THE FINNISH WATERSHED SIMULATION AND FORECASTING SYSTEM (WSFS)

Bertel Vehviläinen[1] and Markus Huttunen[2]
Finnish Environment Institute SYKE, PO Box 140, 00251 Helsinki, Finland
[1]bertel.vehvilainen@ymparisto.fi
[2]markus.huttunen@ymparisto.fi

***Abstract*:** WSFS (Watershed Simulation and Forecasting System) is widely used in Finland for real time simulation of hydrological cycle and for watershed forecasting and simulation of climate change effects on water resources. The system is based on a catchment rainfall-runoff model with the same basic structure as the HBV-model widely used in Scandinavia. The new version of the WSFS covers the total land area of Finland including cross-boundary watersheds 390 000 km^2. The sub-basin division of the model is the third level watershed division (60-100 km^2) or 1x1 km^2 grid in research models. The input values of the model are daily precipitation and temperature. The simulated variables are areal precipitation, evapotranspiration, snow storage, soil moisture, lake evaporation, surface, sub-surface and ground water storage, runoff, discharges and water levels of main rivers and lakes. An elevation model is used to calculate the height and gradient effects on areal precipitation, temperature and snow melt. In an ongoing project the catchment model is improved by using land-use, vegetation and soil type data to increase the accuracy of the spatial variation of the simulated components (snow, soil moisture, soil frost, evapotranspiration). The remote sensing data in the model are satellite data from snow covered area and precipitation from weather radar. In the Envisnow EU-project methods to use ENVISAT snow data within WSFS are developed. An important part of the WSFS in forecasting is the automatic model updating system developed in SYKE. Model updating is done against water level, discharge and snow line observations. The updating procedure corrects the model simulation by changing the areal values of temperature, precipitation and potential evaporation so that the observed and simulated discharges, water levels and water equivalent of snow are equal. The inputs of the WSFS are (1) precipitation from 200 observation stations, (2) temperature from 40 stations, (3) precipitation from weather radar over 30 000 km^2 and (4) Class-A pan evaporation from 22 stations. During updating procedure the models in WSFS are corrected against (1) water level observations from 435 stations (163 in real time), (2) discharge observations from 322 stations (128 in real time), (3) snow water equivalent observations from 178 stations and (4) data of snow covered area from satellite pictures in real time.

***Keywords*:** hydrology, modeling, forecasting, simulation, weather radar, satellite, snow, discharge, runoff, water level

THE FINNISH WATERSHED SIMULATION AN FORECASTING SYSTEM (WSFS) DAS FINNISCHE SIMULATIONS- UND VORHERSAGEVERFAHREN FÜR WASSEREINZUGSGEBIETE

Übersicht: Das WSFS wird in Finnland in weiten Umfang benutzt für Echtzeitsimulationen des Wasserkreislaufes, für die Vorhersage von Wassereinzugsgebieten und die Simulation der Auswirkungen des Klimawandels auf den Wasserhaushalt. Das Verfahren beruht auf einem Regenhöhen-Oberflächenabfluss Modell mit dem gleichen Grundaufbau wie das in Skandinavien weit verbreitete HBV-Modell. Die neue finnische WSFS Version erfasst die gesamte finnische Landfläche was zusammen mit den grenzüberschreitenden Wassereinzugsgebieten ein Fläche von 390.000 qkm ergibt. Die Unterteilung der Teileinzugsgebiete richtet sich nach der Klasse-3-Unterteilung der finnischen Wassereinzugsgebiete, oder ist für Forschungszwecke enger gerastert (1x1qkm). Die Eingabewerte des Modells sind die täglichen Niederschläge und die Temperatur. Die simulierten Werte sind die Gebietsniederschlagshöhe, die Gesamtverdunstung, die Schneemenge, die Erdfeuchte, die Verdunstungsmenge der Seen, die Oberfläche, das Boden- und Grundwasser, die Abflussmenge sowie die Austrittsmengen der grossen Flüsse

und Seen. Ein Höhenmodell wird verwendet um den Einfluss von Höhe und Landschaftsverlauf auf Gebietsniederschlagshöhen, Temperatur und Schneeschmelze zu ermitteln. In einem laufenden Projekt wird das Erfassungsmodell durch die Einbeziehung von Landnutzung, Vegetation und Bodentyp verbessert, um die Richtigkeit der räumlichen Änderungen der simulierten Werte (Schnee, Erdfeuchte, Bodenfrost, Verdunstungsquote) zu erhöhen.

