flow from an adjoining upstream cell;

Q_T = artificial transfer of flow out of or back to the current cell;

L = proportional loss for flows out of the upstream cell;

' indicates that flows have been routed; and

Σ refers to the summation over those of the 8 adjoining cells which flow into the current one.

Further descriptions of these parameters can be found later in this report.

Routing between cells and transmission losses As the overall model time step is monthly, flow routing is generally not important except in very large basins where the difference in timing between flows from different parts of the basin could be significant. The method used is the simple Muskingum method which provides a time delay and a degree of dispersion of the flood wave, but does not require any physical data. For situations where there is loss of water in transmission along the rivers through seepage into the banks or bed, a simple proportional loss can be applied to account for it. Transmission losses tend to occur particularly in dry areas where the soil is sandy, and the effects are especially noticeable in large rivers flowing through arid regions.

Artificial water transfers between cells In addition to the natural transfer of water between cells along the river channels, there are also artificial transfers along canals and pipelines, both within basins and between different basins. The term ‘artificial transfers’ is used only to refer to transfers between cells. Two types of transfer were defined. Either the amount to be transferred is a fixed amount, or it is determined as the amount of water to

satisfy the demand in the cell into which the transfer is arriving. In both cases, the amounts are, of course, limited so that they do not exceed the available flow in the cell supplying the transfer in each month of the simulation. Transfers determined by the demand also have a second, fixed limit which corresponds to the capacity of the canal or pipeline. Transfers from outside the basin into the basin (such as the Msta-Tvertsa transfer), or from inside the basin to the outside world (such as the Chusovaya-Iset transfer), can also be specified in this part of the model.

Effects of lakes, reservoirs and wetlands Where there are lakes, reservoirs or wetlands, these can cause considerable alteration to the flow pattern, and therefore they need to be included in the model. In the following most of the discussion refers to lakes, but reservoirs and wetlands can generally both be treated in the same way.

Lakes can be modelled by a simple water balance procedure:

$$S_i = S_{i-1} + Q_{in} + (P-E).Area - Q_{out}$$

where S_i = storage at end of month i; S_{i-1} = storage at end of month i-1; Q_{in} = inflow in month i; Q_{out} = outflow in month i; P, E and Area = precipitation, evaporation and lake surface area, all in month i-1. For lakes and reservoirs open water evaporation is used, while wetlands are assumed to be covered in vegetation which tends to increase the rate of evaporation loss, and so a higher evaporation rate is used. For most cases the relation of the surface area to the storage will not be known, but it is reasonable to assume that they are linearly related. If the storage becomes greater than the capacity of the lake, it spills, and this additional outflow is added to Q_{out}.

The problem with this procedure is that the outflow is not known, but this is what we are in fact most interested in. To overcome this, a simplifying assumption is made by assigning lakes to one of two types. In the first type, which would apply to natural lakes and wetlands and to reservoirs where the main function is to regulate the outflow, it is assumed that the principal effect of the lake is to reduce the variability of flows by increasing dry season flow. This can be done by assuming outflow is a function of net inflow:

$$Q_{out} = Q_{net\text{-}in} (S/S_{max})^{1.5}$$

where $Q_{net\text{-}in}$ = average net inflow in month i = average of Q_{in}+P-E, estimated on a monthly basis; S_{max} = the capacity of the reservoir (S, the storage in the previous month, can be greater than S_{max} for certain cases). Where the size of the lake is small in relation to the annual inflow, the function has little effect, but as the size of the lake increases the outflow becomes more and more smoothed, until with very large lakes, it provides an almost constant outflow.

The second type, which would normally only apply to reservoirs where the main function is to store as much water as possible, the outflow is found by assuming that there is none unless the reservoir spills. The part of the inflow in a particular month which causes storage to exceed the maximum capacity becomes the outflow for that month. In this type there would normally be direct abstractions from the reservoir, and these are included in the water

balance equation. This type has the opposite effect to the first type of lake in that the variability of the flow tends to be increased.

When working on a gridded basis, there are two different situations which need to be considered with regard to lakes:

1. The lake is contained entirely within the cell. In this case the total flow for the cell is routed through the lake, except in cases where there are very small lakes not on the main river; these are ignored. Where there are several lakes in a cell they can be lumped together and treated as one.
2. The lake extends over more than one cell. In this case, the cells which do not include the lake outlet are treated as normal except that when calculating the local runoff only the area of land is used (*i.e.* the area of the cell less the area occupied by the lake). Thus, for cells totally within the water body, local runoff is zero. The lake routing model is run only for the cell which includes the outlet, and this uses the total area and storage of the lake, not just those within the cell.

As noted above, these procedures will generally apply equally to lakes, reservoirs and wetlands. However, in a number of special cases, the capacity and surface area of a lake or wetland may increase significantly in response to a period of substantially increased inflow. For these cases a different formulation of the lake model was used in which the lake (or wetland) size is not limited to the maximum capacity S_{max}, but can adjust to different values if the flows change.

Water consumption and return flows The equation for summation of flows between cells includes water consumed, Q_C, and return flow, Q_R. Water consumed is the water diverted for use (irrigation, water supply, etc.) in the adjoining upstream cell. This is calculated from the demands in that cell (as discussed below), with the proviso that, if demands exceed the available supply, the water consumed is taken as the available supply only. It can be seen that the values specified for demands will affect the model results for surface water availability because the water consumed in an upstream cell reduces the amount of water which flows to the next cell downstream. Demands can be supplied from either surface or groundwater or from a combination of the two, and a method is needed to determine that part of the demand which is to be supplied from surface water. This has been done by setting up some default assumptions, as follows:

- Domestic demands in urban areas are met by 80% surface water, 20% groundwater;
- Domestic (plus livestock) demands in rural areas are met by 30% surface water, 70% groundwater;
- Industrial demands are met by 70% surface water, 30% groundwater;
- Irrigation demands are met by 100% surface water.

These figures have been selected in earlier studies as being those most likely to apply in the majority of situations. For each cell, the model has the capability for these default values to be over-ridden by different figures reflecting the true situation for that cell. However, the

ability to do this is constrained in many cases by lack of adequate data. Normally, the figures for allocation of demand are followed as far as possible, but where the demand from one of the sources exceeds the available supply, the model attempts to satisfy the remainder of the demand from the alternative source. Where water is consumed, there may also be return flows, and the model also accounts for these. For industrial water uses, the return flows may be almost as much as the water abstracted, while for domestic uses they will be a smaller proportion, depending on whether or not there is a sewage system and how effective it is. Irrigation return flows can be assumed to be zero provided the irrigation diversions are calculated as a reasonable estimate of the water demand. In the Caspian Sea basin, data on actual net abstractions (i.e. abstraction minus return flow) were available and used in preference to the estimation of demand from data such as population estimations.

Surface water which is available for use The model has the facility, following the procedures discussed in the previous sections whereby a 30-year sequence of monthly flows is generated for each cell in the basin, to derive estimates of the water which is available for use within the cells. In previous studies, a number of different measures were considered, including:

- The mean flow, or a proportion of it, say 50 %;
- The average of the driest months in each year;
- The driest month in each year which occurs with a certain degree of reliability.

However, in the model of the Caspian Sea basin, the primary concern was prediction of total inflow to the Sea and thus the Sea level, and so estimates of water availability in the basin were not studied here.

Estimation of water demands There are a variety of different water demands which can be considered. They have been divided into categories here as follows:

- Domestic: rural and urban water supply, all uses;
- Industrial;
- Agricultural: irrigation and livestock.

Domestic water supplies are mostly for consumption by people for drinking, cooking and washing, but small industrial demands may sometimes be included here as well. They are estimated from the population in each cell and data for water requirements per head. Urban and rural demands were separated because water use is generally higher in urban areas. An additional category of ‘special’ urban water requirement was introduced to allow for the still higher demands which are expected in capital cities and other very large and important urban areas.

Future domestic water demands can be derived from a combination of scenarios of population growth (which can be assigned separately to rural and urban areas) and changes in per capita requirements (rural and/or urban). In this way, spatial changes in population water demand resulting from, for example, migration to urban areas, can be examined.

Industrial demand refers to large-scale industrial water users which are not included in the rural or urban water supply. There is often a lack of information on industrial water use, but where such demands are known they can be entered into the appropriate grid cell and added to the other demands.

The agricultural sector is often the largest water consumer, and irrigation demands are often the most important. Irrigation schemes vary widely in scale; the main areas need to be identified and, where possible, the following information should be assembled for each scheme: gross and net irrigated area, irrigation efficiency, main crops, and cropping pattern. In contrast to the other types of demand, which can reasonably be assumed to be constant throughout the year, irrigation demands vary through the year depending on the potential evaporation, precipitation, crop type, number of cropping seasons and time of planting of each. Estimates of the demands throughout the year can be made following FAO guidelines for crop water requirements (Doorenbos and Pruitt, 1977) and taking account of effective precipitation (Dastane, 1974).

In the Caspian Sea basin, data on actual abstractions and water use were available and entered directly into the model, meaning that the above estimations of demand were not required.

A1.2 Setting up the model for the Caspian Sea basin

The methodology outlined in the preceding section was applied to the whole of the Caspian Sea basin. The basin includes parts of five countries: Azerbaijan, Iran, Kazakhstan, Russia and Turkmenistan. The basin has a land area of around 3 million km^2, and a total of 1,364 0.5° by 0.5° cells were needed to cover it. This section describes the assembly and treatment of the large amounts of data which were needed.

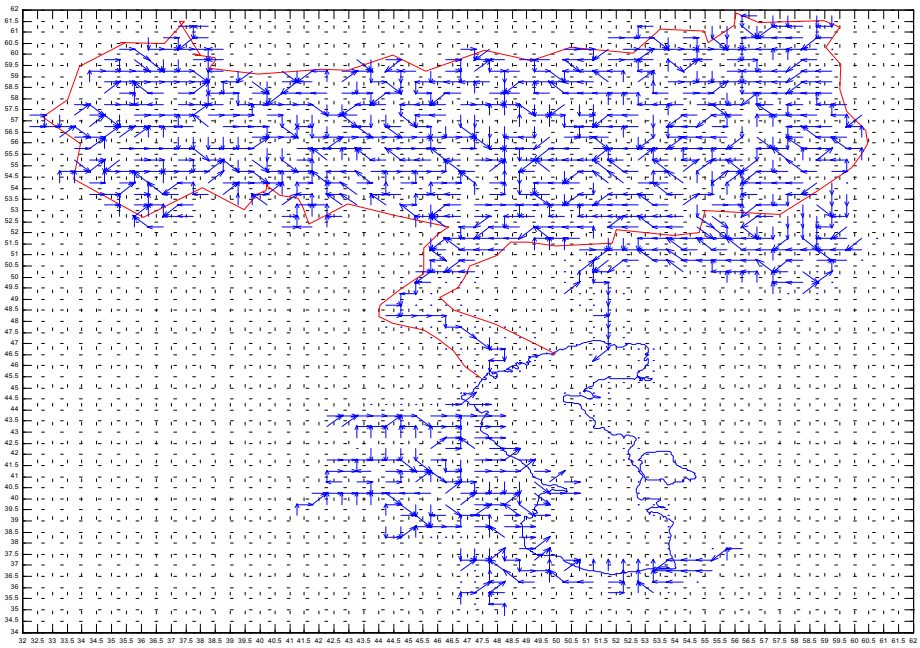
The task of assembling and improving global data sets is largely the responsibility of the UN agencies. The UNEP (Geneva and Nairobi) has developed a Global Resources Information Database (GRID) based on the ARC/INFO GIS system as part of the Global Environmental Monitoring System (GEMS). GRID holds information on population, livestock and many other data. GRID-Geneva have also introduced a customised geo-referenced location and retrieval system of non-GRID held environmental data (WEBSEARCH).

Surface water - data preparation and assumptions This covers the specification of **land areas, cell linkages and artificial transfers.** The area of land in each cell is needed both for modelling and for estimation of demands. These areas were calculated using the latitude of the half degree cells, with the reduced area of cells occurring at the basin boundary being estimated by eye. Where lakes and wetlands occurred, the reduced land area after taking these into account was estimated in a similar manner.

To examine the flow directions and linkages between adjoining grid cells, land elevation data were obtained from the US National Oceanic and Atmospheric Administration/Environmental Protection Agency (NOAA-EPA) Global Ecosystems Database CD-ROM (1992). This

provides minimum, maximum and modal elevations on a regular grid of 10-minute resolution. The minimum and modal elevation data were abstracted and the averages of both calculated for each 0.5° cell. As an initial estimate of the flow direction between cells, it was assumed that the direction for which the difference in average elevation was greatest was the appropriate one. Thus, for each cell, one of the eight possible directions of flow to an adjoining cell could be assigned. In addition, some cells were defined as having no outflow into an adjoining cell. This was done in desert areas where there are no significant rivers, and it was also used to define the downstream end of basins either at the coast or for internal drainage areas. The flow directions can then be compared, cell by cell, with maps of the region (generally at a scale of 1:1 million) to check whether they were the best approximation to reality, and corrected as necessary. By these means, the network of interconnecting cells which most closely approximates the actual drainage basins was determined. Because the region has large areas of rather flat topography, it was found that the elevation data were inadequate to estimate the direction of drainage for many of the cells, and the directions for the cells had to be estimated by eye.

Additional links between cells were set up by defining artificial transfers within and from/to



outside the basin. The final cell linkage network used in the model is shown in **Figure A25**.

Figure A25: **Caspian Sea basin cell linkages**

Mean monthly **precipitation** data were available from the 0.5° gridded data set supplied by the Climate Impacts LINK Project based at the Climatic Research Unit, University of East Anglia, UK. (Precipitation anomaly series for 1961-90, number of rain-days, evaporation and temperature data sets, all on a 0.5° grid, were also supplied by the same source; these data

sets are referred to as CRU data in the following). Mean annual precipitation based on these data is illustrated in Figure A26.

