The JGrass-NewAge system for forecasting and managing the hydrological budgets at the basin scale: the models of flow generation, propagation, and aggregation

G. Formetta$^1$, R. Mantilla$^2$, S. Franceschi$^3$, A. Antonello$^3$, and R. Rigon$^1$

$^1$University of Trento, 77 Mesiano St., Trento, 38123, Italy
$^2$The University of Iowa, C. Maxwell Stanley Hydraulics Laboratory, Iowa 52242-1585, USA
$^3$Hydrologis S.r.l., Bolzano, BZ, Italy

Received: 16 April 2011 – Accepted: 22 April 2011 – Published: 29 April 2011

Correspondence to: G. Formetta (formetta@ing.unitn.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

This paper presents a discussion of the predictive capacity of the first implementation of the semi-distributed hydrological modeling system JGrass-NewAge. This model focuses on the hydrological balance of medium scale to large scale basins, and considers statistics of the processes at the hillslope scale. The whole modeling system consists of six main parts: (i) estimation of energy balance; (ii) estimation of evapotranspiration; (iii) snow modelling; (iv) estimation of runoff production; (v) aggregation and propagation of flows in channel, and (vi) description of intakes, out-takes, and reservoirs. This paper details the processes, of runoff production, and aggregation/propagation of flows on a river network. The system is based on a hillslope-link geometrical partition of the landscape, so the basic unit, where the budget is evaluated, consists of hillslopes that drain into a single associated link rather than cells or pixels. To this conceptual partition corresponds an implementation of informatics that uses vectorial features for channels, and raster data for hillslopes. Runoff production at each channel link is estimated through a combination of the Duffy (1996) model and a GIUH model for estimating residence times in hillslope. Routing in channels uses equations integrated for any channels’ link, and produces discharges at any link end, for any link in the river network. The model has been tested against measured discharges according to some indexes of goodness of fit such as RMSE and Nash Sutcliffe. The characteristic ability to reproduce discharge in any point of the river network is used to infer some statistics, and notably, the scaling properties of the modeled discharge.

1 Introduction

Hydrological forecasting over time has focused on different issues. Determining the discharge of rivers during flood events has been a central topic since more than a century; firstly through the rational model of Mulvaney (1851), later through the use of instantaneous unit hydrograph models (Sherman, 1932; Dooge, 1959), and more
Models developed to reproduce a whole set of hydrological quantities for operational purposes came from water resource management and agriculture needs. In this context were produced the implementations of large modeling systems, of which the progenitor is the Stanford watershed model (Crawford and Linsley, 1966), like the Sacramento model (e.g. Burnash et al., 1973), and the PRMS model (Leavesley et al., 1983). They were based on the metaphor of intercommunicating compartments (reservoirs), each representing a process domain, each one with its proper residence time. The recent “Distributed Model Intercomparison Project”, DMIP, (Reed et al., 2004), revealed some of the many differences among the models reported above, and provided a first set of tentative comparisons. Despite the major emphasis of the project was still in reproducing discharges, a more marked attention to the prediction of the entire hydrograph, instead of only the peak of the hydrograph was evident: a necessary element for the overall management of basins and particularly for the management of droughts.

To see the topic from a different point of view, in literature there exist an even larger variety of models, with varying degrees of complexity, and simplifications. The two extremes of modeling are given by fully distribute models (for a recent review, see Kampf and Burges, 2007 and Rigon et al., 2006), and lumped models (e.g. Beven, 2001). In the first class, the physics is modeled at grid (pixels) level using the fundamental laws of conservation of energy, mass, and momentum, in the second, the ruling equations are simplified in order to obtain some statistics of the hydrological budget without unnecessary representation of the full spatial variability.

