A GIS embedded approach for Free & Open Source Hydrological Modelling

Massimiliano Cannata, Institute of Earth Sciences (SUPSI), Lugano, Switzerland

Maria Antonia Brovelli, Politecnico di Milano, Como, Italy

The project: Interreg IIIA (Switzerland-Italy)

Development of an hydrogeological risk management system in the Lake Maggiore area.
Managment System Components

General Circulation Model
- GCMs describe climate behavior by means of physically based atmospheric modelling and are used to produce meteorological forecasts or perform climatic studies.

Routing Models
- Solve mass balance equations to evaluate water discharges through and over the soil.

Land Surface Models
- Simulation of the water resource distribution at the ground with estimates of each water component quantity.

Hydrologic/Routing Models

Water resource Applications
- Water resource exploitation (intakes, dams, etc.) leads to a modified circulation model with economic, environmental and social goals.

Mesoscale Models
- Numerical weather prediction models and climate models. They work in a projected system, use observed data and are linked with other surface models. They usually provide short/medium term forecast.

Space-time
- Every physical phenomena is a direct function of its position in space and time.

The watershed system

Distributed: represents the hydrologic variables continuously in the space (raster maps);

Physically based: all the used variables have a physical meaning;

Continuous: generates estimations at consecutive time steps (hourly resolution);

Modular: assembled by means of the combination of different stand-alone sub-models simulating different rainfall-runoff involved processes.

Fully embedded: all the routines developed by using GIS commands or GIS libraries.

Open Source: generated by using only free and open source software.

system status ($t_0$) - input data - Processes simulation system status ($t_1$)
Data sources

- Meteorological stations (point wise observations of: temperature, humidity, wind speed, rainfall intensity, solar radiation)
- Meteorological forecasts (mesoscale models raster outputs T,H,W,R)
- Remote sensing (radiometric images > T,ET, ...)
- Geo-morphological maps (raster maps: digital terrain model, landuse, pedology)
- Parameters information (data from literature describing their behaviour on the basis of season, land use, and soil type: LAI, vegetation height, vegetation density, Albedo)

Call for GIS

- Different formats: (JPEG, BINARY, ARCGRID, TIFF, GIF, ASCII, ...)
- Different data modelling: (vector, raster, image, attribute)
- Different coordinate and reference systems: (CH1903LV05, UTM, GAUSS-BOAGA, ...)
- Different time resolution: (hourly, daily, monthly, ...)
- geo-database with spatial component
- storage
- spatialization
- simulation
### Geo-database with time component

#### Free & Open Source

- **PostgreSQL**
  - vector attributes
  - literature values, etc.

- **GRASS DB**
  - vector geometry
  - raster map

- **GRASS CORE**
  - GIS commands: visualization, interpolation, etc.

- **on-line stations database**
- **vector time series data format definition**
- **raster time series data format definition**
- **Disk space management**
  - (1 month ~ 2GB)

---

### Spatialization

**Temperature**: 5 model analyzed with leave-one-out cross-correlation approach:
1. INVERSE DISTANCE (idw)
2. SPLINES
3. POLINOMY OF DEGREE 3 IN X,Y PLUS ONE Z TERM
4. POLINOMY OF DEGREE 1 IN Z
5. POLINOMY OF DEGREE 2 IN Z

Following the MAE criteria the best model is the linear gradient (model 4) in Z

**Rainfall**: meteorological radar data validation with rain gauge observations by using semi-variograms, linear and non linear correlation indexes.

<table>
<thead>
<tr>
<th>RADAR</th>
<th>RAIN GAUGE</th>
<th>rain</th>
<th>no-rain</th>
<th>tot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>rain</td>
<td>no-rain</td>
<td>tot</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16218</td>
<td>6482</td>
<td>22700</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18059</td>
<td>210318</td>
<td>228368</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34268</td>
<td>216800</td>
<td>251068</td>
</tr>
</tbody>
</table>

The spatial component of radar data is quite good while the magnitude is not comparable with gauges data.
Processes

**INPUT MAP GENERATION**
- Meteorological maps from interpolation of climatic stations observations
- Watershed characterizing parameters maps: Spatio-temporal interpolation of literature values
- Surface, slope, aspect