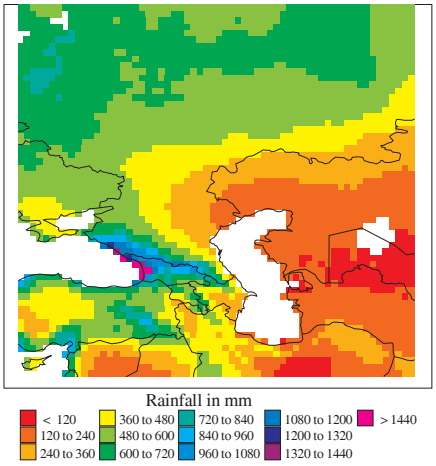


Figure A26: Caspian Sea basin mean annual precipitation (mm) for baseline period 1961-90

In addition to mean monthly precipitation, the model requires long series of monthly precipitation anomalies so that time series of monthly precipitation can be derived for each cell. Anomaly series on a 0.5° grid were available for 1961-90 from CRU.

Values for the average number of **rain-days** in each month are also needed, and these were also provided by CRU, as noted above. Monthly mean **temperature** data and anomalies were also obtained from the same source, and the baseline 1961-90 data are illustrated in Figure A27.

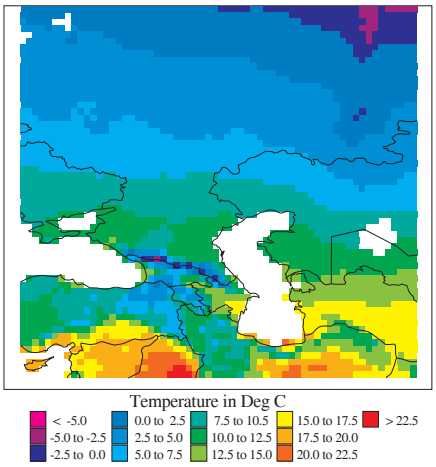


Figure A27: Caspian Sea basin monthly mean temperature (°C) for baseline period 1961-90

Monthly reference crop **potential evaporation** (PE) estimates were made for each cell using the Penman method (Doorenbos and Pruitt, 1977), based on data for temperature, radiation, humidity and wind speed. These input data were provided by CRU, as noted above. Mean annual potential evaporation (reference crop) based on these data is illustrated in Figure A28.

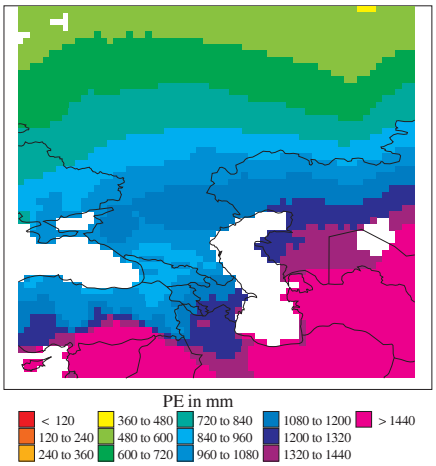


Figure A28: Caspian Sea basin mean annual potential evaporation (mm) for baseline period 1961-90

By choosing different values for the albedo (0.25 for grass, 0.05 for open water), the Penman method can also be used to estimate open water evaporation rates for use in water balances of lakes and reservoirs. It was found that Penman open water evaporation estimates were on average about 20% higher than reference crop, and thus a factor of 1.2 was used in the model to convert reference crop PE for each cell to open water PE. Similarly, a conversion factor is needed for calculating the water balance of wetlands. As these are often covered, at least partially, by vegetation, they tend to have evaporation rates even higher than open water. In taking an overview of studies of potential evaporation of different land types, Savenije and Hall (1995) quote a figure for the PE of swamps or wetlands as 1.5 (or more) times greater than open water (that is, a factor of 1.8 compared to reference crop). Clearly, it is not possible to select a factor for the PE of wetlands which would be reliable for all sites, but a figure of 1.4 (compared to reference crop) was selected as a plausible compromise between the values of 1.2 for open water and the very high value of 1.8 for well-vegetated wetlands.

The modelling process distinguishes between 'forest' and 'grass' when determining evaporation and defining soil moisture characteristics, and some data on **vegetation cover and land use** are needed to estimate the percentage forest cover in each grid cell. The

Olson ecosystem classification was used for this. The data are available at 0.5° by 0.5° resolution on the GED CD-ROM (NOAA-EPA, 1992). The Olson classification was derived from a combination of published vegetation maps, remotely sensed data and observations. An alternative method, suggested by DeFries and Townshend (1994), could also be used in further refinement of the model. This method allows estimation of vegetation type based on the temporal variation of the Normalised Difference Vegetation Index (NDVI) satellite data. It can be applied to the monthly NDVI data which are also available on the GED CD-ROM, and the dominant vegetation type (from broadleaf, evergreen forest to bare ground and cultivated crops) can then be estimated for each cell. With this procedure, results vary somewhat depending on the year chosen for analysis. For the purposes of this study, the vegetation types (as defined by the Olson classification) then need to be simplified to values of percentage forest cover, as set out in Table A11. Land not covered by forest was treated as being grassland.

Table A11: ***Assumed percentage forest cover for Olson ecosystem types***

Olson ecosystem number	Class	Forest cover (%)	Olson ecosystem number	Class	Forest cover (%)
1	Conifer forest	100	10	Forest/field; dry evergreen broadleaf forest	50
2	Broadleaf forest	50	11	Wetlands, swamps	0
3	Mixed forest	50	12	Desert	0
4	Grassland/shrub + tree	0	13	Shrub/tree; succulent/thorn	0
5	Tropical/subtropical forest	100	14	Crop/settlement	0
6	Scrub and woodland or fields	0	15	Conifer snowy rainforest	100
7	Semi-desert	0	19	Mangrove	100
8	Field/woods complex or savannah	25	27	Tundra margins	0
9	Northern boreal taiga forest	25			

In order to define the hydrological model parameters, spatial data on **soil hydrological properties** need to be defined: soil field capacity (the amount of water that can be held in

the soil against gravity) and saturation capacity (the amount of water held in the soil when all pore spaces are full) are required. Soil water holding capacities are primarily related to soil texture. The information on soil texture was obtained from the FAO soil map of the world, which was available in digital format as an ARC/INFO coverage. From this, the dominant soil type in each of the cells in the model was determined, and the soil texture information extracted and assigned to one of seven classes, following FAO-Unesco (1974) (Figure A29).

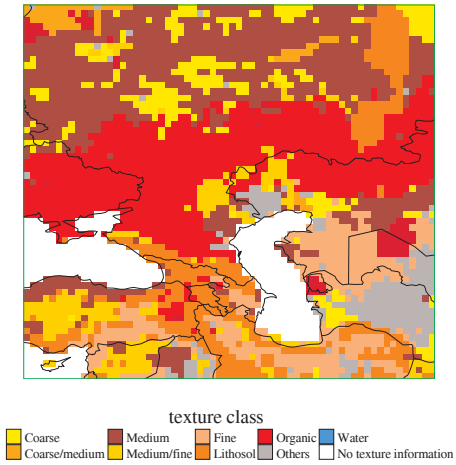


Figure A29: Caspian Sea basin FAO soil texture classes

Field and saturation capacity, expressed as percentages, were estimated for each soil texture class, following Vörösmarty *et al.* (1989), using empirical equations produced by Saxton *et al.* (1986). These equations predict soil moisture content given percentage sand and clay contents. These data and the resulting capacity values are shown in Table A12.

Table A12: ***Soil properties and calculated field capacity and saturation capacity***

Soil texture class	FAO texture class	% sand	% clay	Field capacity (%)	Saturation capacity (%)
Sand	Coarse	92	5	13.1	35.5
Sandy loam	coarse-medium	65	10	20.0	41.3
Silt loam	medium, coarse-fine, coarse-medium fine	20	15	29.4	46.8
Clay loam	medium-fine	35	35	33.1	50.4
Clay	Fine	20	60	48.3	54.5

To convert these capacity values into figures expressed as millimetres of storage (as needed by the PDM model), values of root depth were assumed for each soil texture class, distinguishing between forest and grass. The root depth values were taken from Vörösmarty *et al.* (1989). These data and the calculated field capacity and saturation capacity for each texture class are shown in Table A13. The field capacity for lithosols was taken directly from Vörösmarty *et al.* (1989), the lithosol saturation capacity was assumed to be 50% by volume, and organic soils were assumed to have field and saturation capacities of 50 and 100 mm respectively.

Table A13: ***Field capacity and saturation capacity by soil texture class***

Soil texture class	Root depth (m)		Field capacity, f_c (mm)		Saturation capacity, c_{max} (mm)	
	Forest	Grass	Forest	Grass	Forest	Grass
Sand	2.5	1.0	328	131	888	355
Sandy loam	2.0	1.0	400	200	826	413
Silt loam	2.0	1.3	588	382	936	608
Clay loam	1.6	1.0	530	331	806	504
Clay	1.2	0.7	580	338	653	381
Lithosol	0.1	0.1	27	27	50	50
Organic	-	-	50	50	100	100

Data on **lakes, reservoirs and wetlands** were needed in order to model their effect on flows. Ideally, the data required for each are: the general location and the location of the outlet (so that they could be assigned to the appropriate cells), the surface area and capacity, whether the water is fresh, salt or soda, and for reservoirs the type of use. For the larger water bodies in the Caspian Sea basin, attribute data were taken from a report by the Central Research Institute of Complex Usage of Water Resources (1986). For the smaller water bodies, visual techniques and large scale maps were used, and each wetland, lake or reservoir was assigned to the appropriate cell or cells, the proportions of the cells occupied were estimated, and the total surface areas of the water bodies were estimated. Where no better information was available, it was generally assumed that wetlands and small lakes had an average depth of one metre when full.

Other inputs for the surface water model were as follows:

For the routing between cells, the parameter k in the Muskingum method which is effectively the time delay was estimated from the difference in average elevation between the cells using initially:

$$k = 1.5\, e^{-D/250}$$

where k is the time delay in days, and D is the difference in elevation in metres. This somewhat arbitrary formula was developed to give very little delay in the steeper areas and increase gradually for flatter slopes. In previous studies (*e.g*. Meigh *et al*., 1998), when the simulated hydrographs were compared to observed for some of the larger catchments, it was found that the travel times corresponded closely. The other parameter in the Muskingum method, x, was set at 0.25 for all cells, a value which is generally found to be appropriate in a range of circumstances.

Losses in the flow between cells were considered to be negligible in the basin, and so proportional loss for all cells was set at zero.

A.2. Work carried out in the Intermediate Phase

A.2.1. Verification of input data

Following some preliminary calibration runs carried out in Phase I, it was decided to verify that the input baseline climate data (precipitation and evaporation) were good representations of measured data in the region during the 1961 to 1990 baseline period. This was carried out by comparing annual total values at several sites with maps of measured data. The precipitation data were found to be a good approximation of the measured data in the northern parts of the Caspian Sea basin, but less accurate in the southern region, particularly over the Caucasus (Kavkaz) mountains.

Discussions were held with Ms Anna Mescherskaya, of the Geophysical Observatory in St Petersburg, regarding the possibility of obtaining another alternative set of precipitation data to compare with the data already being used. The data that she offered were for the Volga basin, and averaged over large administrative regions. Considering that the raw data used to derive Ms Mescherskaya's mean values were unavailable for derivation of the necessary grid coverage, and the data already being used in the model were of good quality in that region, these alternative data were not used except to confirm that the data being used in the model were sufficiently accurate.

The potential evaporation data were verified in a similar manner, through comparison with maps of observed values. However, the data being used in the model were found to be generally rather high across the whole basin, thus explaining the rather low modelled runoff, and so were globally adjusted to provide more realistic values.

The data provided by Azerbaijan at the end of Phase I, particularly with regard to reservoirs in the basin, were verified in the model. However, data for modelling purposes had still not been received from Iran and Turkmenistan. The cell linkages were checked by hand, particularly in the Kavkaz region, to ascertain the cause of a problem with the model of the River Sulak. The flow network was corrected for the Sulak basin, and the corrected cell linkage network for the whole Caspian Sea basin is shown in Figure A9.

A.2.2. Model development

Originally, model calibration was performed by globally adjusting the parameters, such as for the PDM rainfall-runoff model and the snowmelt model, in the initialisation file for each model run. However, due to the large size of the Caspian Sea basin, and the many varied climates and flow regimes represented therein, it was decided to adjust the model to allow parameter calibration on selected pertinent sub-basins. The nine sub-basins used in the model are listed in Table A4, along with their key attributes, and plotted on the map of Figure A10. These sub-basins were selected to represent the range of hydrological regimes across the basin, and because good calibration data were available. Additionally, various other gauges (River Volga at Volgograd, River Ural at Kushum, Rivers Alazani and Arax (tributaries of the River Kura)) were also used to help verify the model calibration but were not treated as sub-basins. At some sites, naturalised data were available, and these were also used to validate

model results where the reservoir sub-model had not been used. The first aim of the calibration was to improve the modelling of annual total flow volumes, and then to improve the mean monthly hydrograph shape. Time series results (1961 to 1990) were also used to assist with calibration.