Die Schneebedeckung wird mittels Satellitenfotos, der Niederschlag mittels Wetterradar ermittelt. Im Envisnow-Programm der EU werden Methoden entwickelt um die Schneedaten von ENVISAT im WSFS zu verwenden. Einen wichtigen Beitrag zur Vorhersagefähigkeit von WSFS ist das im Finnish Environmental Institute entwickelte Model-Updating-System. Das Model-Updating reagiert auf Änderungen von Wasserstand, Durchflussmenge und Schneebedeckung. Es korrigiert indem Temperatur- und Durchflusswerte sowie die zu erwartende Verdunstungsmenge so geändert werden, dass die beobachteten und simulierten Durchflüsse, Niederschläge und die Wasseräquivalente des Schnees konsitent sind. Die Eingabewerte für WSFS sind (1) Niederschlagsmengen von 200 Beobachtungsstationen, (2) die Temperaturen von 40 Messstationen, (3) Niederschlagswerte des Wetterradars für 30.000qkm und (4) die Verdunstungswerte von 22 Messpunkten. Für das Update wird das WSFS korrigiert gegenüber (1) Wasserstandsbeobachtungen von 435 Stationen (davon 163 in Echtzeit), (2) Abflussmengen von 322 Messstationen (davon 128 in Echtzeit), (3), Wasser-Schnee-Äquivalentmessungen von 178 Stationen und (4) Daten der prozentualen Schneebedeckung durch Satellitenfotos in Echtzeit.

General description of WSFS

The main operating part of the Watershed Simulation and Forecasting system (WSFS) is a distributed watershed model, which simulates the hydrological cycle using standard meteorological data.

The other independent systems to which the WSFS is connected are the hydrological data register, operative watershed management systems, automatic real-time water level and discharge observation station network, synoptic weather stations reporting through the Finnish Meteorological Institute (FMI), weather forecasts from European Centre of Medium-Range Weather Forecasts (ECMWF) via the FMI. Connection to a diffuse load simulation system (VEPS) has also been built up.

The WSFS automatically reads watershed data from the registers, runs forecasts and distributes results to Internet and registers for different users. The different stages in watershed forecasting are: (1) weather data transfer in real-time from the FMI, (2) automatic collection of watershed data from registers, (3) automatic watershed model updating according to information obtained, (4) forecast runs by watershed models, (5) distribution of forecasts through Internet: www.vyh.fi/tila/vesi/ennuste/index.html.

1. Watershed model

1.1. Hydrological model

The basic component of the watershed model in WSFS is a conceptual rainfall-runoff model, which simulates runoff with daily precipitation, potential evaporation, and temperature as input (Figure1). The rainfall-runoff model is originally based on the HBV-model structure (Bergström, 1976). The main parts of the rainfall-runoff model are precipitation, snow, soil moisture, subsurface, and ground water models. A watershed is divided approximately into 100 km^2 sub-basins with parameter set calibrated against observation or given according to the available GIS-information. Each sub-basin is further divided into 1 km^2 grid. Precipitation and temperature is first interpolated for each 1 km^2 grid using elevation model and the then the runoff model is simulated separately for each grid. Each grid has snow, soil moisture, sub-surface and ground water storages. The basic rainfall-runoff model is then calibrated