A solid paradigm of simplification is offered by the theory of the geomorphological unit hydrograph which provides flow values at a single point of the river network (i.e. at the outlet of the basin). In this case, many models with few parameters are able to reproduce the expected result with an acceptable degree of confidence. This is
granted because the outlet discharge is an additive stochastic process (e.g., Rinaldo and Rodriguez-Iturbe, 1996), in which the topology and the geometry of the river network is more important than the details of the local dispersive dynamics (e.g., Rinaldo et al., 1991). In addition Leopold and Maddock (1953) observed that the overall action of hydrological and geomorphological forces act in maintaining approximately constant the flow velocity. This simplification is not appropriate when spatial prediction are required (e.g. discharge at intranet location). To this end it is necessary to make use of detailed information on topography (as derived from modern LiDAR or SAR sensors), and a large variety of remote sensed information, which provide new tools for representation of the physics of flow transport along the channels of the river network and processes into the hillslopes.

The JGrass-NewAge model (Franceschi et al., 2011) was conceived and structured to meet these demands, to forecast not only floods, but also of droughts, to calculate the water balance at several points in the river network of a basin, and to provide statistics revealing the internal (spatio-temporal) variability of some of the quantities analyzed. To obtain this, the model implements innovative informatics, which is described in Antonello et al. (2011), to allow modifications of its parts and parameterizations without changing the whole, and therefore supporting the comparison of different schemes of simplification, and of the parametrization of hydrological processes.

To achieve this, the model partitions the basin into hillslopes and channels (giving to the model a hillslope-link, HL, structure), where the hillslopes are the basic hydrologic units. It is at this scale that the energy and water mass budgets statistics are calculated. The channels are described as vector elements (features) that are topologically interconnected in a simple directed graph. This concept could be confounded with the concept of hydrological runoff units (HRUs) promoted in Ross et al. (1979), Flügel (1995), and used, for instance in Krause (2002), and in Viviroli et al. (2009). Any HRU instead is derived from the intersection of various classes of superimposed information layers, and are clusters representing areas of the basin where similar hydrological behavior is expected, while in JGrass-NewAGE the HL structure is derived
by the watershed delineation operation, and represents the set of flow lines that converge to an outlet and/or an outlet cross section. Thus HRUs can be seen as sub-partitions of the HL, and these sub-classes provide in JGrass-NewAge statistics at hillslope (or small watershed) level, instead that single estimations of the hydrological quantities. These sub-partitions (and the relative sub-parameterizations, when applicable) are process dependent, as, for instance, the hillslope HRUs for evapotranspiration could be different from the ones for snow modeling. For computational reasons, the partitioning of the area is not usually designed to identify all the hillslopes physically present in the system, but to define small watersheds with dimensions, in the current application, to 2–3 km$^2$ on average.

Any element of the river network is represented in the model as a vectorial entity (OGC feature), connecting the hillslopes. The network includes anthropogenic infrastructure that regulates the flow regimes and thus it is possible to simulate urban discharge inputs and outputs, managements of dams, artificial channels and irrigation water withdrawals. The way in which all of these elements are implemented in the JGrass-NewAge system includes the definition of a topological hierarchy based on a modification of the Pfafstetter’s ordering scheme (Verdin and Verdin, 1999; de Jager and Vogt, 2010). HRUs instead can be either treated as vectorial features, or rasters, according to convenience.

Elements that deal with hillslope runoff production and its aggregation in the channel network are described below, and cover the part of the system called “Adige” from the name of the river on which the model has been applied for the first time.

## 2 JGrass-NewAGE runoff production, aggregation, and routing

The “Adige” component is made up of three part: a model of runoff generation, a model of runoff propagation in hillslope, and a model for routing in channels. The model for runoff-generation used in JGrass-NewAge follows, with minor variations, Duffy (1996). Basically, it considers the integration of the equation of continuity on the partial volumes
of a hillslope occupied by the saturated and unsaturated portions of a hillslope as a dynamic system in two state variables: the volume stored per unit area in the saturated, $S_1 \ (L)$, (where (L) means a length such as mm) and unsaturated, $S_2 \ (L)$ storages. This is common to many other models, including (Castelli et al., 2009; Majone et al., 2010). Differently from most of the models, the equations used are linear and those deemed non-linear functions are approximated using a second order polynomial (Duffy, 1996). The equations of the two reservoirs are coupled and generate runoff, while the runoff routing itself, is described by a simple model of residence times (Rinaldo and Rodriguez-Iturbe, 1996). The runoff produced by each hillslope, kinematically propagated downhill, is then propagated in channels through a simplified model, derived from the CUENCAS model (Mantilla and Gupta, 2005), essentially a non linear variant of the Saint Venant equation (e.g. Bras and Rodríguez-Iturbe, 1985), model, which represents the flow equation in each channels’ link. The resulting system of equations allows an estimate of the varying discharge value in each link of the river network.