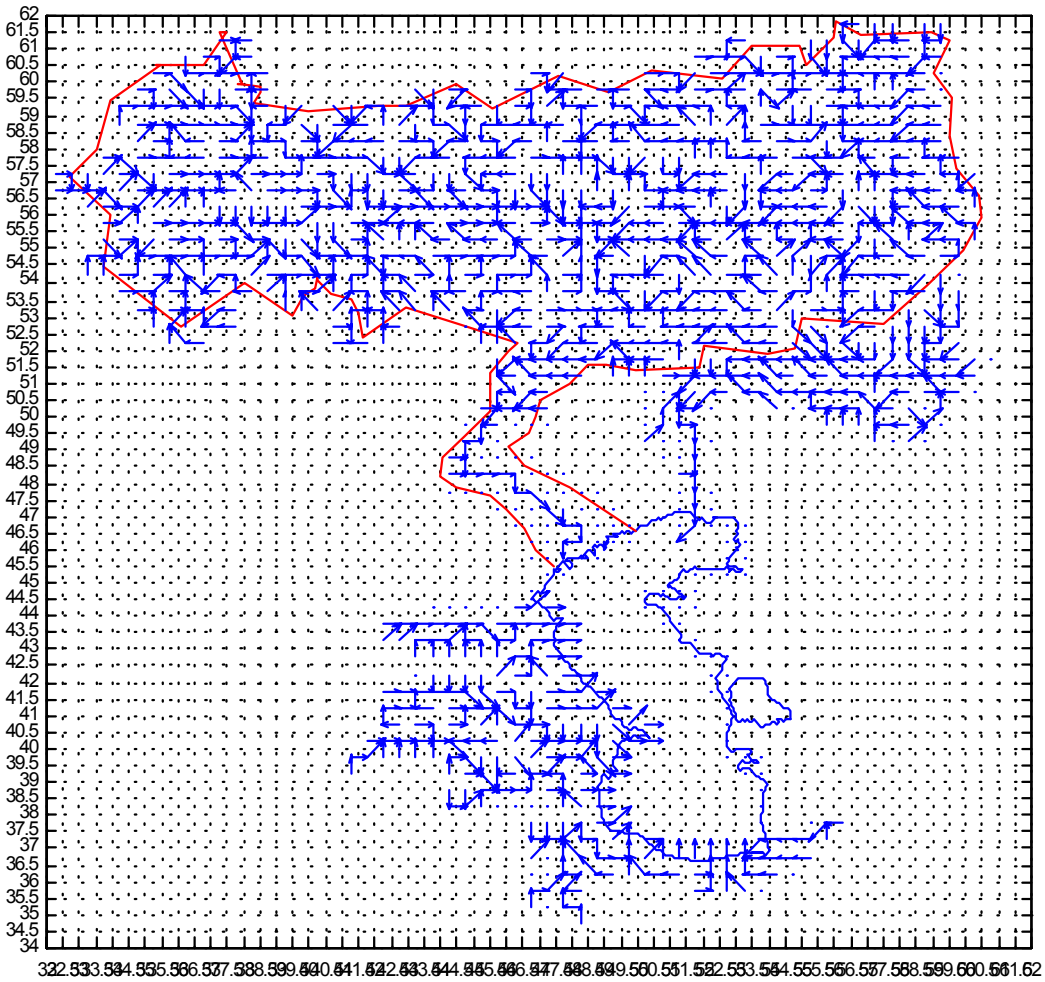


Figure A9 ***Corrected Caspian Sea basin cell linkages***

A minor alteration was made to the model to attempt to slow down the hydrograph recession in the low flow period of October to December, since this was originally considered too steep. The groundwater runoff routing equation was changed from:

$$bflow = grout * (gstor/100)^3$$

to

$$bflow = grout * (gstor/100)$$

where bflow is the routed groundwater runoff;

grout is the PDM baseflow routing parameter; and

gstor is the amount of drainage of the groundwater store.

However, this model alteration was seen to have only minor effects on the results.

Table A4 ***Sub-basins and gauges used to refine the model calibration***

Sub-basin ID	Name of sub-basin or gauging station	Location of gauge (cell reference)	Location of gauge (grid reference)	Model basin area (km^2)
1	Southern Caspian Sea basin	N/A	N/A	447,882
2	Oka (Volga)	22201455	55.5N,42.0E	232,251
3	Belaya (Volga)	23501455	55.5N,55.0E	114,850
4	Vyatka (Volga)	23101460	56.0N,51.0E	131,120
5	Ural inflow to Sea	23151365	46.5N,51.5E	239,899
6	Staritsa (Upper Volga)	21451460	56.0N,34.5E	17,250
7	Sura (Volga)	22601450	55.0N,46.0E	46,706
8	Vetluga (Volga)	22551465	56.5N,45.5E	37,985
9	Samara (Volga)	23001430	53.0N,50.5E	43,569
-	Volga inflow to Sea	22751355	45.5N,47.5E	1,477,348
-	Terek inflow to Sea	22751335	43.5N,47.5E	50,276

-	Sulak inflow to Sea	22751330	43.0N,47.5E	18,082
-	Kura inflow to Sea	22901290	39.0N,49.0E	180,819

A.2.3. Calibration procedure

To help understand how best to calibrate the model, a sensitivity analysis was performed. This was a way to ascertain, with some degree of confidence, the sensitivity of the simulated mean annual runoff to changes in the values of the key modelling parameters. The aim of the procedure was to assist with the model calibration, and this gave a broad, preliminary understanding of model sensitivity to some parameters which were felt likely to be important. It should be stressed that this was only intended to give some direction to model calibration, and was not taken as the definitive proof of model sensitivity; further, more detailed sensitivity analysis would be required to enable a sound understanding of the Caspian Sea basin model's sensitivity, particularly with respect to other, as yet untested elements of the model, and to combinations of model parameters.

Two cells were selected for the sensitivity analysis: one in the Oka sub-basin located within the Volga basin (grid cell code 21701445), and the other in the Alazani sub-basin, located within the Kura basin (grid cell code 22601315). The model was run several times, altering one of the parameters each time to assess the impact of the changes on the mean annual runoff to each of the two test grid cells. The sensitivity analysis involved the seven parameters listed below:

- The 'b' parameter of the PDM rainfall-runoff model, determining the spatial variability of the maximum soil capacity across the catchment (0.0 representing a constant field and 1.0 being a uniform variation, typically assumed to be 0.25);
- The Vörösmarty factor;
- The surface routing parameter (srout) of the PDM rainfall-runoff model;
- The groundwater routing parameter (grout) of the PDM rainfall-runoff model;
- The snowmelt factor (factmelt) of the snowmelt module;
- The threshold between precipitation falling as rain or snow (tempsnow) of the snowmelt module;
- The threshold temperature for snowmelt (tempmelt) of the snowmelt module.

The results of the sensitivity analysis are discussed below. First, the principal parameters used to calibrate the model, grouped into sub-model categories, are described in more detail.

Figure A10 ***A map of the Caspian Sea basin showing gauges and sub-basins used for calibration***

Parameters of the snowmelt module Snowmelt is a major component of the hydrograph of the Volga River, which itself provides around 80% of the total inflow to the Caspian Sea. The effect of snowmelt is to provide a delay, sometimes of several months, between precipitation and the resulting melt-waters which contribute to runoff. The snowmelt module used in this version of the PDM includes the following components which were adjusted during the calibration:

- A model for partitioning precipitation into rainfall and snow, typically according to the air temperature relative to some user defined 'freezing', or threshold, temperature;
- A threshold temperature for snowmelt;
- A snowmelt factor governing the depth of daily snowmelt proportional to some threshold air temperature above freezing;

- In applying a temperature-based approach for modelling snowmelt, and for partitioning precipitation into rainfall and snow, monthly mean temperatures were disaggregated into daily mean values. These daily values were obtained by linear interpolation between monthly values, assuming that the monthly value applies in the middle of the month, e.g. 15th Jan, 14th Feb...., and then applying the random component. This standard deviation, initially set at 2°C (according to a previous study for western and central Europe using the UK's Climate Research Unit (CRU) data set (Arnell, 1999)) was used in the calibration procedure.

Parameters of the PDM The main parameters of the PDM requiring calibration, as described in the Phase I Report, were:

c_{max} – maximum capacity of the soil moisture store (saturation capacity),

f_c – field capacity of the soil,

b – parameter describing the frequency of occurrence of the soil moisture stores,

s_{rout} – surface routing parameter,

g_{rout} – baseflow routing parameter.

It was not practicable to calibrate the model for each grid cell with reference to observed river flow data. Therefore, in order to allow it to be applied over the very large areas which were needed in this study, the parameters were determined from the physical characteristics of the cells. The advantage of this approach to rainfall-runoff modelling is that calibration for individual cells is not necessary as the parameters are derived solely from observed physical characteristics. This means that flows can be generated with a reasonable degree of confidence for each of the very large number of cells needed to provide coverage of large parts of the earth. Conversely, its disadvantage is that the available physical characteristics may not be sufficient to reflect the true variability of hydrological processes across the region. In this modelling study, these parameters were able to be calibrated by sub-basin. Further detail on the selection of model parameters is given in Meigh *et al*. (1998).

As described in the Phase I Report, soil properties (percentage field capacity and percentage saturation capacity) were converted into figures expressed as millimetres of storage (as needed by the PDM model) by assuming values of root depth for each soil texture class, distinguishing between forest and grass. The root depth values were taken from Vörösmarty *et al.* (1989). These data and the calculated field capacity and saturation capacity for each texture class are shown in Table A5. A 'Vörösmarty factor' was used to globally adjust these field and saturation capacities during the calibration procedure.

Table A5 Field capacity and saturation capacity by soil texture class

Soil texture class	Root depth (m)		Field capacity, f_c (mm)		Saturation capacity, c_{max} (mm)	
	Forest	Grass	Forest	Grass	Forest	Grass
Sand	2.5	1.0	328	131	888	355
Sandy loam	2.0	1.0	400	200	826	413
Silt loam	2.0	1.3	588	382	936	608
Clay loam	1.6	1.0	530	331	806	504
Clay	1.2	0.7	580	338	653	381
Lithosol	0.1	0.1	27	27	50	50
Organic	-	-	50	50	100	100

A summary of all of the calibration runs performed is presented in Table A6 below.

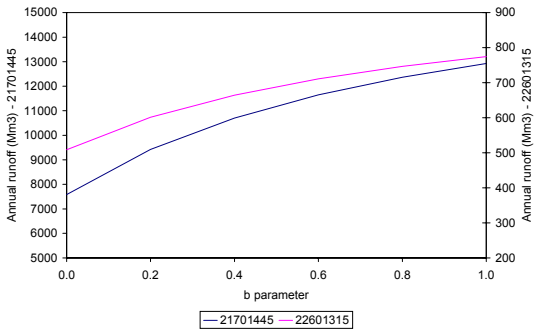
Table A6 *Calibration runs performed*

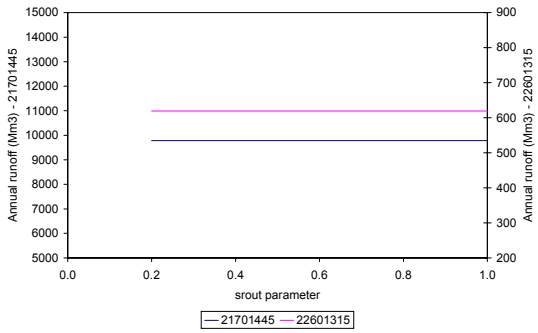
Run ID	Based on run	Details of alterations made to base run
Casp1		Proportional loss = 0 Lake & wetland depth set =1 metre if unknown Reservoir outflow equation constant = 1; Power = 1.5
Casp2	Casp1	Use routing/loss sub-model = No Use lake/reservoir/wetland sub-model = No
Casp3	Casp2	Threshold temperature for snowmelt = 0.5°C
Casp4	Casp2	Olson 100% forest cover changed to 80%
Casp5	Casp2	Vörösmarty factor = 1.2
Casp6	Casp2	Std. Dev. of daily temperature in snowmelt model = 2.5°C

Casp7	Casp2	Snowmelt factor = 1 mm/°C/day
Casp8	Casp2	Vörösmarty factor = 2.0
Casp9	Casp2	Both snow thresholds = 0.0°C
Casp10	Casp2	Muskingum k adjusted in parameters file to be $5*e^{(-D/250)}$
Casp11	Casp2	Vörösmarty factor = 2.0 Snowmelt factor = 1 mm/°C/day
Casp12	Casp11	PDM parameter grout = 0.025 (i.e. quartered) PDM parameter srout = 0.25 (i.e. quartered)
Casp16	Casp12	Using gwsm-cs3.exe and subchge.txt to test new version of program
Casp17	Casp16	Evaporation values reduced globally by 10%
Casp18	Casp12	Evaporation multiplied by a=0.854 and subtract b=5.4
Casp19	Casp18	PDM parameter grout = 0.1 Vörösmarty factor = 0.3
Casp20	Casp18	PDM parameter grout = 0.01 Vörösmarty factor = 0.3
Casp21	Casp19	Igstt=4, then gw store drains set for ivclasses 1 to 7 respectively: 0.01mm,0.0075, 0.005, 0.005, 0.0025, 0.005, 0.005. (i.e. halved)
Casp22	Casp21	Igstt=4, then gw store drains set for ivclasses 1 to 7 respectively: 0.005mm,0.00375, 0.0025, 0.0025, 0.00125, 0.0025, 0.0025. (i.e. quartered)
Casp23	Casp19	Using gwsm-cs4.exe and subchge.txt to test new version of program
Casp24	Casp19	PDM parameter b = 0.1
Casp25	Casp19	Using new version of gwsm-cs4.exe with 9 sub-basins, reservoir outflow doubled in Oct, Nov & Dec, and recession reduced in same period.
Casp26	Casp19	Using gwsm-cs5.exe to test new version of program, and reservoir outflow doubled in Oct, Nov & Dec.
Casp27	Casp21	Altered parameters by sub-basin, and corrected Sulak flow connections, plus reservoir component switched on, using gwsm-

		cs4 (which has reservoir outflow doubled and recession slowed-down in Oct-Dec)
Casp28	Casp27	Using gwsm-cs5 (which has reservoir outflow doubled in Oct-Dec)
Casp29	Casp27	Reservoir component switched off
Casp30	Casp29	Revised parameters
Casp31	Casp30	Std Dev. of daily temperature in snowmelt module = 0.5°C.
Casp32	Casp30	Revised parameters
Casp33	Casp32	Std Dev. of daily temperature in snowmelt module = 0.5°C.
Casp34	Casp30	Revised parameters; new parameters file (Samara basin adjusted); Std Dev. of daily temperature in snowmelt module = 0.5°C
Casp35	Casp34	PDM b=0.1 (instead of 0.24) in gwsm.ini file
Casp36	Casp34	Vörösmarty factor reduced from 0.3 to 0.2 in gwsm.ini file
Casp37	Casp36	PDM b=0.4 in gwsm.ini file

The sensitivity analysis indicated to which changes in which parameters the model was most sensitive. The results can be viewed in Figure A11.