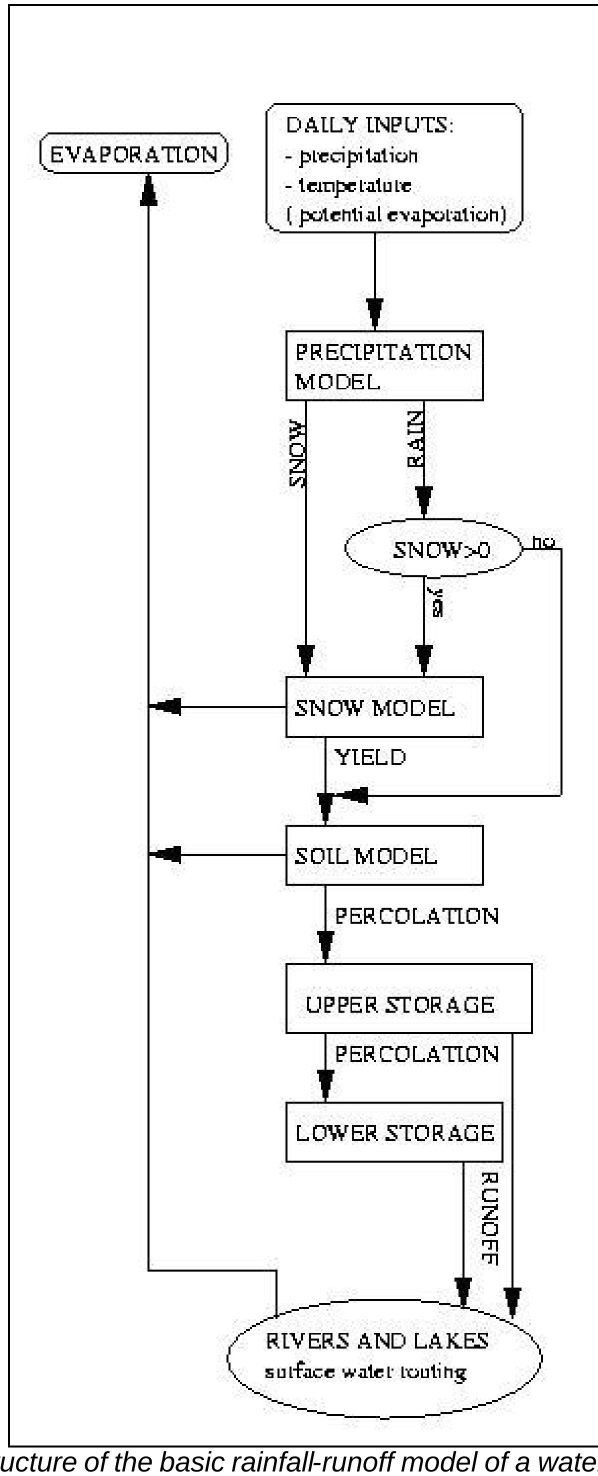


Figure 1. General structure of the basic rainfall-runoff model of a watershed model in WSFS

more or less specifically for each sub-basin in the watershed depending on the available data. The combined runoff from different sub-basins is routed through river and lake models in the order the rivers and lakes are situated. The combination of hydrological rainfall-runoff models, river routing models, and lake models forms the watershed model.

1.2. Precipitation model

Precipitation model simulates the quantity of areal sub-basin precipitation and the form of precipitation. Station precipitation error, mainly due to wind effect, is corrected by a constant, which is different for solid and liquid precipitation and depends on the openness of the station. This correction constant is also specific for sub-basins in the model and takes into account, elevation, and other terrain effects resulting from location of the rain gauge in the basin. If the rain gauge is situated in the higher part of the basin it usually measures more precipitation than is obtained on the average in the sub-basin. When the rain gauge is located below or at mean altitude of a basin the measured precipitation is usually less than the average over the basin. Thus the correction coefficients may be also under 1.0 both for liquid and solid precipitation, but are typically 1.03 - 1.06 for liquid precipitation and 1.2 - 1.3 for snow (Vehviläinen, 1992). Since the correction constants are specific for a sub-basin the combination of precipitation stations used should be constant, or if stations are changed the omitted station values are first approximated from existing stations.

Precipitation changes from snow to water within a temperature range approximately from -2.5 to 2.5 ^{o}C (Vehviläinen, 1992); included in these threshold values are elevation, coastal and other effects influencing the form of precipitation. The temperature range is specific for a sub-basin with a certain combination of temperature stations. If the combination is changed the omitted station values should be calculated from the values of operative stations.

1.3. Snow model

Snow model simulates snow accumulation and snowmelt. Inputs are the earlier calculated areal precipitation and daily temperature. Snowmelt is simulated by degree-day model with increasing degree-day value during the melt period. Open and forest snowmelts are simulated separately, which is essential for correct simulation of long melt periods with cold and warm spells and to create appropriate distribution of areal snow cover. The parameters in snowmelt model are more or less specific for a sub-basin and stations used. Other important processes in snow model simulation are liquid water retention in snowpack, refreezing of melted water, and simulation of snow-covered area and temporary surface storage during snow cover. Temporary storage causes delay in water outflow from the sub-basin due to snowdrifts and snowpack restricting water-flow through the terrain.

1.4. Soil moisture simulation

Soil moisture is simulated with a storage model in which input is rainfall, snowmelt and potential evaporation; output term is actual soil evaporation, which is simulated according to the saturation of soil. When the soil becomes saturated, the actual evaporation approaches potential evaporation; outflow from soil moisture storage into the subsurface storage is an exponential function of the degree of the saturation of soil. Soil moisture storage is active and changing during summer, when risk for flood and long drought can be forecasted based on the state of soil moisture storage and precipitation forecast. When soil moisture storage is full abundant rainfall causes flooding, when empty, the soil surface is dry, rainfall creates little runoff and inflow into lakes and rivers remains low.

1.5. Subsurface and groundwater storage

Water from soil moisture storage recharges the subsurface storage. The outflow from the subsurface storage creates mainly the runoff peaks during high flow. From the subsurface storage water percolates into the groundwater storage from which outflow is baseflow. The model structure is based on the ‘old’ hydrological concept of runoff formation, where runoff is divided into surface, subsurface and groundwater flows. New knowledge based on isotope studies in which runoff formation is due more to the functioning of recharge/nonrecharge areas is not taken into account directly. The use of many sub-basins leads generally to a similar simulation of recharge/nonrecharge areas: the sub-basins near river or lake usually have small soil, subsurface and groundwater storages and respond more quickly to rainfall and melt. The upper sub-basins have larger storages and longer response times, and the outflow remains higher due to them for longer periods; thus quantitatively, the ‘old’ hydrological concept of runoff formation works well in watershed forecasting.

1.6. River models

The river model is a routing model which simulate the delay and damping of peak flow in a river stretch. The river routing model is built up automatically based on the GIS-data of the watershed. The river model uses lake percentage of the sub-basin and the length of the river stretch on the sub-basin to simulate the delay and damping of the flow. Hydraulic models have also been used, when water levels are needed along the river. These are based on Saint Venant equations (see e.g. Forsius, 1984).