**SNOWMELT AND ACCUMULATION**
- Abbot approach: simplified energy balance of the snowpack

**NET SOLAR RADIATION**
- Calculated in function of the date, the hour and the cloudiness

**POTENTIAL EVAPOTRANSPIRATION**
- Penman-Monteith approach: radiative term and aerodynamic

**RUNOFF**
- Topkapı approach: surface, sub-surface and channel flow with non linear storage equations

**CANOPY INTERCEPTION**
- Rutter approach: max interception capacity, drainage to the ground and evapotranspiration losses

**SNOWMELT AND ACCUMULATION**
- Abbot approach: simplified energy balance of the snowpack

**INPUT MAP GENERATION**
- Meteorological maps from interpolation of climatic stations observations
- Watershed characterizing parameters maps: Spatio-temporal interpolation of literature values

**Global radiation**

**Snow melt and accumulation**

**Evapotranspiration**

**Canopy interception**

**Runoff**

**Modules implementation**

New modules were implemented either by using GRASS libraries and Shell script for the sub-process simulation. The results were validated.
HydroFOSS

The general model (HydroFOSS) is assembled by a shell script linking and running in the correct sequence all the required modules and commands. (data access, spatialization, simulation, results storage)

- initialization -

Input:
DEM
LANDUSE
SOILTYPE
DATES (FROM, TO)
CONNECTION
TABLE NAME (PARAMETERS)

meteorological map derivation
thematic map derivation
global solar derivation
snow melt and accumulation
potential evapo-transpiration
canopy interception
net rain
real evapotranspiration
runoff

- loop for each hour -

Case study: the Verzasca basin

- The Verzasca river flows from the Alps to the Lake Maggiore and has a length of 34 Km with a mean discharges of 1 m³/s that, during floods events, can easily rise up to 250 m³/s.
- The drained area has an extension of 218 Km² and registers a mean annual precipitation of 1976 mm with a mean of 129 rainy days a year.
- The basin shows the typical Alpine characteristics with its high elevation ranges (from 193 m to 3596 m a.s.l.), narrow valley (2 Km wide), and high mean slope (33.93°).
Case study: basic data

- Swiss elevation model DHM25 map
- Modified "Arealstatistik 1972" landuse map
- Modified "Bodeneignungskarte" soil types map
- 30 on-line meteorological stations

Case study: derived data

- r.slope.aspect
- r.watershed
- r.mapcalc

- Aspect
- Basins
- Flow direction
- Flow accumulation
- River width
Case study: spatialized climatic data

- v.surf.idw (npoints=1)
- linear regression + r.mapcalc

Case study: the snow module

- snowpack water equivalent [mm]
- snowpack energy content [Kcal/m²]
- snowmelt [mm]
Case study: the snow module

- Rainfall
- Energy content [Kcal/m²]
- Water equivalent [mm]
- Snowmelt [mm]

Case study: the evapotranspiration module

- 6 May 2004 h: 12.00
  - Cloudy day
- 8 May 2004 h: 12.00
  - Sunny day
Case study: the canopy interception module

<table>
<thead>
<tr>
<th>canopy storage level [mm]</th>
<th>interception drainage [mm/h]</th>
<th>interception losses [mm/h]</th>
<th>rainfall [mm/h]</th>
</tr>
</thead>
</table>

Rainfall

Evaporation from liquid surface [mm/h]

Interception drainage [mm/h]

Interception losses [mm/h]

Canopy storage level [mm]
Case study: runoff initialization

- $K_s$ – the saturated hydraulic conductivity; $p(S_{Type})$ [*]
- $\Theta_s$ – the saturated soil water content; $p(S_{Type})$ [*]
- $\Theta_r$ – the residual soil water content; $p(S_{Type})$ [*]
- $\alpha$ – the soil type dependent parameter; $p(S_{Type})$ [*]
- $L$ – the soil thickness; $p(S_{Type})$ [*]
- $n$ – the Manning coefficient; $p(LU)$ [*]
- Basin – the basin index; (rast)
- Direction – the flow direction index; (rast)
- Accumulation – the flow accumulation value; (rast)
- Slope – the terrain slope; (rast)
- h_type – the hydrological type; (rast)
- width – the riverwidth; (rast)
- $L_{use}$ – the landuse category; (rast)
- $S_{type}$ – the soiltype category; (rast)

[*] --> parameters suitable for calibration

Case study: sensitivity analysis & calibration

Parameters

- SQL Update

Parameters

- initialization table
- runoff water
- actual evapotranspiration
- starting overland water volume
- starting soil water volume
- starting channel water volume
- timestep

hydroFOSS.runoff

overland water volume
soil water volume
channel water volume
starting water volume
soil water discharge
channel water discharge
saturated water excess

h.sample.series

estimates

observations

UCODE_2005
Case study: sensitivity analysis

Simulations are computationally expensive (~30 s/h).