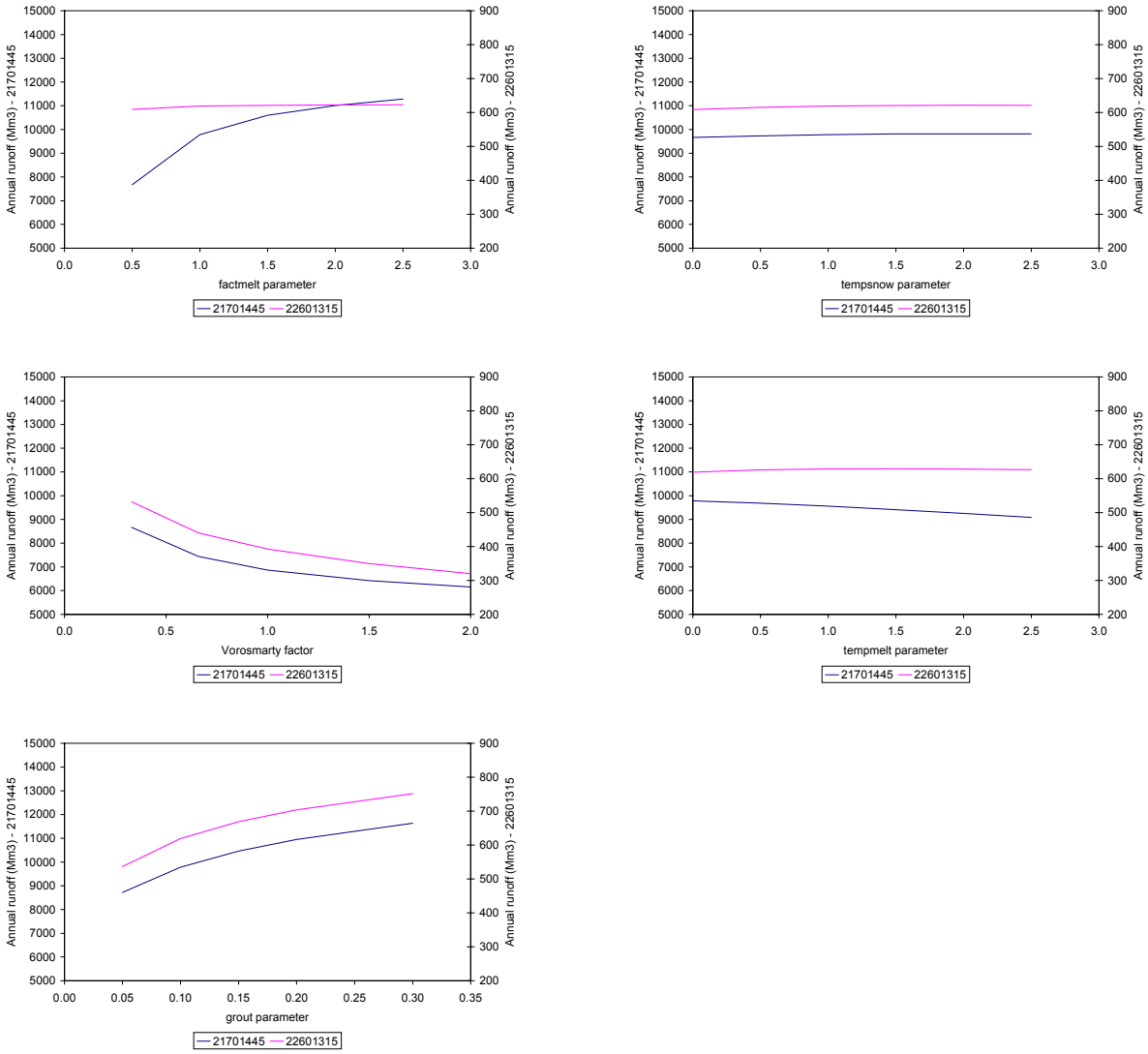


Figure A11 ***Results of the sensitivity analysis for two cells in the Caspian Sea basin for grid cell 21701445 (Volga basin) and 22601315 (Kura basin)***

The results for the Volga basin cell (21701445) show that its mean annual runoff is most sensitive to changes in the b parameter of the PDM rainfall-runoff model, but also to changes to the Vörösmarty factor, to the groundwater routing parameter and to the snowmelt factor. The range of mean annual runoff volumes relative to the parameters' minimum values (across the range of values tested) was 70%, 29%, 33%, and 47% for the b parameter, Vörösmarty factor, groundwater routing parameter and snowmelt factor respectively. The runoff volume is relatively insensitive to adjustment of the PDM surface routing parameter, to the threshold temperature between precipitation falling as rain or snow, and to the threshold temperature for snowmelt.

The results for the Kura basin cell (22601315) show that it is similarly most sensitive to changes in the b parameter of the PDM rainfall-runoff model, and also to the Vörösmarty factor, to the groundwater routing parameter, but not to the snowmelt factor (this is to be expected since this catchment receives a much lower volume of snowfall from its precipitation). The range of mean annual runoff volumes relative to the parameters' minimum values across the range of values tested was 52%, 40%, and 40% for the b parameter, Vörösmarty factor, and the groundwater routing parameter respectively. The runoff volume is also relatively insensitive to adjustment of the PDM surface routing parameter, to the threshold temperature between precipitation falling as rain or snow, and to the threshold temperature for snowmelt, as for the Volga basin grid cell.

A.3. References

Arnell N.W. (ed.), 1996a. *Impact of climate change on hydrological regimes and water resources in Europe*. Progress report 1994/95, University of Southampton, UK.

Arnell N.W., 1996b. *Global warming, river flows and water resources.* Wiley, Chichester.

Arnell, N.W., 1999. A simple water balance model for the simulation of streamflow over a large geographic domain, *Journal of Hydrology*, **217**, 314-335.

Arnell N.W. and Reynard N.S., 1996, The effects of climate change due to global warming on river flows in Great Britain. *J. Hydrol.*, **183**, 397-424.

Braun, L.N., Renner, C.B., 1992. Application of a conceptual runoff model in different physiographic regions of Switzerland, *Hydrological Sciences Journal*, **37**(3), 217-231.

Calder I.R., 1990. *Evaporation in the uplands*. Wiley, Chichester.

Central Research Institute of Complex Usage of Water Resources, 1986. *Preparation of data on the reservoirs of the USSR for basins, republics and types of usage.* Report by CRICUWR, Minsk, Russia.

Dastane N.G., 1974. *Effective rainfall in irrigated agriculture.* Irrigation and Drainage Paper 25, Food and Agriculture Organization, Rome.

DeFries R.S. and Townshend J.R.G., 1994. NDVI-derived land cover classifications at a global scale. *Int. J. Remote Sensing*, **15**, No.17, 3567-3586.

Doorenbos J. and Pruitt W.O., 1977. *Guidelines for predicting crop water requirements.* Irrigation and Drainage Paper 24 (revised), Food and Agriculture Organization, Rome.

FAO-Unesco, 1974. *Soil map of the world, Volume 1, Legend*. Food and Agriculture Organization, Rome/ United Nations Educational, Scientific and Cultural Organization, Paris.

Hydrometeorological Publishing House, 1968. *Climate Directory of the USSR*. Volumes 1 to 29, Leningrad, Russia.

Meigh J.R., McKenzie A.A., Austin B.N., Bradford R.B. and Reynard N.S., 1998. Assessment of global water resources – Phase II, Estimates of present and future water availability for Eastern and Southern Africa. *Institute of Hydrology / Department for International Development*, 104 pp.

Moore R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. *Hydrol. Sci. J.*, **30**, 2, 273-297.

Moore, R.J., Bell, V.A., Austin, R.M., Harding, R.J., 1999. Methods for snowmelt forecasting in upland Britain, *Hydrology and Earth Systems Science*, 3(2), 233-246.

NOAA-EPA Global Ecosystems Database Project, 1992. *Global Ecosystems Database Version 1.0.* User's Guide, Documentation, Reprints, and Digital Data on CD-ROM. USDOC/NOAA National Geophysical Data Center, Boulder, Colorado, USA.

Reynard N.S., 1996. Changes in surface water availability. In: Hulme M. (ed.), *Climate change and southern Africa: an exploration of some potential impacts and implications in the SADC region*. Climatic Research Unit, University of East Anglia, Norwich, 31-39.

Savenije H.H.G. and Hall M.J., 1995. Evaporation from wetlands, a loss or a benefit?. Unpublished paper to a meeting on "The role of wetlands in the hydrological cycle", Institute of Hydrology, November 1995.

Saxton K.E., Rawls W.J., Romberger J.S. and Papendick R.I., 1986. Estimating generalised soil water characteristics from texture. *Trans. Am. Soc. Agric. Engrs.*, **50**, 1031-1035.

Sene K.J., 1999. *Snowmelt module for GWAVA*. Private correspondence.

Vörösmarty C.J., Moore B., Grace A.L., Gildea M.P., Melillo J.L., Peterson B.J., Rastetter E.B. and Steudler P.A., 1989. Continental scale models of water balance and fluvial transport: an application to South America. *Global Biogeochem. Cycles*, **3**, 241-265.

Wigley T.M.L. and Raper S.C.B., 1992. Implications of revised IPCC emissions scenarios. *Nature*, **357**, 293-300.

Annex B:
GWAVA
Global Water Availability Assessment Model
Version 3.3 User Notes

B.1. GWAVA Version 3.3 User Notes: General

B.1.1. Background

The **G**lobal **W**ater **AV**ailability **A**ssessment method (GWAVA) was developed by the Centre for Ecology & Hydrology and the British Geological Survey in order to provide an improved methodology for the assessment of water resources in relation to water demand for application at the global scale. The key objective was to develop a method which was based on a more realistic representation of the world than had been achieved in previous assessments (for instance, Stockholm Environment Institute, 1997) and which provides more meaningful and more detailed results. The approach is based on a 0.5 by 0.5 degree latitude-longitude grid, and it attempts to improve on previous work by:

- adequately representing the spatial and temporal variation in both water availability and water demands, thus avoiding the distortions of national averages;
- simulating the effects of natural features (lakes, wetlands, glaciers, snow) and human-made features (reservoirs, abstractions and long-distance transfers) which have major effects on the availability and flow of water;
- accounting for the full range of demands, including domestic use, industry, irrigation and livestock;
- combining assessment of both surface water and groundwater;
- treating the problems of international river basins in a realistic manner; and
- providing for the assessment of future scenarios of the balance between water availability and demand due to climate change, population growth, urbanisation and increasing per capita consumption.

The main output of GWAVA is the comparison of water availability and demand at the scale of the grid cell, enabling the variability and complexity of the water resources situation to be appreciated. It is an approach which can be applied worldwide, allowing policy makers and others to make better informed resource allocation decisions, and facilitating the assessment of impacts of human activities (such as global warming) on global water distributions. For more details of the development of the GWAVA approach and the methodologies which are incorporated in the model, refer to Meigh *et al.* (1998, 1999).

B.1.2. Objective of these User Notes

This document is intended solely as a guide to how to set up and run the GWAVA model. Information on the modelling methodology can be found in other reports (Meigh *et al.*, 1998, 1999).

The model is a FORTRAN program. The current version to which these notes refer is **gwava33.exe**. It runs on a PC under Windows 95 or later, in a MS-DOS box.

The example files required for testing the model and manipulating data, as described in these user notes, are supplied on a CD-ROM. These files should all be copied onto the model's working directory on the user's computer. The complete set of files, and a brief description of each, is as follows:

List and description of files for running the GWAVA model

File name	Description
Initialisation files	
gwava.ini	Initialisation file for input files and global parameters of the GWAVA model
Executable programs (& necessary associated files)	
gwava33.exe	Executable program for the GWAVA model
ChkPara.exe	Program to check the parameters file for errors
Elevmod10min.exe; fnocmod.img	Program to extract average elevation data; source data
Elevmax10min.exe; fnocmax.img	Program to extract maximum elevation data; source data
Elevmin10min.exe; fnocmin.img	Program to extract minimum elevation data; source data
f_order.exe; test_ord.lst	Program to calculate cell processing order; list of cell codes & flow directions
getsoil.exe; Caspian_faotext.txt	Program to extract FAO soil data & regional data set.
get_olso.exe; owe13a.img; get_olso.ini	Program to extract Olson vegetation classification data; Olson source data; Look-up-table to convert Olson classifications to percentage forest cover for GWAVA
ChkRain.exe	Program to check climate baseline data for errors
getbaseclim.exe; cru_prec_6190.dat; cru_wet_6190.dat; cru_pen_6190.dat; cru_temp_6190.dat	Program to extract baseline climate data from global data sets; global baseline climate data on precipitation, rain-days, potential evaporation, temperature
get_ptanom.exe, global precipitation and temperature anomalies (on GED CD-ROM), gunzip.exe; cygwin1.dll	Program to extract precipitation and temperature data from CD-ROM, calculate anomalies and put into format for GWAVA; program and library to unzip the data from CD-ROM
ChkDems.exe	Program to check general water demands file for errors
getcc.exe; regional climate change data caspian_hadcm3_50_tmp.dat; caspian_hadcm3_50_ppn.dat; caspian_hadcm3_50_pe.dat; plus equivalents for CGCM1 and Echam)	Program to extract climate change data from global data files
code.lst	List of grid cell codes (required by all data extraction programs)

Input data files	
test_par_mountains.xls/.txt	Grid-based parameters file
cellcode.xls	Spreadsheet to calculate grid cell codes
cellarea.xls	Spreadsheet to calculate grid cell areas
subchge.xls	Sub-catchment calibration file
resinfo.txt	Large reservoir operation data file
areastor.txt	Large reservoir area:storage characteristics data file
testrnev.xls/.txt	Climate baseline data
test0505.txt	Precipitation and temperature anomaly file
testdems.xls/.txt	General water demands file
testmondems.xls/.txt	Monthly water demands file
testtran.txt	Data on transfers between cells
test2050.xls/.txt	Climate change scenario file
test_gw.xls/.txt	Groundwater availability file
mount.dat	Mountainous region info file

B.2. Data preparation

Input data need to be prepared as a series of ASCII files. Some of these are best prepared as spreadsheets in the first instance, and then converted to ASCII format to provide the final model input file; the sample input data files listed in Section B1 are therefore listed as both *.xls* and *.txt* files for completeness.