1.7. Lake models

Lake models are water balance models. The input terms are inflow from rivers, runoff from the surrounding land, and direct precipitation, lake outflow and lake evaporation. During winter the diminishing storage due to ice layer remaining on the lake bottom is taken into account in heavily regulated lakes and reservoirs. The calibration of discharge rating curve for lake outflow is possible in the watershed model when water level observations from an unregulated lake are available. It is a quick and inexpensive way to construct a rating curve for a lake. If daily inflow is large compared with lake volume the simulation time step could be shortened to one hour or even less to maintain the simulation stable.

For regulated lakes it is possible to give regulation rules and change them automatically or manually as needed in forecasting. Testing of different regulation procedures in a difficult flood situation is often needed to minimize the damage caused by flooding.

2. Model calibration

The optimization criteria in the calibration are the weighted sum of square of differences between observed and simulated water equivalents of snow, discharge, and water level. All available observations are used in the calibration and thus up to 100 different calibration criteria can be available in a watershed model calibration.

The calibration method used is the direct search Hooke-Jeeves optimization (Hooke and Jeeves, 1960, Kuha, 1993), which has been developed in to a fully automatic procedure using quite much computer time but little interactive manual work time. There are 5 - 10 important basin-specific parameters in each sub-model which must be calibrated for each sub-basin and 5 - 10 'constant' parameters which are slightly tuned for each sub-basin, thus the total number of calibrated and tuned parameters is within 10 - 20 within a sub-basin and the number. To manage the calibration properly it is started by the same parameter set for all sub-basins. This stage is divided into two steps: first the parameters of precipitation and snow models are calibrated against observed areal snow values after which other parameters are calibrated against the total water balance (discharge or water level and outflow) of the watershed. The sub-basins are then divided into 2 - 4 homogeneous groups for which the same procedure is repeated: snow and precipitation models are calibrated against observed snow measurements and runoff parameters against total water *balance*. Finally each sub-basin is calibrated against all observation points downstream. For each parameter has been defined limits how much the parameter value of one sub-basin may differ from the mean value of all sub-basins and from the values of the adjacent sub-basins. These limits are used to define the a priori believes of how much sub-basins may differ from each other. The reason for defining the limits is to decrease the possibility for over calibration. Finally all the parameters of each sub-basin are calibrated together against the nearest water balance and snow observations. The final result of this calibration procedure is a realistically distributed watershed model in which the calibration has been done against snow and water balance (discharge, water level) observations.

3. Data assimilation and updating

One important part of the WSFS is the automatic model updating system developed in SYKE. This system takes care that the watershed models are in the best possible state before forecast evaluated according to the real time observations and makes the updating possible: a task unreasonable to do manually due to the large amount of observation points (250 water levels and 160 snow lines) and sub-basins (5000). The model updating data are gathered from different registers. When new data become available the updating procedure corrects the model simulation by changing the areal values of temperature, precipitation and potential evaporation so that the observed and simulated discharges, water levels and water equivalent of snow are equal.

Daily corrections for precipitation, temperature and potential evaporation over long periods for sub-basins lead to large number of parameters to be optimized. To reduce the dimensionality in the updating, one 'correction-term' for each day with values of -1 ... +1 was introduced; the precipitation, temperature and potential evaporation corrections are calculated from this term. With a positive correction term precipitation and temperature are increased and potential evaporation is decreased within prescribed limits and vice versa. Furthermore, the correction is the same for 30- day groups. If the result of optimization is not satisfactory, groups of 15, 8, 4 days and even 1 day may be used to improve the accuracy. The same correction terms are also used for all sub-basins above a water level or a discharge observation point. Typically the updating is done for the last 100 days with 4-day groups. This correction procedure is also denoted data assimilation, as for example in Chui and Chen (1991).

The optimized error function consists of 12 different error terms. The error terms are:

- difference in simulated and observed discharge
- difference in simulated and observed volume
- difference in simulated and observed water equivalent of snow
- total amount of correction in precipitation, temperature and evaporation. smaller corrections are preferable.
- areal variability of the correction terms. In the basins close to each other the correction terms should not different much.
- effect of the correction terms on water balance of the basin during the first day of the forecast. The correction procedure should not make unnecessary changes to the water balance.

An example of the time series of the error terms is in Figure 2**.** The direct search Hooke-Jeeves optimization algorithm (Hooke and Jeeves, 1960, Kuha, 1993) is used in the updating. It was found to be more reliable and numerically stable than gradient-based Quasi-Newton and Levenberg-Marquardt algorithms; this is perhaps due to the fact that the function (watershed model) to be optimized is not differentiable and is high-dimensional. Another good feature of the Hooke-Jeeves algorithm is that it is not a 'line-minimizer' type of algorithm, i.e. the optimization routine does not try to correct the entire error in the simulation by corrections for the first few days.