In calibration phase multiple model running over long periods (months) are needed.

Collaboration with the Swiss National Supercomputing Centre (CSCS) to reduce computational time.

Composite Scaled Sensitivity (CSS)
The total amount of information provided by the observations for the estimation of the parameter $j$

(summarize all the sensitivities for one parameter)

<table>
<thead>
<tr>
<th>Manning</th>
<th>Ks</th>
<th>Teta_s</th>
<th>Teta_r</th>
<th>L</th>
<th>SAT</th>
</tr>
</thead>
</table>

Dimensionless Scaled Sensitivity (DSS$_i$)
The importance of the parameters $j$ to the calculation of an observed value $i$
• During the recession limb of the hydrograph the soil conductivity is the most important parameter.
• During the rising limb of the hydrograph the soil thickness is the most important parameter.
• During the regular flow the most important parameter is the river Manning coefficient.

Leverage
measures the influence of the observation $i$ on the regression due to its location in the space of the inputs

observations at the beginning of the rising limb are the most important in the good model fitting
Case study: runoff calibration

calibration process upon 8 parameters:
 n_11, ks_2, ks_3, ks_4, L_2, L_3, L_4, and SAT

In order to reduce the iterations, only the most important parameter of ks and L set was directly calibrated. The others parameters were proportionally scaled.

The calibrated parameters values are:
- river Manning coefficient 2.3E-02 [s*m^{-1/3}];
- regosol soil conductivity 0.678E-03 [m*s^{-1}];
- regosol soil thickness 0.374 [m];
- starting soil saturation 7.6 [%];

channel water discharge [m^3/s]
A: before the event (17/08/2004, 16:00).
B: during the peak (20/08/2004, 06:00).
C: during the recession (20/08/2004, 16:00).

overland water discharge [m^3/s]
A: before the event (17/08/2004, 16:00).
B: during the peak (20/08/2004, 06:00).
C: during the recession (20/08/2004, 16:00).

subsurface water discharge [m^3/s]
A: before the event (17/08/2004, 16:00).
B: during the peak (20/08/2004, 06:00).
C: during the recession (20/08/2004, 16:00).
Conclusions

Geographical Information Systems are projected in the era of modelling in order to allow a conscientious environmental management of the resources.

- The GIS capability to model the rainfall-runoff has been proven by means of the development of procedures and new commands simulating all the physical processes driving the water flows (HydroFOSS model).

- A GIS embedded approach establishes a valuable resource for solving numerous management problems.

- Free and Open Source (FOSS) GIS Software, largely facilitates the task of new procedures and model development due to the accessibility of its code, and availability of documentation.

- The followed modular approach, in conjunction with the standard GIS commands and the accessibility of the code, guarantees, to an expert user, to fully customize the application.

- A new standard model for time series data has to be defined as a key point in GIS embedded modelling.

Future work

- Definition of a standard time series data format (raster and vector).

- Development of new modules (for a lumped model and different time resolution).

- Validation of the single developed modules (lumped and distributed).

- Development of data pre-processing procedures.

- Further sensitivity analysis & calibration of the models and comparison with commercial models.

- Development of Web Interface for a distributed user friendly model usage (GRASS+PHP+AJAX).

- Writing of a user manual book.

- Code optimization to speed up processing.
## References

<table>
<thead>
<tr>
<th><a href="http://w3.ist.supsi.ch/geomatica/">SUPSI SmartGEOSS project</a></th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://grass.itc.it/">http://grass.itc.it/</a></td>
</tr>
<tr>
<td><a href="http://www.postgresql.org">http://www.postgresql.org</a></td>
</tr>
</tbody>
</table>

### Contacts:

- massimiliano.cannata@supsi.ch
- maria.brovelli@polimi.it