The input files required are:

- parameters – 0.5x0.5 degree grid
- climate data – 0.5x0.5 degree grid
- precipitation and temperature anomalies – 0.5x0.5 degree grid
- general water demands – 0.5x0.5 degree grid
- monthly water demands (optional)
- large reservoir operation information (optional)
- large reservoir area:storage characteristics (optional)
- sub-catchment calibration information (optional)
- transfers between cells (optional)
- climate change scenario – 0.5x0.5 degree grid (optional)
- groundwater availability – 0.5x0.5 degree grid (optional)
- mountainous region information (optional)

B.2.1. Parameters file [example: test_par_mountains.xls/.txt]

The file has one line for each 0.5 by 0.5 degree grid cell. Each line contains the following data items (in this order):

- grid cell code
- sub-catchment code
- mountain region code
- area of cell
- flow direction
- cell processing order
- soil class
- percentage forest cover
- proportional loss
- routing parameter

- percentage lake
- percentage wetland
- lake category
- lake characteristic: shape parameter
- lake characteristic: surface area
- lake characteristic: capacity
- lake characteristic: outflow constant
- lake characteristic: outflow power

The data are read in free format, so there must be at least one space between each item. When the data preparation has been completed, the program **ChkPara.exe** can be used to check whether there are any errors in the file format or the ranges of the values (output file from this is **chk_para.out**).

Grid cell code The first task which needs to be carried out before any other data preparation is to define a list of grid cell codes. The list defines the catchment (or catchments) being studied. Where the catchment boundary crosses a grid cell, a decision on whether to include or exclude each cell needs to be made. The cells should be listed starting at the most north-westerly point. List the cells moving east, then go to the next line 0.5^{o} south and again list the cells moving east, and so on, until the whole catchment is covered. The cell codes consist of eight digits which are assigned using the algorithm:

$$\text{cell code} = 100000\,(E+180) + 10\,(N+90)$$

where E is the degrees of longitude (W treated as negative) and N is the degrees of latitude (S treated as negative), and the point on the cell used is the south-western corner. As an example, the cell which is at 37-37.5^{o} E and 8-8.5^{o} S has the code 21700815. The spreadsheet **cellcode.xls** can be used for this calculation.

Sub-catchment code A code number specifying that the cell belongs to a particular sub-catchment (zero indicates that the cell does not belong to a specific sub-catchment). The sub-catchments are those which have been chosen to assist with the calibration of the model (i.e. for which calibration data exist). A sub-catchment has a unique whole number identifier, and every cell within that sub-catchment is allocated the same identifier. This must match the codes in the sub-catchment calibration parameters file (see Section B2.2). A maximum of twenty sub-catchments can be specified.

Mountain region code A code number specifying that the cell contains significant mountains and should be modelled using the glacier module (can be applied to all cells with mean elevation over 2000 metres). If the code is set to 1, the glacier model will be run, if set to 0, the glacier model will not be run for this cell.

Area of cell The area of cell is the land area (including lake, reservoirs, etc) in km^2. The range allowed is 0.001 to 3200 km^2. Where the cell is completely within the catchment this is dependent only on the latitude, and can be calculated using **cellarea.xls**. Where the coast or the catchment boundary cuts the cell, the land area can be estimated either by eye, or by using a GIS (Geographic Information System) to overlay the grid cells onto a map of the boundaries.

Flow direction The main direction of flow out of each cell needs to be determined. Directions should be assigned values 1 to 8 as follows:

NW=1	N=2	NE=3
W=4		E=5
SW=6	S=7	SE=8

The value 0 can also be assigned, indicating either the basin outlet to the sea, the end point of an internal drainage basin, or a cell which has no significant runoff (in extreme desert areas, for instance). Where a lake covers an entire cell, it is still necessary to define the flow direction so that the complete network linkage is obtained.

Initial selection of flow direction can be made using the difference between the cell's average elevation and the average elevations of the surrounding cells. Average elevation data on a 10-minute resolution can be obtained from the binary file **fnocmod.img** (taken from the GED CD-ROM). To extract these data on a 0.5-degree resolution use the program **Elevmod10min.exe**. The input file (example: **code.lst**) contains the list of cell codes. The output file is **elevmodof10min.dat**, containing the cell codes followed by the average elevation for the cell.

The direction values obtained by this approach may not be reliable. A thorough manual check, by comparison to topographic maps, is needed to ensure that the flow network defined is a reasonable representation of the real catchment.

Cell processing order The order in which the grid cells must be processed needs to be determined so that, when the flows are calculated for a particular cell, all those cells which contribute flow from upstream have already been processed. This can be done using the program **f_order.exe**. Two input files are needed (examples: **test_ord.lst**, **testtran.txt**). The first file contains the list of cell codes followed by the flow direction. The second contains information on transfers between cells (see Section B2.9). The output file is **code_ord.dat**; this is the same as **test_ord.lst** except that a third column is added, containing the processing order values. An additional output file **code_ord.out** provides details of which cells are upstream of each cell. The processing order must be recalculated whenever a flow direction value is altered or a transfer between cells is changed. The values can be in the range 1 to 1000.

Soil class (ivclass) This is used to determine the main parameters of the rainfall-runoff model. The dominant soil texture in each grid cell can be determined from the FAO Soil Map of the World (available as an Arc/Info coverage). The soil class value is assigned according to the following scheme: sand = 1, sandy loam = 2, silt loam = 3, clay loam = 4, clay = 5, lithosol = 6, organic = 7. In addition, the value should be set to 99 when the entire cell is covered by a lake or wetland. The global data are contained in 15 coverages, representing 11 regions of the world (Africa, Australasia, Central America, Central Asia, North East Asia, Europe and west of the Urals, Far East, North America, Near East, South America, South East Asia). Once a specific regional data set has been extracted using ArcInfo, the data for a particular catchment under study can be extracted from this data file using **getsoil.exe**. The input file **code.lst**, containing all the required grid cells in the correct order, will also be needed. The resultant output file, containing the grid cell followed by the soil code, in the correct order, will be called **faosoil.dat**.

Percentage forest cover This is also used to determine parameters of the rainfall-runoff model. The dominant vegetation classification for each grid cell can be extracted from the GED CD-ROM, either using the Olson ecosystem classification or the Normalised Difference Vegetation Index (NDVI) data for a representative year. The vegetation information needs to be converted into an assumed value of percentage forest cover (the range of values allowed is 0 to 100%). The following schemes have been used in previous work, but they may need adjustment in order to obtain good calibration for a particular region:

Assumed percentage forest cover for Olson ecosystem types

Olson ecosystem number	Class	Forest cover (%)	Olson ecosystem number	Class	Forest cover (%)
1	Conifer forest	100	10	Forest/field; dry evergreen broadleaf forest	50
2	Broadleaf forest	50	11	Wetlands, swamps	0
3	Mixed forest	50	12	Desert	0
4	Grassland/shrub+tree	0	13	Shrub/tree; succulent/thorn	0
5	Tropical/subtropical forest	100	14	Crop/settlement	0
6	Scrub and woodland or fields	0	15	Conifer snowy rainforest	100
7	Semi-desert	0	19	Mangrove	100
8	Field/woods complex or savanna	25	27	Tundra margins	0
9	Northern boreal taiga	25			

forest

The Olson data are in the file **owe13a.img**. To extract these data and convert to percentage forest cover use the program **get_olso.exe**. The input file (example: **code.lst**) contains the list of cell codes. The output file is **olson.dat**, containing the cell codes followed by the Olson ecosystem number and the assumed percentage forest cover. The conversion from the Olson number to percentage forest cover is defined by the data in the file **get_olso.ini**. These values can be changed either on the basis of available vegetation or land use maps, or to obtain improved agreement between model results and observed flow data. The NDVI data can also be obtained from the GED CD-ROM.

Assumed percentage forest cover for vegetation types derived using NDVI

Vegetation type	Forest cover (%)	Vegetation type	Forest cover (%)
Broadleaf evergreen forest	90	Grassland	5
Coniferous evergreen forest and woodland	70	Shrubs and bare ground	5
Mixed deciduous and evergreen forest and woodland	70	Bare ground	0
Broadleaf deciduous forest and woodland	70	Cultivated crops	0
Wooded grassland	25		

Proportional loss To account for transmission losses in flow between cells, a proportional loss can be applied to the flows from upstream coming into the current cell. Values must be in the range 0 to 0.999. A recommended starting value is 0.01 indicating a transmission loss of 1%.

Routing parameter For routing flows between cells the simple Muskingum method is used. The parameter k, approximately equivalent to the time delay in days, needs to be defined (the range allowed is 0 to 50). The other Muskingum parameter, x, is fixed at 0.25. In previous work the values of k were determined from the difference in average elevations between adjoining grid cells by the formula:

$$k = 1.5\, e^{-D/250}$$

where D is the difference in elevation in metres; for elevation data see under “Flow direction”. Note that in some areas, especially in relatively flat country, the elevation differences may be negative; such values should be set to zero.

Percentage lake and percentage wetland These values are used to reduce the local runoff from the cell when there are significant areas of lakes, reservoirs or wetlands. The percentage of lakes includes both reservoirs and lakes. The range allowed for both

percentage lake and percentage wetland is 0 to 100%, individually and for the sum of the two. But values of percentage lake < 0.1% and percentage wetland < 5% are ignored.

Lake category This defines the lakes, reservoirs and wetlands in the cell. The possible values are:

0 no lake, reservoir or wetland

1 there is a lake/reservoir but the outlet is not in this cell

2 outlet of a natural lake

3 downstream end of a lake/reservoir but there is no outlet

4 outlet of a regulation reservoir (for hydropower)

5 outlet of a regulation reservoir (for irrigation)

6 outlet of a regulation reservoir (for irrigation and hydropower)

7 outlet of a reservoir with direct abstraction (water supply, etc)

21 there is a wetland but the outlet is not in this cell

22 outlet of a wetland

23 downstream end of a wetland but there is no outlet

51 there is a salt/soda lake but the outlet is not in this cell

52 outlet of a salt/soda lake

53 downstream end of a salt/soda lake but there is no outlet

62 outlet of a special lake

90 outlet of a large reservoir with detailed information available

If the lake category is either 2, 4-7, 22, 52 or 62 then this is the outlet of a water body and data on its characteristics are needed, as defined below. The basic lake model (described below) is only run when the category has one of these values. Category 90 is used only for the outlet of important, large reservoirs. Detailed information on these should be included in a separate data file as described in Sections B2.3 and B2.4. For this special category, a more detailed model is run. The characteristics are not relevant for other category values. Reservoirs commissioned before the modelling period should be specified; but not those commissioned after the modelling period. If a reservoir was commissioned during the modelling period, it can be specified in the model, but the period of observed flow data used for calibration should be selected accordingly. Where there is more than one lake or wetland in a cell they need to be lumped together. Where small lakes occur in a cell, especially in a

cell which has a large river flowing through it, a decision has to be made whether or not to route the total flow through the lake; that is, whether to assign category 2 or 1 (for a fresh water lake or reservoir).

Lake characteristics: shape parameter, surface area, capacity, outflow constant, outflow power The shape parameter has the value 0 if the lake is assumed to have constant surface area whatever the water depth, or 1 if it is taken to have a V-shape (*i.e.* area varies linearly with depth). The surface area is the area of the water body when full (km^2). The capacity is also when full (million m^3), and it is live volume for reservoirs. The outflow constant and power are the values C and P in the equation which defines the outflow of the water body, Q_{out} (except when lake category = 7):

$$Q_{out} = C\, Q_{net\text{-}in}\, (S/S_{max})^P$$

where $Q_{net\text{-}in}$ = average net inflow in the current month; S_{max} = the capacity of the reservoir; and S = the storage in the previous month. In almost all cases, suitable values can be assumed to be C=1, P=1.5. The outflow constant has a different meaning in the special case of lake category = 7; for this, the value is taken as the monthly direct abstraction from the reservoir in million m^3. When lake category = 62 (outlet of a special lake), it is assumed that the capacity can increase indefinitely beyond the maximum value, S_{max}, rather than the excess being treated as spill and added to the outflow.

The ranges allowed for these parameters are: shape parameter 0 or 1; surface area 0.0001-200,000 km^2; capacity 0.0001-1x10^8; outflow constant 0.1-10, outflow power 0.5-20. To ensure that the program runs correctly, dummy values of these characteristics must be assigned for all cells, whatever the lake category value.

B.2.2. Sub-catchment calibration parameters file (optional) [example: subchge.txt]

The purpose of this file is to enable the user to calibrate the model for certain user-selected sub-catchments. Some of the critical modelling parameters have been identified during previous modelling trials, and the user can use this file to allocate values of these parameters for particular sub-catchments. This will over-ride the global parameter values defined in the initialisation file **gwava.ini** (see Section B3). The file has one line for each of the sub-catchments selected by the user (in numerical order), indicated by a unique sub-catchment index in the parameters file (see Section B2.1). Each line contains the following data items (in this order):

- sub-catchment index
- adjustment factor for Vörösmarty parameters (over-rides 'fact')
- surface water routing parameter in the PDM rainfall-runoff model (over-rides 'srout')
- groundwater routing parameter in the PDM rainfall-runoff model (over-rides 'grout')
- snowmelt factor, controlling how fast the snow melts (mm/degC/day) (over-rides 'factmelt')

- threshold temperature defining whether precipitation falls as rain or snow (degC) (over-rides 'tempsnow')
- threshold temperature for snowmelt (degC) (over-rides 'tempmelt')
- baseflow regression equation power for glacier cells (over-rides 'bfpower' for those glacier cells)
- baseflow regression equation power for non-glacier cells (over-rides 'bfpower' for those non-glacier cells)

B.2.3. Large reservoir operation data file (optional) [example: resinfo.txt]

This file contains more detailed information on those important, large reservoirs identified in the parameters file as category 90 (see Section B2.1). For every cell that contains the outlet of such a large reservoir of category 90, this file should have a corresponding detailed set of information on monthly reservoir storage profiles, and monthly direct abstraction and guaranteed outflow information, in the same order as the parameters file. This file also specifies when a reservoir comes on-line, and indicates the date of any changes in operation. The file contains the following information:

- grid cell code (which contains the outlet of the reservoir)
- year (indicating the year it came on-line, and/or the start year of a new pattern of operation/abstraction)
- maximum gross storage in the reservoir (million cubic metres), i.e. the total volume held up to spillway crest level
- dead storage (million cubic metres), i.e. the maximum gross storage minus the live (or usable) storage volume
- 12 monthly mean storage volumes (million cubic metres); this provides a normal operating curve/regime for the reservoir
- 12 monthly minimum release volumes (million cubic metres per month); this is the volume of flow which must be released from the reservoir in a particular month to guarantee a certain level of flow in the river downstream, e.g. for navigation purposes or environmental considerations
- 12 monthly abstraction volumes (million cubic metres per month).