Wintertime precipitation correction is not permitted before total snowmelt because the effects of snow and snow correction are not realized before the snow is melted. Winter is very troublesome in updating because ice in river causes unreliable discharge observations, if they are based on water levels and discharge curve. The only reliable discharge data from these places are direct measurements made once or twice during winter. The accuracy of snow simulation is supported by the use of all available precipitation measurements even with 1-2 months delays. During snow-free periods the precipitation affects with 0-2-day delay on the water balance, i.e. water levels and discharges. Thus updating can be done almost up to the last observed values of water level and discharge in basins with short response time.

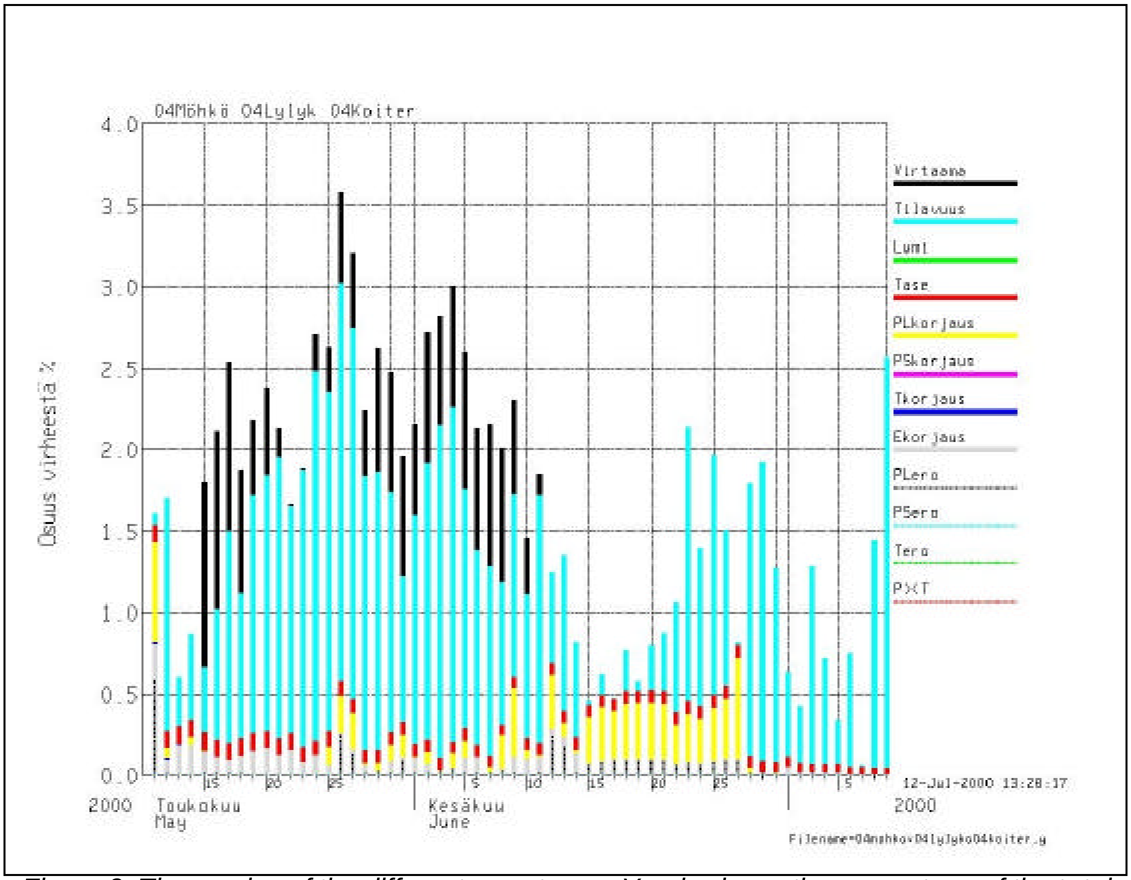


Figure 2. Time series of the different error terms. Y-axis shows the percentage of the total error. In the updating procedure12 different error terms are minimized. These error terms include errors in simulated discharge (Virtaama), volume (Tilavuus), amount of modifications in the precipitation (PL, PSkorjaus), temperature (Tkorjaus) and the areal variability of the correction terms.

4.Use of GIS data

An elevation model is available from Finland as is also land use and soil type data as grid based information covering the whole land area of Finland. Grid size of the elevation model and land use data is 25 mx25m and in soil type data it is 85mx85m. Accuracy of the altitude data in the elevation model is a few meters, which is enough in the estimation of areal precipitation and temperature, but it is not accurate enough for drawing maps of flooded areas. In the land use data the land use is classified into 50 different classes. In the soil type data the number of classes is 10.

In the latest version of the watershed model simulation is carried out in 1 km^2 grid size. From the GIS data the mean altitude and distributions of different land use and soil type classes is calculated for each grid. The simulation is carried out first by estimating daily precipitation and temperature for each grid and then the rainfall-runoff model (Figure 1) is simulated in each cell separately. The result of the simulation is the runoff from each grid cell, which is then collected to runoff from each sub-basin. The sub-basin runoff is routed forward using river models, as described earlier. The river model is also based on the GIS-data of the watershed.