### 2.1 Runoff generation

Partitioning the soil into an unsaturated fraction, identified by the subscript 1, and a saturated fraction, subscript 2, and considering the boundary as a moveable front separating the two soil layers, the continuity equation defines the following system represented schematically in (Fig. 1) Eqs. (1) and (2):

$$\frac{dS_1}{dt} = f_{01} - g_{12} - f_{10}$$

$$\frac{dS_2}{dt} = f_{02} + g_{12} - f_{20}$$

where $S_1 \ (L)$ and $S_2 \ (L)$ represent the volume of water in unsaturated and saturated soil fractions per unit area of hillslope; $f_{01} \ (L \ T^{-1})$ and $f_{10} \ (L \ T^{-1})$ are the inbound and outbound flows between the unsaturated soil volume and channels; $f_{02} \ (L \ T^{-1})$ and
\( f_{20} \ (L \cdot T^{-1}) \) are input and output flows between the saturated soil volume and the surrounding environment respectively; \( g_{12} = g_{21} \ (L \cdot T^{-1}) \) represents the flow between the two soil volumes.

The analytical relationships used to express the terms of the functions introduced in Eqs. (1) and (2) are as follows:

\[
\begin{align*}
    f_{01} & = p \cdot [1 - d_4(S_2 - S_2^0)] \\
    f_{00} & = p \cdot A_s = d_4(S_2 - S_2^0) \\
    g_{12} & = d_0(S_1 - S_1^0) + d_1(S_1 + d_2)(S_2 - S_2^0)^2 \\
    f_{20} & = d_3(S_2 - S_2^0)
\end{align*}
\]

where \( p \ (L \cdot T^{-1}) \) is the precipitation, \( S_{01} \) and \( S_{02} \ (L) \) are the water volumes per unit area in the unsaturated and saturated part respectively, \( A_s \ (L \cdot T^{-1}) \) is the hillslope fraction of area that reached saturation, and \( f_{00} \) is the part of precipitation that goes into runoff. The terms \( d_i \) are positive parameters. Note that \( g_{12} \), exchanging between \( S_1 \) and \( S_2 \), contains a nonlinear term which couples the two state variables, while the external flows \((f_{01} \) and \( f_{02} \)) are linear functions. The relationship between the saturated area and water volume in the saturated system is linear. The model represents the mechanism of runoff production because saturation excess in which \((p - f_{01})\) represents the surface runoff in the saturated surface area. The term \( f_{20} \) represents the runoff generated from the saturated part of the terrain, the larger the reservoir, the greater the amount of runoff \( (i.e. \ value \ of \ f_{20}) \). The parameters \( S_{01} \) and \( S_{02} \) are, as already mentioned, the excess water volumes in case of a completely drained hillslope \( (\text{due to gravity and no precipitation, i.e. } p = 0) \). The parameters \( d_0, d_1, d_2 \) describe the shape of the charging hillslope relationship that combines the two state variables. The parameters which describe the external flows all have a physical meaning: \( d_3 \) is the subsurface flow constant, while \( d_4 \) correlates the saturated surface area \((A_s)\) with the saturated volume of soil.
2.2 Time lag of hillslope runoff delivery

The kinematical lag in surface and subsurface flows generated by the hillslope and flowing into a generic network element can be neglected only in the case of very large basins (e.g. D’Odorico and Rigon, 2003; Botter and Rinaldo, 2003). Thus, in JGrass-NewAge, the kinematic lag in flows, caused by different water velocities along the hillslopes and in the channels, is calculated according to the instantaneous unit hydrograph theory and by defining:

\[ q_{\text{sup}} = \int \text{IUH}_{\text{sup}}(t-\tau) \cdot f_{00}(\tau) \, d\tau \]  \hspace{1cm} (7)