This pattern of information is repeated for as many years as there are changes of operation (*i.e.*, there will be at least one line for each major reservoir).

B.2.4. Large reservoir area:storage data file (optional) [example: areastor.txt]

This file works in conjunction with the general information file described in Section B2.3. It contains detailed information on the area versus storage characteristics of those important, large reservoirs identified in the parameters file as category 90 (see Section B2.1). For every cell that contains the outlet of such a large reservoir of category 90, this file should

have a corresponding set of area;storage information in the same order as the parameters file. The file contains the following information:

- grid cell code (which contains the outlet of the reservoir)
- 5 area values, increasing in area from the minimum area (corresponding to minimum live storage) to the maximum area (corresponding to maximum live storage) (square kilometres)
- 5 storage values, increasing from the reservoir storage volume at the minimum live storage level to the maximum live storage level (million cubic metres).

The five areas should correspond with the five storage volumes, as in the following

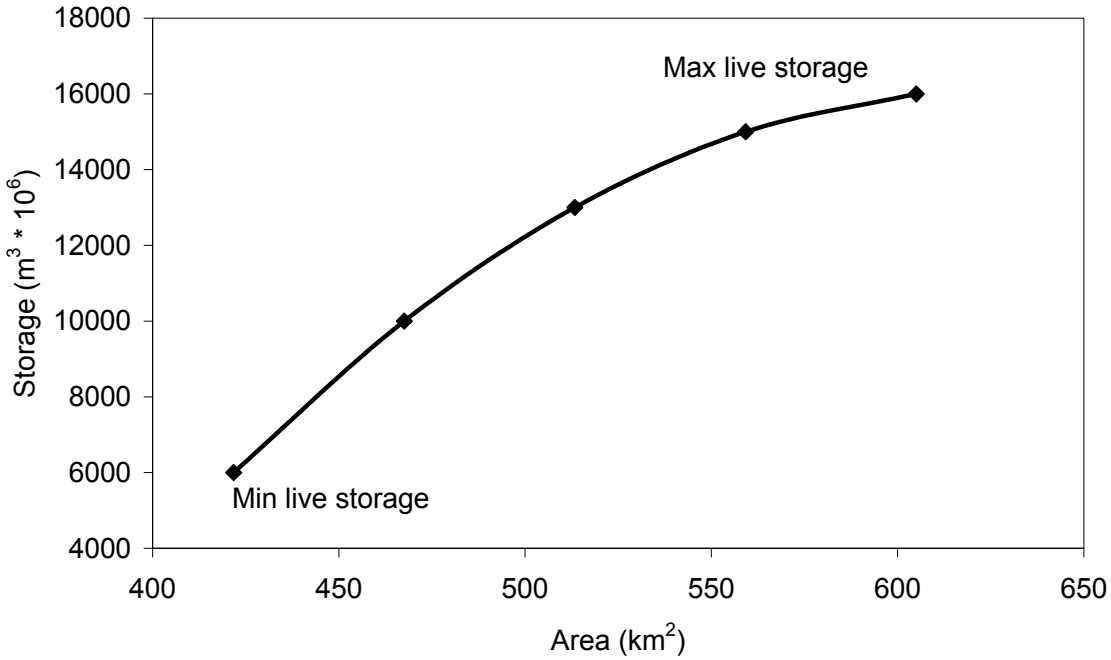


illustration:

B.2.5. Climate baseline data file [example: testrnev.xls/.txt]

The file has one line for each 0.5 by 0.5 degree grid cell. Each line contains the following data items (in this order):

- grid cell code
- 12 mean monthly precipitation values (mm) (January-December)
- 12 mean monthly values of number of rain-days per month (January-December)
- 12 mean monthly reference crop potential evaporation values (mm) (January-December)
- 12 mean monthly values of temperature (oC) (January-December)

The grid cell codes must be in the same order as in the parameters file. The data are read in free format, so there must be at least one space between each item. When the data preparation has been completed, the program **ChkRain.exe** can be used to check whether there are any errors in the file format or the ranges of the values (output file from this is **chk_rain.out**).

The mean values used here refer to the baseline period being used for the model run (often 1961-90). These gridded data can be obtained from the Climatic Research Unit (CRU) website (http://ipcc-ddc.cru.uea.ac.uk/). When using this source of data, reference crop PE values need to be calculated based on the temperature, humidity, sunshine and wind data. For the calculation of lake and wetland water balances, other values of PE are needed, and the model assumes that for evaporation from open water PE is 20% greater than reference crop, and from well-vegetated wetlands it is 40% higher. However, the global data for 1961-90 have already been extracted from the CRU website and the PE calculations have been performed (using the Penman method). Data are in the files: **cru_prec_6190.dat**, **cru_wet_6190.dat**, **cru_pen_6190.dat** and **cru_temp_6190.dat**, for precipitation, rain-days, evaporation and temperature, respectively. The data for the required cells can be extracted from these files and organised into the necessary file structure using the program **getbaseclim.exe**, again requiring the input file **code.lst**. The output file will be called **baseclim.out**.

B.2.6. Precipitation and temperature anomaly file [example: test0505.txt]

The precipitation anomaly data specify the percentage change for individual months from the means given in the climate baseline data file. The temperature anomalies are absolute changes from the means in the climate baseline data file. The anomaly data are specified on the same 0.5 by 0.5 degree grid as used elsewhere. The file has the following format (the cells must be in the same order as in the parameters file):

- 1st line – start year and end year. The total period defined by these years must be the same as the total period defined by the start and end years of the model run specified in the initialisation file gwava.ini (see Section B3); i.e., as long as the difference between the values is the same as the difference between those given in gwava.ini the values themselves can be different. The period should also be the same as that used to calculate the mean climate values as specified in the preceding section.
- Grid cell code.
- Line of data for each year of the run period: Year, 12 monthly precipitation anomaly values (% change) (Jan-Dec), 12 monthly temperature anomaly values (absolute change, oC) (Jan-Dec). The years in the first column must be all the years as defined in the first line of the file.
- Repeat the cell code and the following lines of data (for each year of the run period) for each cell to cover the basin.

The data are read in free format, so there must be at least one space between each item.

The 0.5x0.5 time series data for calculating the anomalies are available from the CRU website. The global data files have been downloaded onto a CD-ROM; there is a separate file for each year with the name formats: **glo.pre.19xx.Z** for precipitation (in mm) and **glo.tmp.19xx.Z** (in °C*10) for temperature, where xx is the year. These compressed files can be uncompressed, the relevant data extracted, the anomalies calculated and organised into the required format, using the program **get_ptanom.exe**. For this program to work, the file **code.lst** is again required, along with the baseline climate ASCII file described in Section B2.5, and two further files, to enable decompression, called **gunzip.exe** and **cygwin1.dll**.

B.2.7. General water demands file [example: testdems.xls/.txt]

The file has one line for each 0.5 by 0.5 degree grid cell. Each line contains the following data items (in this order):

- grid cell code
- total population in cell
- urban population in cell
- city code
- population of cattle in cell
- population of sheep and goats in cell
- urban demand per capita for cell
- rural demand per capita for cell
- industrial demand
- ratio of surface to ground water for this cell
- crop area (this and next two items repeated 8 times for up to 8 different crops)
- crop type code
- planting month for crop
- annual irrigation volume, where available
- monthly demands code

The grid cell codes must be in the same order as in the parameters file. The data are read in free format, so there must be at least one space between each item. When the data preparation has been completed, the program **ChkDems.exe** can be used to check whether there are any errors in the file format or the ranges of the values (output file from this is **chk_dems.out**). Note that to ensure that the program runs correctly, all the data items must be pre-assigned values for all cells; the file should be filled up with zeroes, except for the **ratio of surface to ground water** which should be set to –1.

In the special case where data on the actual quantities of significant water demands are available, then most of this file's data items are not needed. If only annual average data are

available, the actual quantities of significant water demands can be specified in the column for **Industrial demand**. This is normally the value for water consumed by industry in the cell. The value should be specified as the annual total of net abstraction (*i.e.*, gross abstraction minus return flows) in the cell in units of Mm^3. If detailed monthly demand data are available, these can be specified in an extra file of monthly demands and the monthly demands code set as described below.

In general cases, and where data on the actual quantities of significant water demands are not available, the specification of the data items should be as follows:

Population data The total population and the urban population in the cell are needed to allow estimation of water demands for domestic use. Ranges allowed for each are 0 to 10,000,000, but the urban value must not be greater than the total value.

City code This is normally 0, but for especially large or important cities a value of 1 can be specified. When this is done a higher rate of water use per capita is used than for ordinary urban areas.

Animal population data Animal populations (cattle, sheep/goats) in the cell are needed to allow estimation of water demands for livestock. Ranges allowed for each are 0 to 500,000.

Demands per capita The values of demand per capita (litres/capita/day) for urban and rural areas are used with the population figures to calculate the total requirements for domestic use. If values for individual cells are available they are specified here, but if only general information on typical demand figures is available, the values can be specified in the initialisation file (see Section B3). Demands are assumed to be constant throughout the year. The ranges allowed for each are 0 to 10,000 litres/capita/day.

Industrial demand The industrial demand is a gross figure for water consumed by industry in the cell in Mm^3/year. If data on quantities consumed for domestic use, livestock, etc are available, they can be included here, rather than calculated from the population data. The demand is assumed to be constant throughout the year. The range allowed is 0 to 15000 Mm^3/year.

Ratio of surface to ground water When calculating the amount of flow from a cell which proceeds to the next cell, the proportion of the demand which is supplied from surface water is needed, since where demand is supplied from groundwater it does not have any impact on the downstream flows. If the value of the ratio is set in the range 0 to 1, then this is the ratio used for the cell, and it is applied to all types of demand (so, 0=all groundwater, 1=all surface water). Alternatively the ratio can be set to -1 in which case default values specified in the initialisation file (see Section B3) are used for this cell.

Crop/irrigation data Information on irrigated crops consists of: the crop area (in ha); the crop code; and the planting month for the crop. This group of three values is repeated to allow information on up to 8 different crops to be specified for the cell. The water demands, varying by month, are calculated based on this information and the potential evaporation for the cell. The demands for different crops are summed for the cell. The crop code is used to

define the type of crop and a set of crop factors. The crop factors are proportions of a reference crop value, and are hard-coded (based on ones suitable for southern and eastern Africa); however, a new set of factors more suitable for the area under study could be substituted. The current ones are given in the table below (based on Doorenbos & Pruit, 1977). In this, the data are always given starting in month 1, but in computing the water demands the values are shifted to the appropriate months as defined by the planting month. This is given a number between 1 and 12 (1=Jan; 12=Dec).

Typical crop factors – General

Crop code	Crop	Month											
		1	2	3	4	5	6	7	8	9	10	11	12
1	Wheat/Barley	0.62	1.13	1.15	0.68	0	0	0	0	0	0	0	0
2	Rice	1.13	1.13	1.20	1.20	1.00	0	0	0	0	0	0	0
3	Peanuts	0.51	0.85	1.10	1.03	0.23	0	0	0	0	0	0	0
4	Cotton	0.50	0.71	1.11	1.20	1.18	0.95	0.37	0	0	0	0	0
5	Millet	0.62	1.13	1.10	0.26	0	0	0	0	0	0	0	0
6	Maize	0.53	0.91	1.15	1.08	0.37	0	0	0	0	0	0	0
7	Sugar cane	0.45	0.80	1.05	1.10	1.25	1.25	1.25	1.25	1.25	1.25	0.95	0.70
8	Citrus	0.65	0.65	0.70	0.70	0.70	0.75	0.75	0.75	0.75	0.75	0.70	0.70
9	Vegetables (unspecified)	0.51	0.93	1.14	0.47	0	0	0	0	0	0	0	0
10	Unknown short crop*	0.55	0.90	1.15	1.00	0.60	0	0	0	0	0	0	0
11	Fodder	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0	0	0	0	0
12	Horticultural fruits	0.50	0.69	0.98	1.00	0.58	0	0	0	0	0	0	0
13	Sorghum	0.53	0.92	1.10	0.91	0.10	0	0	0	0	0	0	0
14	Bananas	0.55	0.35	0.63	0.58	0.63	0.73	0.90	1.08	1.15	1.18	1.13	1.13
15	Tea/Coffee	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

*Can be used in cases where no information on crop type is available (based on a rough average of a range of crops).

A special set of crop factors that apply only to South Africa, Swaziland and Lesotho

are also included; see the next table. There are three sets of crop factors which apply to different parts of the countries, as defined in the table. Crop codes 101 to 111 apply to the northern part of the country, 201 to 201 apply to the south-west, and 301 to 310 to the south-east; for more information on the location of the regions see Midgley *et al.*, 1994. When these crop codes are used it is also necessary to specify the annual irrigation volume (in million m^3) in the last column of the file – this value is the volume of water used for irrigation, totalled over all crops in the cell. With this method, the monthly water volumes are first calculated based on the crop factors in the usual way, and the monthly volumes are then adjusted linearly so that the total volume of irrigation water used matches that specified in the last column.