The elevation model is used for estimation of areal precipitation and temperature. The daily precipitation and temperature of each grid cell is interpolated from the data of nearest observation stations. The estimated values are corrected by altitude difference between the stations and the grid cell. In the Finnish conditions the correction for precipitation is +10% and temperature –0.5°C per 100 m increase in altitude.

Under development is an interpolation method for precipitation that can take into account wind speed and direction together with gradient of landscape and distance from coast line. In Finnish conditions these terms have an effect especially on wintertime snow precipitation.

The land use data is used in the watershed model by dividing each grid cell into forest and open areas. Forest is still divided to different types: fir, birch and pine forest. Snowmelt in the different forest types and open area are simulated separately (Chapter snow model). The soil type data is not currently used in the watershed model but a model with a calibration procedure that could find similar parameter values for the runoff model for the basins with same soil type is under development.

5. Use of remote sensing data

The remote sensing data used in the model includes satellite data of snow coverage and precipitation data from weather radars. Precipitation data from weather radars is used in the Kyrönjoki basin 4800 km^2 in the Western Finland. The data from a weather radar, which covers the main parts of Kemijoki river basin with 60 000km^2 in Lapland is also in use.

The satellite data of snow coverage is used to estimate the accuracy of simulated snow water equivalent during snow melt period. So far there is no automatic method to use the information of the snow coverage from satellite to update the watershed model simulation. However there are two ongoing projects where the data assimilation methods are developed.

6. WWW-user interface

The WSFS can be run through a WWW-user interface (Figure 5), which allows one to add observations, change weather forecasts and regulation rules of lakes. These features are important when new observations are added to the system or regulation rules are updated during critical periods of floods. With the WWW-interface users from other institutes and companies can use the same model and thus only one model version has to be kept up. This reduces maintenance costs remarkable compared to a system where each user has own model version to run and to keep up. The user-interface contains a lot information of real time situation and forecast from the watersheds as figures, maps (Figures 3 and 4) and numerical data which can be used to make decisions of flood mitigation procedures, lake regulation or to use the simulated data in other simulation systems and registers at the sites and sub-basins from where no observations are available. The simulated WSFS data is further transferred to SYKE's main hydrological register for reuse. In the WSFS user-interface the forecasts accuracy evaluation is also available for some important catchments.