\[ q_{\text{sub}} = \int \text{IUH}_{\text{sub}}(t-\tau) \cdot f_{20}(\tau) \, d\tau \]  \hspace{1cm} (8)

where \( \text{IUH}_{\text{sub}}(t) \) and \( \text{IUH}_{\text{sup}}(t) \), \( (L \cdot T^{-1}) \), are the instantaneous unit hydrographs which vary in time as a result of surface and subsurface contributions. In these calculations, the velocity of the two flows is different: usually less than 0.4 m s\(^{-1}\) for the surface flow, and is kept constant for a hillslope, while the mean velocity for subsurface flows is based on calculations using Darcy’s law average according to water paths:

\[ v_{sb} = \overline{K_s} \cdot \overline{\nabla z} \]  \hspace{1cm} (9)

where \( \overline{K_s} \) \( (L \cdot T^{-1}) \) is the average hydraulic conductivity and \( \overline{\nabla z} \) \( (L) \) is the average slope of the hillslope, and therefore \( v_{sb} \) is made dependent on slope. In practice \( \overline{K_s} \) is considered as a subsurface flow calibration parameter. The IUH, are estimated, on the base of the rescaled width function (Rinaldo and Rodriguez-Iturbe, 1996; D’Odorico and Rigon, 2003), considered variable in time, as a function of the total saturated area, \( A_s \), of any hillslope. To simplify the calculations, the rescaled width function is not used.
directly but interpolated using (Nash, 1957) hydrograph with the appropriate parameters:

\[
IUH(t) = \frac{1}{k \cdot \Gamma(n+1)} \left( \frac{t}{k} \right)^n e^{-\frac{t}{k}}
\]  

(10)

where \( k (T) \) is the mean residence time along the hillslopes and \( n \), is a second calibration parameter. In JGrass-NewAge, the mean and variance of Nash distribution are derived from the mean and variance of the width function for each basin under different saturation levels (different values of \( A_s \)), once surface and subsurface runoff velocities have been assigned. The parameters to be included in Eq. (10) are:

\[
k = \frac{\sigma^2}{\mu}
\]  

(11)

\[
n = \frac{\mu^2}{\sigma^2} - 1
\]  

(12)

where \( \sigma^2 \) and \( \mu \) represent the variance and mean of the distribution of residence times for each hillslope for a given level of saturation.

2.3 Runoff aggregation

The flow generation model along hillslopes delivers discharge to the channel network conceptualized in the model as a oriented tree graph. For each link the continuity equation is:

\[
\frac{dQ_i(t)}{dt} = K \left( Q_i(t) \right) \cdot \left[ A \cdot R_i(t) + \sum_{\text{trib}} Q_{\text{trib}}(t) - Q_i(t) \right]
\]  

(13)

\( i = 1, 2, \ldots, H \)

where, \( H \) is the total number of network links, \( Q_i(t) \ (L \ T^{-1}) \) is the output discharge from \( i \)-th link, \( K \left( Q_i(t) \right) \ (T^{-1}) \) is the Chezy coefficient, \( R_i(t) \ (L \ T^{-1}) \) is the runoff intensity
per unit area from the upstream hillslopes, \( Q_{\text{trib}} \) \( (LT^{-1}) \) is the flow of upstream links, and \( A L^2 \) is the drainage area of the link in question. The analytical form of the Chezy propagation coefficient implemented in JGrass-NewAge is:

\[
K_Q = \frac{3}{2} \cdot Q_{\text{sp}}^{\frac{1}{3}} \cdot C^{\frac{2}{3}} \cdot b^{\frac{1}{3}} \cdot L^{-\frac{1}{3}} \cdot i_f^{\frac{1}{3}}
\]  

(14)

where \( C \) \( (L^{1/3}T^{-1}) \) is the Chezy coefficient, \( b \) \( (L) \) and \( L \) \( (L) \) represent the width and average length of the link respectively, \( i_f \) \([-]\) is the average slope of the link, and \( Q \) \( (LT^{-1}) \) is the channel discharge. For a more detailed discussion of the terms in Eq. (14) see Menabde and Sivapalan (2001) and Mantilla et al. (2005).