Typical crop factors – South Africa, Swaziland and Lesotho

Crop code	Crop	Month											
		1	2	3	4	5	6	7	8	9	10	11	12
Northern (zones I, II, III, VI; drainage regions A, B, C, D, F, U, V, W, X)													
101	Vegetables	0.40	0.70	0.70	0.70	0	0	0	0	0	0	0	0.30
102	Fibre crops	0.74	0.82	0.61	0.43	0.37	0.12	0	0	0	0	0.30	0.49
103	Small grains	0	0	0	0	0	0.30	0.57	0.97	0.97	0.38	0	0
104	Summer grains	0.97	0.33	0	0	0	0	0	0	0	0.50	0.90	1.10
105	Sub-tropical fruit	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
106	Oil & protein seed	0.73	0.93	0.95	0.43	0	0	0	0	0	0	0	0.33
107	Citrus	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67
108	Vineyards/ Grapes	0.39	0.36	0.27	0.25	0.15	0.15	0.15	0.15	0.15	0.19	0.29	0.34
109	Deciduous fruit	0.50	0.59	0.42	0.20	0.20	0.20	0.20	0.20	0.23	0.27	0.31	0.41
110	Pasture	0.80	0.80	0.80	0.70	0.60	0.50	0.50	0.50	0.60	0.70	0.80	0.80
111	Sugar cane	0.90	0.92	0.93	0.91	0.88	0.82	0.78	0.74	0.72	0.74	0.81	0.85
South-western (zone IV; drainage regions E, G, H, J, K, L)													
201	Vegetables	0.67	0.70	0.70	0.14	0	0	0	0	0	0	0.30	0.39
203	Small grains	0	0	0	0	0.16	0.34	0.79	1.00	0.87	0.17	0	0
204	Summer grains	0.97	0.33	0	0	0	0	0	0	0	0.50	0.90	1.10
205	Sub-tropical fruit	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
207	Citrus	0.55	0.60	0.65	0.70	0.65	0.60	0.45	0.45	0.45	0.45	0.45	0.50
208	Vineyards/ Grapes	0.34	0.33	0.25	0.20	0.17	0.17	0.17	0.17	0.17	0.18	0.28	0.31

209	Deciduous fruit	0.45	0.57	0.48	0.20	0.20	0.20	0.20	0.20	0.20	0.24	0.27	0.36
210	Pasture	0.80	0.80	0.80	0.70	0.60	0.50	0.50	0.50	0.60	0.70	0.80	0.80
South-eastern (zone V; drainage regions M, N, P, Q, R, S, T)													
301	Vegetables	0.70	0.70	0.70	0	0	0	0	0	0	0	0.30	0.40
302	Fibre crops	0.69	0.82	0.62	0.31	0	0	0	0	0	0	0.18	0.29
303	Small grains	0	0	0	0	0	0	0.30	0.50	0.95	1.00	0.65	0
304	Summer grains	0.97	0.33	0	0	0	0	0	0	0	0.50	0.90	1.10
305	Sub-tropical fruit	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
306	Oil & protein seed	0.73	0.93	0.95	0.43	0	0	0	0	0	0	0	0.33
307	Citrus	0.55	0.60	0.65	0.70	0.65	0.60	0.45	0.45	0.45	0.45	0.45	0.50
308	Vineyards/ Grapes	0.39	0.37	0.27	0.20	0.15	0.15	0.15	0.15	0.15	0.19	0.29	0.34
309	Deciduous fruit	0.50	0.59	0.42	0.20	0.20	0.20	0.20	0.20	0.23	0.27	0.31	0.41
310	Pasture	0.80	0.80	0.80	0.70	0.60	0.50	0.50	0.50	0.60	0.70	0.80	0.80

Also note that, for these cases, the crop factors are specified for the appropriate months of the year, and this means that the planting month should always be set to 1.

Monthly demands code This is normally 0, but set to 77 if extra (monthly) demands information is available (see Section B2.8).

B.2.8. Monthly water demands file (optional) [example: testmondems.xls/.txt]

As noted in the previous section, where data on the actual quantities of water demands are available, these can be specified as average annual values in the column for **Industrial demand** in the general water demands file. Where additional data on monthly demand quantities are available, and also where the demands change during the period of the run, the additional information is specified in the monthly water demands file. Data should only be provided for cells which have the **monthly demands code** in the general water demands file

set to 77.

Each line contains the following data items:

- Grid cell code.
- Year (indicating the year that the specified monthly pattern of demands came into operation). The following years are assumed to have the same pattern of demand until a line for the same cell with a later year is found. If the earliest year that is specified for a particular cell is after the start date of the model run, the industrial demands in the period up to the earliest year are taken to be zero.
- 12 monthly demand values in million cubic metres per month (January-December); they should be specified as net abstractions (i.e., gross abstraction minus return flows).

The values specified here supersede any annual value of industrial demand specified in the general water demands file. Lines can be repeated for particular cells as needed, depending on the number of times that the pattern of demand changes. In this file the cells can be specified in any order, but all the lines for a particular cell must be grouped together and they must be in chronological order.

B.2.9. Transfers between cells file (optional) [example: testtran.txt]

Transfers between cells refer to links between cells (canals, pipelines, etc) which are in addition to the normal river network links defined by the flow direction parameter (in the parameters file, Section B2.1). Transfers can also be used to define flow in a delta where the flow out of a cell goes to more than one downstream cell. This file only contains data for cells which have a transfer associated with them. Each line contains the following data items (in this order):

- code for cell which transfer is coming from
- transfer type
- code for cell which transfer is going to
- definition of transfer

The grid cell codes (in the first column) must be in the same order as in the parameters file. The data are read in free format, so there must be at least one space between each item. In order to allow for transfers into or out of the area being modelled, dummy cell codes can be specified; these have a value of 00000099 or less. If the transfer is coming into the basin from outside, the dummy cell is specified in the first column, and these transfers must be listed at the beginning of the file. If the transfer is going from the basin to outside, the dummy cell is specified in the third column. In all cases where a dummy cell is specified, the transfer type must be set to 10 (see below). Note that the dummy cells are only included in the transfers file; they must <u>not</u> be specified in any of the other data files.

Transfer type The transfer type can have the following values:

1 transfer is a proportion of the flow in the originating cell

2-5 transfer is based on demand in receiving cell: 2 = use urban demand; 3 = use irrigation demand; 4 = urban plus irrigation demand; 5 = total demand

10 transfer is a fixed amount

Note that there can only be one transfer into a cell. There can be more than one transfer out of a cell, but these must be of the same type; that is, either all with transfer type = 1, or all with type 2-5, or all with type = 10.

Definition of transfer When transfer type = 1, this value is the proportion of the flow in the originating cell which is to be transferred. The total transfer proportion must not exceed 0.99. When transfer type = 2-5, it is a limit on the size of the transfer (for instance, the pipe or canal capacity) (Mm^3/month). When transfer type = 10, it is the amount of the transfer (Mm^3/month).

B.2.10. Climate change scenario file (optional) [example: test2050.xls/.txt]

The file has one line for each 0.5 by 0.5 degree grid cell. The file should only be included if the effects of climate change scenarios on the model are being investigated. Each line contains the following data items (in this order):

- grid cell code
- 12 % change values for mean monthly precipitation (January-December)
- 12 % change values for mean monthly potential evaporation (January-December)
- 12 absolute change values for mean monthly temperature (January-December)

Changes in precipitation and potential evaporation are specified as percentages. All changes refer to the change from the baseline period defined in the climate baseline and anomalies files, usually 1961-90 (see Sections B2.5 and B2.6).

In addition to the above, the first line of the file contains only the year to which the scenario refers (*e.g.*, 2050). The grid cell codes must be in the same order as in the parameters file. The data are read in free format, so there must be at least one space between each item.

The gridded climate change data can be obtained from the Climatic Research Unit (CRU) website, and PE values calculated in the same way as for the baseline climate data. Data have been extracted by CEH Wallingford from the CRU website and the PE calculations have been performed for the results from three GCM experiments. For the HadCM3 experiment (UK Hadley Centre for Climate Prediction and Research), the data files specific to your region under study are named: **caspian_hadcm3_50_tmp.dat**, **caspian_hadcm3_50_ppn.dat** and **caspian_hadcm3_50_pe.dat** for temperature, precipitation and evaporation, respectively. For the CGCM1 experiment (Canadian Centre

for Climate Modelling and Analysis), **hadcm3** is replaced by **cgcm1_GG1** in the file name, and for the ECHAM4 experiment (German Climate Research Centre), it is **echam_GG1**. For each GCM, the data for the required cells need to be extracted from these files, and combined into a single file as defined above; this can be achieved using **getcc.exe**, which once again requires not only the input climate files, but also the file of grid cell codes, **code.lst**.

B.2.11. Groundwater availability file (optional) [example: test_gw.xls/.txt]

The specification of this file is given below. Note that in regions where groundwater is of negligible significance, the file is not needed.

The file has one line for each 0.5 by 0.5 degree grid cell. Each line contains the following data items (in this order):

- grid cell code
- annual groundwater availability (Mm3)
- monthly groundwater factor

The grid cell codes must be in the same order as in the parameters file. The data are read in free format, so there must be at least one space between each item.

It is normally assumed that the availability of groundwater is uniform throughout the year. In this case the monthly groundwater factor should be set to 12. In cases where it can be assumed that there is variation within the year, a higher value of the factor can be specified. The factor is used to determine the groundwater availability in the month for which it is least. Thus, if a value of 16 is specified, the minimum monthly groundwater availability is 75% of the value obtained when availability is assumed to be constant. The factor should not be less than 12.

B.2.12. Mountainous region data file (optional) [example: mount.dat]

This file contains more detailed information on those cells with a mean elevation over 2000m, identified as such in the parameters file, with a mountain region code of 1 (see Section B2.1). If the glacier module has been selected in the initialisation file, for every cell that has a mountain region code of 1, this file should have a corresponding detailed set of information on elevations and climate gradients, in the same order as the parameters file. The file contains the following information:

- grid cell code (which has a mountain region code of 1)
- Zo, the elevation at which the half-degree baseline climatology was derived (this will usually be the mean elevation in the cell, but can be adjusted for calibration purposes) (metres)
- itype, a code of either 0 or 1. Code 0 indicates that the Pareto distribution should be used to describe the range of elevations in the cell; code 1 indicates that pairs of co-

ordinates describing the elevation distribution will be provided by the user. Code 0 is currently the recommended choice

- Zmin, the minimum elevation in the cell (metres)
- Zaverage, the average elevation in the cell (metres)
- Zmax, the maximum elevation in the cell (metres)
- GradT, the gradient of temperature with elevation (degrees Celsius decrease/increase per kilometre of elevation above/below Zo); a value up to 7°C/km is recommended
- GradP, the gradient of precipitation with elevation (percent increase/decrease per kilometre of elevation above/below Zo); a value of 20%/km is recommended

The gradients of temperature and precipitation are used to adjust the baseline climate according to the elevation distribution in the grid cell. The elevation data can be obtained using three supplied programs. Average elevation data on a 10-minute resolution are contained in the binary file **fnocmod.img** (taken from the GED CD-ROM). To extract these data on a 0.5-degree resolution use the program **elevmod10min.exe**. The input file (example: **code.lst**) contains the list of cell codes. The output file is **elevmodof10min.dat**, containing the cell codes followed by the average elevation for the half-degree cell. Minimum elevation in the half-degree cell, taken as the minimum of 9 10-minute minimum elevation values in the cell, can be obtained using **elevmin10min.exe** with the binary source file **fnocmin.img**, and **code.lst**. The output file is called **elevminof10min.dat**. Similarly, maximum elevation in the half-degree cell, taken as the maximum of 9 10-minute maximum elevation values in the cell, can be obtained using **elevmax10min.exe** with the binary source file **fnocmax.img**, and **code.lst**. The output file is called **elevmaxof10min.dat**.

B.3. Running the model

The details of a model run are defined in the model initialisation file **gwava.ini**. This specifies the input and output files to be used as well as model parameters which are constant over all grid cells. An example of the file is given below, and it is then explained, line by line. The file is read by a formatted read, so the precise format of the file as in the example must be maintained.

Example of initialisation file **gwava.ini**:

3. Initial data file for Global Water Availability Assessment Model (gwava33.exe)
4. Input parameter file . .....test_par_mountains.txt
5. Input subcatchment calibration parameters file .....subchge.txt
6. Input lake/reservoir operation info file .....resinfo.txt
7. Input lake/reservoir area:storage info file .....areastor.txt
8. Input mountainous region info file .....mount.dat
9. Input data file (precip, rain-days, evap, temp) .....testrnev.txt
10. Input precipitation & temp anomaly file test0505.txt
11. Input water demands data file ....testdems.txt
12. Input extra (monthly) water demands file.....testmondems.txt
13. Input transfers between cells data file .....testtran.txt
14. Input climate change scenario file test2050.txt
15. Input groundwater availability data file .....test_gw.txt
16. Output data file - monthly time series (Mm3) .....test_tss001.out
17. Output data file - summary of local flows (Mm3) .....test_loc001.out
18. Output data file - summary of total flows (Mm3) .....test_tot001.out
19. Output data file - demands (Mm3) .....test_dem001.out
20. Output data file - results (Mm3) .....test_res001.out
21. Start year of main run (run length must not be 1961
22. End year of main run (..less than 10 years) 1990
23. Number of years for warm-up - rec minimum is 30 (nyrwup) ...30
24. Output time series tot.qs(itsout=1), local qs(=2), none(=0).0
25. Use clim ch scenario-varying(iscen=1),uniform(=2), none(=0).0
26. If iscen=2, uniform % change for rainfall 0.0
27. If iscen=2, uniform % change for evaporation 0.0

28. If iscen=2, uniform absolute change for temperature 0.0
29. Use groundwater availability data (igwav=1), or not (=0) ...1
30. Use actual demands (iodem=0), or uniform ideal demands (=1) 0
31. Include irrigation demands (iirri=1), or not (=0) 1
32. Do calcs using extra (mthly) dems file (isxdem=1), not (=0).1
33. Use routing/loss model (isrout=1), or not (=0) 1
34. Use lake/reservoir/wetland model (islake=1), or not (=0) ...1
35. Use glacier model (ismount=1), or not (=0) 1
36. Snow-Use snowmelt component (isnowm=1), or not (=0) 1
37. Use subcatchment calibration (isubc=1), or not (=0) 0
38. Use transfers (itran=1), or not (=0) 1
39. Number of days to be added to rain-days (nadd) 5
40. Forest intercept. par. (treefac) - PE increase for forest ..1.1
41. Forest interception model parameter (gam) 2.0
42. Forest interception model parameter (delta) 0.5
43. Adjustment factor for Vorosmarty parameters (fact) 1.0
44. PDM parameter (b) 0.25
45. PDM parameter (grout) 0.1
46. grout set to zero when ann rain < this (groutx) NOT USED ...0.0
47. PDM parameter (srout) 1.0
48. Baseflow recession power (bfpower) 3.0
49. Min flow (monthly q < this, assume q=0) (qamin) NOT USED ...0.0
50. Snow-Std dev of daily temperature (degC) (tempsd) 2.0
51. Snow-Threshold split precip to rain/snow (degC) (tempsnow) .1.0
52. Snow-Threshold temperature for snowmelt (degC) (tempmelt) ..0.0
53. Snow-Snowmelt factor (mm/day/degC) (factmelt) 2.0
54. Snow-Storage time constant- slow drainage (per day) (rk1) ..0.15
55. Snow-Storage time constant- fast drainage (per day) (rk2) ..0.85
56. Snow-Sc* in Moore et al, 1999 (storefac) 0.1
57. Snow-Threshold temp at which drainage stops(degC)(tempcrit).0.0
58. % loss for water supplies - urban (pwlosu) 40.0
59. % loss for water supplies - rural (pwlosr) 20.0