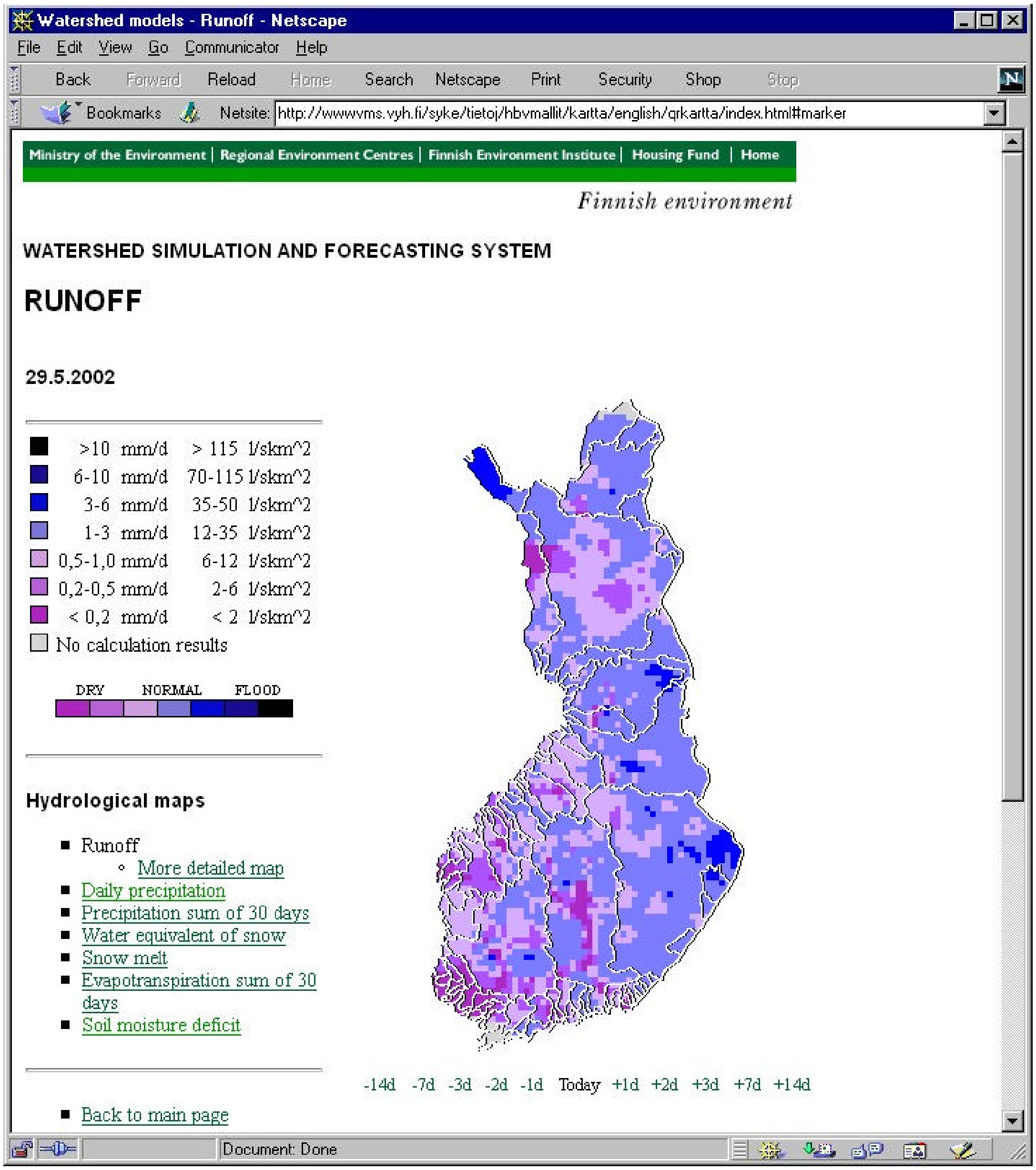
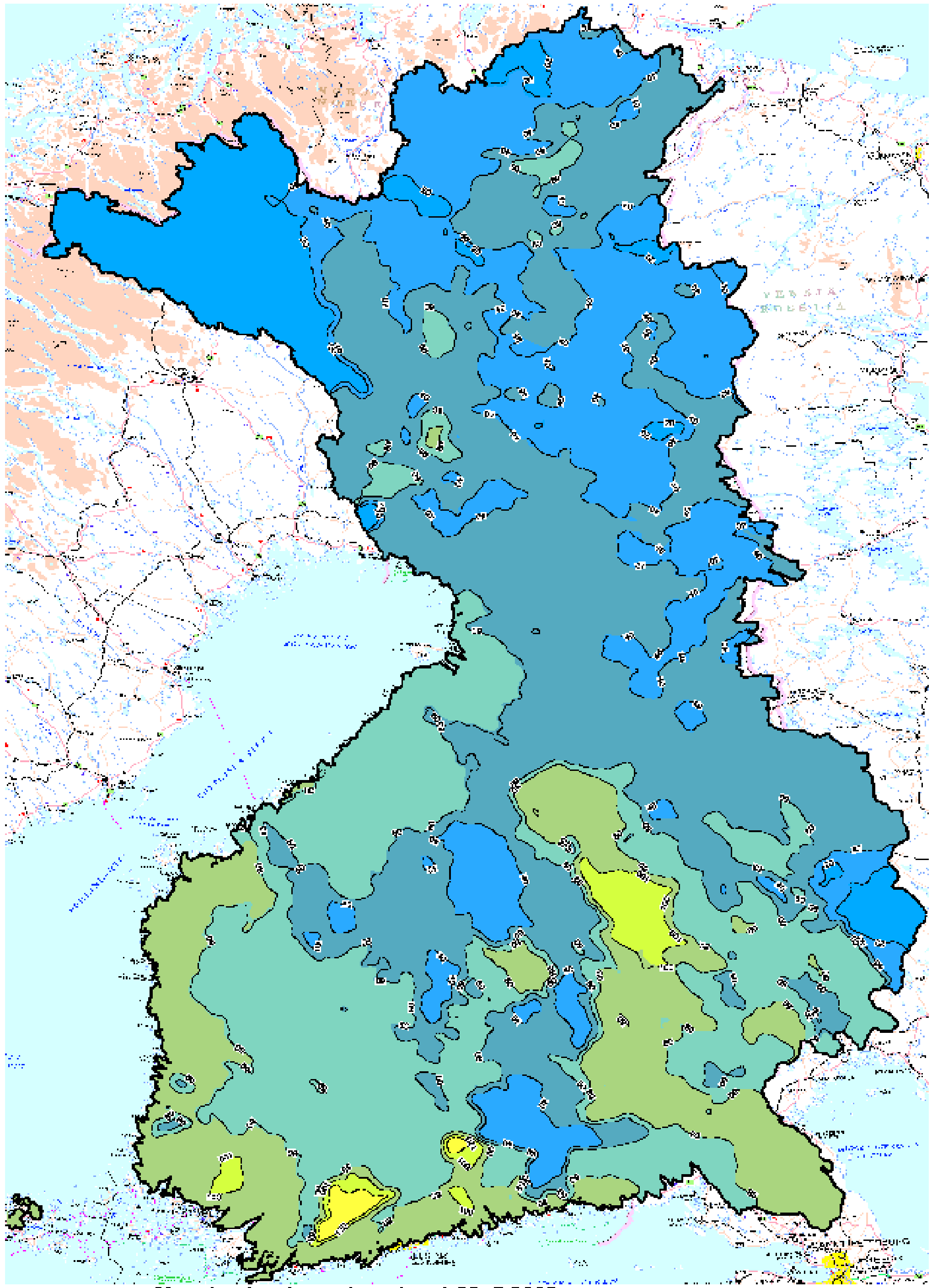


Figure 3. Internet products of WSFS: daily real time runoff maps.