3 An application to river passer

In the following section, we present a preliminary application of the JGrass-NewAge model in the Passer sub-basin, a tributary of the Adige River and shown in (Fig. 2). Passer River was chosen due to the fact that it contains the only hydrometer recording data not affected by water intakes, discharges or withdrawals, but solely due to natural flows. Passer sub-basin, situated in the nord-est of the Adige river basin with outlet to Bozen (Fig. 5), has a drainage area of 350 \( \text{km}^2 \), the minimum and maximum elevation are respectively about 500 (m) and 3100 (m) and in (Fig. 6) the hypsographic curve is shown.

The setting of the JGrass-NewAge infrastructure is performed for the entire Adige basin with outlet to Bozen. The basin is divided in 588 subbasins and a geomorphological analysis is performed using tools implemented in the GIS JGrass.

The relationship Area \((A)\)–Perimeter \((P)\) of each subbasin in a logarithmic plot is shown in (Fig. 3); the result of the linear regression is: \( P \sim A^{0.489} \). In (Fig. 4) it is shown the relation between the means and the variances of slope and total contributing area of each hillslope. Table 1 contains the values of the linear correlation coefficient and intercept and angular coefficient for both means and variances values.
In the first application, shown in (Fig. 7), the simulation period ranged from 1 October 2008 to 31 March 2009 (i.e. during the winter); in summer, the model was tested for the period 01 June 2009 to 31 July 2009. In both the cases were used data provided by the Adige River Basin Authority.

Calibration was performed manually in two steps: first a sensitivity analysis was performed and then, by selecting the most influential parameters, we proceeded to change them simultaneously to minimize the deviation between measured and simulated discharge, according the root mean squared error (RMSE). The model was calibrated for the first month of the simulation, and after estimating “optimal” parameters, was applied for the next five months. Evapotranspiration was considered constant during the period of interest. The calibration period, characterized by the alternation of rainy days and days with no precipitation, as shown in (Fig. 7), allows the calibration to cover a range of different meteorologic conditions. In summer, similarly to the winter simulation, the calibration was performed using the first month of data while the remaining months were simulated using the parameters estimated following the calibration. Figure 8 shows the summer simulation results and the corresponding precipitation hyetograph: the solid curve represents the measured discharge while the dashed line represents the model simulation data.

Observing visually the simulation seems to show some underestimation of the discharge. This can probably derive from underestimation of net precipitation, especially caused by a lack of measurement at higher elevations, and is particularly evident in the forecast for events at the end of July 2009, where volumes of discharges tend to be considerably greater than the precipitation (as derived from spatial extrapolation of measurements).

To perform an overall quantitative assessment of the predictions, three index of goodness of fit were used: the index of Nash Sutcliffe (NS) (Nash and Sutcliffe, 1970), the index of agreement IOA (Willmott et al., 1985), and the percentage model bias (PBIAS). The result are shown in Table 2 and summarized below.
The NS, measure of the ratio of the model error to the variability of the data, assumes values in the range \((-\infty, 1]\); NS < 0 would indicate that the mean value of the observed discharges would have been a better predictor than the model outcomes, and NS = 1 indicates the perfect fit. Taking in account of the manual calibration and of the errors also in measurements the NS values in both the simulation, NS = 0.79 for the winter simulation and NS = 0.78 for the summer, can be considered satisfactory, considerably better that the mean flow, and around the same value found in similar conditions with the state-of-art-models.

The IOA lies between 0 (no correlation) and 1 (perfect fit) and represents the ratio of the mean square error and the potential error (largest value of the squared difference of each pair). With respect to IOA the performance of the model are quite good: IOA = 0.938 for the winter simulation and IOA = 0.921 for the summer one. Since, the indicator is more sensitive to the peaks of the hydrograph, its value suggests that the winter simulation peaks are a little better fitted than in summer.

Finally, the PBIAS gives a measure of whether the model is systematically underestimating or overestimating the observations. According to Marechal (2004) |PBIAS| < 10 gives excellent model performance (summer simulation) and 10 < |PBIAS| < 20 gives very good model performance. In the present case, the performance of the model can be considered good.

As shown in (Fig. 9) the model is able to provide the hydrographs in each link therefore in the inner points of the river network.