60. Irrigation efficiency (%) (effirr) 50.0
61. Proportional return - urban supplies (pretu).....0.0
62. Proportional return - rural supplies (pretr).....0.0
63. Uniform ideal demand (l/c/day), if iodem=1 - urban (udemu) .60.0
64. Urban dem. (l/c/d) for special/important cities (udemusp) ..140.0
65. Weight between special & ordinary urban demands (specfac) ..0.5
66. Uniform ideal demand (l/c/day), if iodem=1 - rural (udemr) .25.0
67. Demand for cattle (l/c/day) (demcow) 37.5
68. Demand for sheep & goats (l/c/day) (demgoat) 20.0
69. Population increase factor - urban (pplusu) 1.0
70. Population increase factor - rural (pplusr) 1.0
71. Population increase factor - cattle/sheep&goats (pplusa) ...1.0
72. The rest are for drainage of gw store in PDM:
73. igstt = 0 no drainage, 1 fixed, 2 cubic fn. on ann. rain,
74. 3 exp. fn. on rain, 4 on depends on ivclass 0
75. igstn = 0 gstore can't go -ve, 1 it can 1
76. Value or constant in function when igstt=1,2,3 (gst_in) ....0.
77. When igstt=4, gw store drainage in mm for ivclass=1 .....0.02
78. When igstt=4, gw store drainage in mm for ivclass=2 .....0.015
79. When igstt=4, gw store drainage in mm for ivclass=3 .....0.01
80. When igstt=4, gw store drainage in mm for ivclass=4 .....0.01
81. When igstt=4, gw store drainage in mm for ivclass=5 .....0.005
82. When igstt=4, gw store drainage in mm for ivclass=6 .....0.01
83. When igstt=4, gw store drainage in mm for ivclass=7 .....0.01
84. Prop. sw to gw to be used - urban demands (swgwpu) 0.8
85. Prop. sw to gw to be used - rural demands (swgwpr) 0.3
86. Prop. sw to gw to be used - industrial demands (swgwpx).....0.7
87. Prop. sw to gw to be used - irrigation demands (swgwpi) 1.0

Explanation of **gwava.ini**, line by line:

Line 1. Header – text doesn't matter.

Lines 2-13. Specify the names of the data files described in Section B2.

Lines 14-18. Specify the names of the output files (see Section B4).

Lines 19-20. Defines the start and end years of the run. The length of the run must not be less than 10 years, and it must match with that given in the precipitation and temperature anomaly file, as described in Section B2.6.

Line 21. Defines the number of years used for "warm-up" before the main model run. This helps to ensure that the model produces stable results. The value should normally be 30 years.

Line 22. Controls the output of monthly time series data. Set it to 1 for output of time series of total flows in each cell; 2 for local flows in each cell; or 0 for no time series output (see Section B4.1)

Lines 23-26. Line 23 controls the use of climate change scenarios in the model. Set it to 1 for climate change scenarios based on GCM results (see Section B2.10); 2 to use a uniform percentage change over all grid cells (the % or absolute changes are then defined in the next three lines); or 0 for no scenario.

Line 27. Controls the use of groundwater availability data in the model (see Section B2.11). Set it to 1 to use the groundwater data; or 0 to ignore it.

Line 28. Determines the source of information on water demands per capita. Set it to 0 to use demands which vary from cell to cell, as defined in the general demands file (Section B2.7); or 1 to use uniform demands across all cells (see lines 61 to 63).

Line 29. Controls the use of the irrigation demand component of the model. Set it to 1 to include irrigation; or 0 to ignore it.

Line 30. Controls the use of the monthly demands component of the model. Set it to 1 to use this; or 0 to ignore it.

Line 31. Controls the use of the routing/loss component for transmission of water between cells. Set it to 1 to use this; or 0 to ignore it.

Line 32. Controls the use of the lake/reservoir/wetland component of the model. Set it to 1 to use this; or 0 to ignore it.

Line 33. Controls the use of the glacier component of the model. Set it to 1 to use this; or 0 to ignore it.

Line 34. Controls the use of the snow component of the model. Set it to 1 to use this; or 0 to ignore it.

Line 35. Controls the use of the sub-catchment calibration component of the model. Set it to 1 to use this; or 0 to ignore it.

Line 36. Controls the use of the transfers component of the model. Set it to 1 to use this; or

0 to ignore it. Note that this could affect the order in which the cells are processed – see under “Cell processing order” in Section B2.1.

Lines 37-47. These are the rainfall-runoff model (and forest interception model) parameters. Generally, they should not be changed from the values given in the example.

Lines 48-55. These are the parameters of the snow component. The values set in the example are typical values which are widely applicable, but they could be adjusted to refine the performance of this component.

Lines 56-69. All these lines refer to specification of the demand component of the model.

Lines 56-57. These are the assumed percentage losses for water transferred from a source to urban and rural supplies.

Line 58. Irrigation efficiency (%) to be used where there is an irrigation demand in the cell.

Lines 59-60. These are the proportions of water supplied to urban and rural areas which are assumed to return to the river.

Line 61. If the value in line 28 is set to 1, this is the urban water demand (litres/capita/day) used for all cells.

Lines 62-63. When the city code for a cell (in the demands file) is set to 1, the special urban demand value (litres/capita/day) specified here (line 62) is used for the proportion of the total urban demand as defined in line 63.

Line 64. If the value in line 28 is set to 1, this is the rural water demand (litres/capita/day) used for all cells.

Lines 65-66. Water demand per head (litres/head/day) for cattle and for sheep/goats.

Lines 67-68. For model testing, factors to increase the human (urban and rural) and livestock populations can be specified here, rather than supplying scenarios of increased populations by grid cell in the demands file.

Lines 70-81. These define additional parameters of the rainfall-runoff model, relating to the capability to allow for a small rate of loss from the flows. This is helpful in arid areas where rivers become dry seasonally. The value in line 72 should either be 4 (with the remaining parameters set as in the example), or it can be set to 0 to remove this option.

Lines 82-85. The proportion of the demand which is supplied from surface water rather than groundwater is needed when calculating the amount of flow which proceeds from one cell to the next.

B.4. Model output

The model produces five different kinds of output file:

- time series of monthly flows
- mean monthly local flows
- mean monthly total flows
- mean monthly demands
- annual data and water availability indices

B.4.1. Output of time series of monthly flows

Each line of the file contains the following values (in this order):

- counter – this runs from 1 for the first cell processed to the total number of cells for the last
- grid cell code
- year
- 12 monthly flow values (January-December)

These lines are repeated for each year of the run before going on to the next cell. The order of the grid cells in the output is that in which they are processed; *i.e.*, it is different from the order of the cells as defined in the input files.

Time series output is optional. It is controlled by the option on line 19 of the initialisation file (see Section B3). The flow values output can be either total flows or local flows, or no time series can be output at all; this can be useful because these files tend to be very large.

B.4.2. Output of mean monthly flows and demands

There are three separate files. Each line of the files contains the following values (in this order):

- counter – this runs from 1 for the first cell processed to the total number of cells for the last
- grid cell code
- 12 mean monthly flow values in Mm3 (January-December)
- 12 standard deviations of monthly flow values (January-December) (not output for total flows)

One file contains the local flows for each cell (*i.e.*, the runoff generated within that cell), the second contains the total flows for the cells (in this case the standard deviations are omitted), and the third contains the demands. The order of the grid cells in the output is that in which they are processed; *i.e.*, it is different from the order of the cells as defined in the input files.

B.4.3. Output of annual data and water availability indices

This contains the main results of the model run. Each line has the following values (in this order):

- counter – this runs from 1 for the first cell processed to the total number of cells for the last
- grid cell code
- three surface water availability indices (SWAI–types 1, 3 & 4)
- three groundwater availability indices (GWAI–types 1, 3 & 4)
- three combined water availability indices (TWAI–types 1, 3 & 4)
- average annual total flow* (Mm3)
- standard deviation of annual total flow (Mm3)
- average of driest monthly total flows in each year (Mm3)
- driest monthly total flow with 90% reliability (Mm3)
- average annual local flow (Mm3)
- standard deviation of annual local flow (Mm3)
- four annual demand values (Mm3) – domestic, industrial, livestock, irrigation
- monthly demand used for comparison with flow in WAI–type 2 (Mm3)
- annual groundwater recharge estimated by rainfall-runoff model (Mm3)
- annual groundwater recharge estimate derived from losses between cells (Mm3)
- lake index
- warning index

The order of the grid cells in the output is that in which they are processed; *i.e.*, it is different from the order of the cells as defined in the input files.

Water availability indices The water availability indices are defined as follows:

Definition of water availability indices

	Index	Definition
Surface water only	SWAI–type 1	Total annual runoff / Total annual demand

* This is average annual total flow before demand in the cell is removed and therefore may not be the same as in sections 4.1 and 4.2.

	SWAI–type 2	90% reliable driest month runoff / Minimum monthly demand
	SWAI–type 3	Minimum over all months of: (90% reliable monthly runoff – Demand for that month)
	SWAI–type 4	(SWAI–type 3) / (90% reliable monthly runoff + Demand for that month)
Ground water only	GWAI–type 1	Annual groundwater yield / Total annual demand
	GWAI–type 2	Minimum monthly groundwater yield / Minimum monthly demand
	GWAI–type 3	Minimum over all months of: (Monthly groundwater yield – Demand for that month)
	GWAI–type 4	(GWAI–type 3) / (Monthly groundwater yield + Demand for that month)
Com-bined	TWAI–type 1	(Total annual runoff + Annual groundwater yield) / Total annual demand
	TWAI–type 2	(90% reliable driest month runoff + Minimum monthly groundwater yield) / Minimum monthly demand
	TWAI–type 3	Minimum over all months of: (90% reliable monthly runoff + Monthly groundwater yield – Demand for that month)
	TWAI–type 4	(TWAI–type 3) / (90% reliable monthly runoff + Monthly groundwater yield + Demand for that month)

Lake index Surface water availability estimates are generally based on the 90% reliable monthly flow. But, in cases where there is a lake or water supply reservoir in a cell (occupying more than 5% of the area of the cell; but wetlands and salt or soda lakes are disregarded), an alternative method of estimating water availability is also used. The value is taken as 50% of the mean annual runoff or half the capacity of the storage, whichever is less. This value is taken as the water availability if it is greater than the value provided by the normal approach. The lake index is normally -1, but if the lake estimate is used it is set to 1.

Warning index This value is normally 0. In calculating the type 3 water availability index, the minimum of the 90% reliable monthly runoff less the demand for that month is used. It would usually be expected that the minimum runoff and the minimum demand roughly coincide. If this is not so, it may indicate that there is a problem with the specification of the demands, and this index is set to 1.

B.5. Calibrating the model

In order to improve the model results, with respect to the real-world system being studied, it is necessary to calibrate the model. The most important requirement for calibration is a set of suitable gauged flow series for the rivers being modelled.

B.5.1. Selection of suitable gauged flow series

Great care should be taken to ensure that a gauged flow series being used for calibration represents the same series as has been modelled. Important considerations are:

- The gauged series should lie on the main river reach being modelled in a cell, not on a tributary of that reach, nor on a different, minor river in the same cell;
- A gauged series should be used for each main river being modelled, and each gauge should ideally lie at the downstream end of each river, to capture all the catchment characteristics in a single series;
- The gauged series should match the period being modelled. In most cases this will be 1961-90. Care should be taken if a reservoir has been modelled which was commissioned during the modelling period; ensure that the gauged flow series is downstream of the reservoir.

B.5.2. Suggested calibration procedure

After setting-up and running the model for the first time, the model results should be compared with the gauged flow series for calibration. Errors and inaccuracies in the model can often be identified at this stage, and the model adjusted accordingly. Typically, the mean annual total flows at the gauged locations should be compared, to ensure that the volumes are similar. Then, the summary of mean monthly total flows produced by the model can be compared with the mean monthly total gauged flows. A graph of the comparable series will show where the shape of the monthly hydrograph needs to be improved; for instance, the peak flow may be too high in the model, or the baseflow component may be too low.

After identification of any inaccuracies, the model inputs can be adjusted accordingly and the model re-run.

A sensitivity analysis is a sensible procedure, whereby the value of a particular parameter is given a range of values and the model re-run for each value in the range. In this way, the parameters to which the model is most sensitive can be identified and given suitable attention in model runs.

B.6. References

Doorenbos J. and Pruit W. O., 1977. *Guidelines for predicting crop water requirements*. Irrigation and Drainage Paper 24 (revised), Food and Agriculture Organization, Rome.

GED CD-ROM: NOAA-EPA Global Ecosystems Database Project, *Global Ecosystems Database Version 1.0.* User's Guide, Documentation, Reprints, and Digital Data on CD-ROM, USDOC/NOAA National Geophysical Data Center, Boulder, Colorado, USA, 1992.

Meigh J.R., McKenzie A.A., Austin B.N., Bradford R.B. and Reynard N.S., 1998. Assessment of global water resources – Phase II, Estimates of present and future water availability for Eastern and Southern Africa. *Institute of Hydrology / Department for International Development*, 104 pp.

Meigh J.R., McKenzie A.A. and Sene K.J., 1999. A grid-based approach to water scarcity estimates for eastern and southern Africa. *Water Resources Management*, **13**, 85-115.

Midgley D.C., Pitman W.V. and Middleton B.J., 1994. *Surface Water Resources of South Africa 1990*. 13 vols, Water Research Commission Report Nos 298/1/94 to 298/6.2/94.

Moore R.J., Bell V.A., Austin R.M., and Harding R.J., 1999. Methods for snowmelt forecasting in upland Britain, Hydrology and Earth Systems Science*, **3**(2), 233-246.*

Stockholm Environment Institute, 1997. *Comprehensive Assessment of the Freshwater Resources of the World*. 8 vols.