Vesistömallit - FEI Watershed models



Maankosteuden vajaus (mm) 29. 5.2002
Forecast day 29. 5.
Soil moisture deficit (mm) 29. 5.2002

Figure 4. Internet products of WSFS: real time soil moisture (mm) deficit.

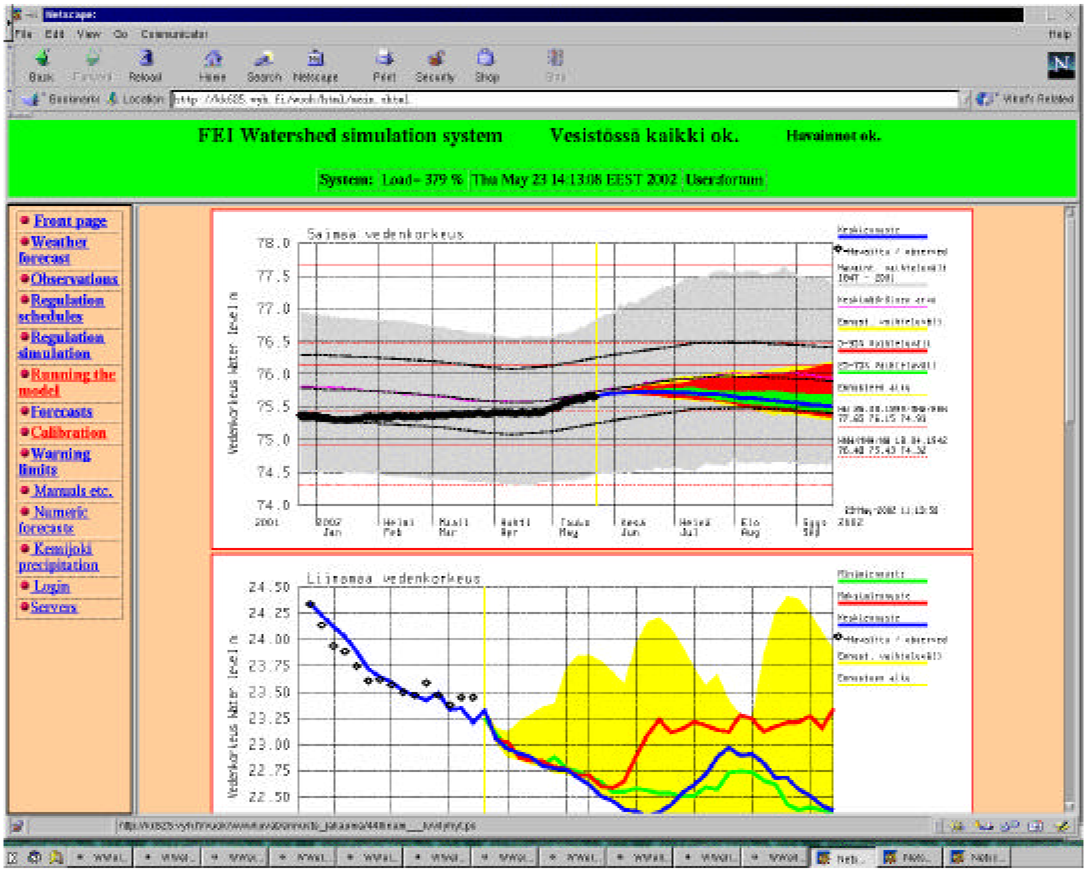


Figure 5. WWW user-interface of WSFS with operation choices at left and a selection of important water level forecasts.

7. References

Bergström, S. (1976) *Development and application of a conceptual runoff model for Scandinavian*. SMHI. Nr RH7. Norrköping.

Chui, D. K. and Chen, G. (1991) *Kalman filter with real-time applications*, Springer Series in Information Sciences, Vol. 17, Springer-Verlag

Forsius, J. (1984). *Computing unsteady flow and tracer movement in a river*. Publications of the Water Research Institute, National Board of Waters. Finland.

Hooke, R. and Jeeves, T. (1961) *Direct Search Solution of Numerical and Statistical Problems*. Journal of the ACM, Vol. 8, No. 2. p. 212-229.

Kuha, J. (1993) *Funktion arvojen vertailuun perustuva epälineaarinen optimointi*. Pro gradu-tutkielma. Tietojenkäsittelyopin laitos. Helsingin yliopisto.
61+18p. In finnish.

Vehviläinen, B. (1992) *Snow cover models in operational watershed forecasting*. Publications of Water and Environment Research Institute 11. Helsinki. 112 p.