4 Conclusions

The novel idea behind JGrass-NewAge is to provide not only a new hydrological tool but an informatics’ infrastructure in which any model can be built in components that can be independently modified or changed. This paper presents the concepts behind the runoff-routing components of the system, and a validation of the model against a year of measurements available. These were assessed with the use of three indicator statistics.
The results of the current study demonstrate the ability of the JGrass-NewAge model to reproduce observed flow discharges during both dry and wet periods, even if not with the same set of parameters.

Besides on structural model defects, some discrepancy between simulated and measured discharge can be related to error in rainfall measurements, as follows from comments in the text.

Variation of parameters between summer and winter which was necessary to obtain reasonably good results, can be considered as a consequence of variation of hydraulic conductivity (depending on temperature), neglecting evapotranspiration, and outflow from glaciers, which were kept constant, in the present work, since there was no way to assess their influence with measures.

The structure of the model which allows for forecasting the discharge in any link end inside the river network allows for producing statistics of simulation that can be compared with those deriving from hydrological regional studies, and shows that a model like JGrass-NewAge can potentially provide that information that it is normally obtain just on a statistical basis.

Acknowledgements. This project was financed by the ASI MORFEO project that supported one of the authors.

References


Beven, K.: Rainfall-runoff modelling: the primer, John Wiley & Sons Ltd, The Atrium. Southern Gate, Chichester, West Sussex PO19 8SQ, UK, 2001. 945

Bras, R. and Rodríguez-Iturbe, I.: Random functions and hydrology, Addison-Wesley, Reading, Massachusetts, 1985. 948
Krause, P.: Quantifying the impact of land use changes on the water balance of large catch-
Mulvaney, T.: On the Use of Self-Registering Rain and Flood Gauges in Making Observations of the Relations of Rainfall and of Flood Discharges in a Given Catchment, Transactions and Minutes of the Proceeding of the Institute of Civil Engineers of Ireland, 1851. 944
Ross, B., Contractor, D., and Shanholtz, V.: A finite-element model of overland and channel
Table 1. Linear correlation coefficient and linear regression analysis results for slope-tca relation.

<table>
<thead>
<tr>
<th></th>
<th>L.C.C.</th>
<th>Ang. Coeff.</th>
<th>Intercepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope-Tca Mean</td>
<td>−0.9396</td>
<td>−5.57</td>
<td>−0.7031</td>
</tr>
<tr>
<td>Slope-Tca variance</td>
<td>−0.8487</td>
<td>−11.35</td>
<td>−33.68</td>
</tr>
</tbody>
</table>
Table 2. Index of goodness of fit for winter and summer simulation.

<table>
<thead>
<tr>
<th></th>
<th>IOA</th>
<th>NS</th>
<th>PBIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter simulation</td>
<td>0.938</td>
<td>0.789</td>
<td>13.344</td>
</tr>
<tr>
<td>Summer simulation</td>
<td>0.921</td>
<td>0.781</td>
<td>6.452</td>
</tr>
</tbody>
</table>
Fig. 1. Runoff generation in JGrass-NewAge System.
Fig. 2. Representation of the Adige River Basin (shown in light blue) and the Passer River sub-basin (shown in dark blue) with the outlet located at Saltusio.
Fig. 3. Areas perimeters relation in a log-log plot.
Fig. 4. Hillslopes Area – slope relationship in a log-log plot: mean (on the left) and variance (on the right).
Fig. 5. The Passer river basin.
Fig. 6. The Passer river basin: hypsographic curve.
Discharge October 2008 - March 2009

Precipitation October 2008 - March 2009

Fig. 7. Application of the JGrass-NewAge model for the period 10 January 2008 to 31 March 2009 (winter): the solid curve represents the measured discharge, while the dashed line represents the model simulation data.
Fig. 8. Application of the JGrass-NewAge model for the period 01 June 2009 to 30 July 2009 (summer): the solid curve represents the measured discharge, while the dashed line represents the model simulation data.
Fig. 9. Application of the JGrass-NewAge model for the period 10 January 2008 to 31 March 2009 (winter): hydrographs in each link of the river basin.