Hydrological model assessment for flood early warning in a tropical high mountain basin

María Carolina Rogelis[1], Micha Werner[1, 2], Nelson Obregón[3], and Nigel Wright[4]

[1]UNESCO-IHE, PO Box 3015, 2601DA Delft, The Netherlands
[2]Deltares, PO Box 177, 2600MH Delft, The Netherlands
[3]Universidad Javeriana, KR 7 No 40-62, Bogotá, Colombia
[4]School of Civil Engineering, University of Leeds, Leeds LS2 9JT, UK

Correspondence to: Maria Carolina Rogelis
(c.rogelisprada@unesco-ihe.org)

Abstract. A distributed model (TETIS), a semi-distributed model (TOPMODEL) and a lumped model (HEC HMS soil moisture accounting) were used to simulate the discharge response of a tropical high mountain basin characterized by soils with high water storage capacity and high conductivity. The models were calibrated with the Shuffle Complex Evolution algorithm, using the Kling and Gupta efficiency as objective function. Performance analysis and diagnostics were carried out using the signatures of the flow duration curve and through analysis of the model fluxes in order to identify the most appropriate model for the study area for flood early warning. The impact of varying grid sizes was assessed in the TETIS model and the TOPMODEL in order to chose a model with balanced model performance and computational efficiency. The sensitivity of the models to variation in the precipitation input was analysed by forcing the models with a rainfall ensemble obtained from Gaussian simulation. The resulting discharge ensembles of each model were compared in order to identify differences among models structures. The results show that TOPMODEL is the most realistic model of the three tested, albeit showing the largest discharge ensemble spread. The main differences among models occur between HEC HMS soil moisture accounting and TETIS, and HEC HMS soil moisture accounting and TOPMODEL, with HEC HMS soil moisture accounting producing ensembles in a range lower than the other two models. The ensembles of TETIS and TOPMODEL are more similar.

1 Introduction

Models constitute the heart of early warning systems, providing a description of the hazard and its evolution in time (Basher, 2006). Hydrologic and hydrodynamic models with varying levels of complexity are used to provide advance warning of the likely timing and magnitude of flooding, and to help to understand the complexities of a flood event as it develops (Sene, 2008). A key aspect is to ensure that all relevant hydrological processes are included, and that appropriate computational weight is given to each process on the basis of its relative importance (Clark et al., 2008). This task is highly complex, since different models represent hydrologic processes differently, and all of them are imperfect (Duan et al., 1992).

Hydrologic modelling is affected by four main sources of uncertainty: input uncertainty, output uncertainty, structural uncertainty and parametric uncertainty (Renard et al., 2010). Structural uncertainty is defined as the modelling uncertainty due to

the selection of an appropriate model, which includes the defined hydrological processes (perceptual model) and description of these processes (conceptual model) (Zhang et al., 2011), and their mathematical implementation. Uncertainty induced by model structures can be more significant than parameter and input data uncertainty, but such uncertainties are difficult to assess explicitly or to separate from other uncertainties during the calibration process (Beven and Binley, 1992). The identification of the most appropriate model and model structure and its associated uncertainty to be implemented in a flood forecasting system is crucial, since the acceptable reproduction of hydrological processes builds up reliability into the hydrological model. This is essential when the model is to be used for forecasting and extrapolation (Reusser, 2010), where getting the "right answers for the right reasons" (Kirchner, 2006) or realism (Kavetski and Fenicia, 2011) is an important component of the confidence of the forecasting system. However the range of schemes available for assessing the impact of model structures on modelling uncertainty is still quite limited (Zhang et al., 2011).

The suitability of a rainfall-runoff model structure for a certain catchment has recently been studied through the use of flexible hydrological model structures, which focus on the diagnosis of their differences (Clark et al., 2008). These flexible hydrological model structures include: the Framework for Understanding Structural Errors (FUSE) introduced by Clark et al. (2008); the SUPERFLEX modelling framework proposed by Fenicia et al. (2011) that develops the earlier FLEX model (Fenicia et al., 2008); and the Framework for Assessing the Realism of Model Structures (FARM) proposed by Euser et al. (2013), where consistency and performance are analysed through principal component analysis. The criteria to be used for model evaluation both in these frameworks and in standard calibration procedures are an active research topic. Metrics such as the Nash–Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) or the root mean square error (RMSE) are often used to evaluate simulation results. However, their drawbacks (Pfannerstill et al., 2014; Fenicia et al., 2007) call for a more comprehensive approach. The use of vector search techniques to optimize model parameters is an alternative to incorporate multiple criteria within calibration to provide a number of alternative parameter sets that are optimal, on the basis of the Pareto-dominance concept (Efstratiadis and Koutsoyiannis, 2010). Fenicia et al. (2007) compared a pareto-optimality based calibration approach with a procedure that replicates the steps that are undertaken during manual calibration finding that given their strengths both calibration approaches can be combined. Other approaches rely on signature measures (Pfannerstill et al., 2014; Yilmaz et al., 2008) that define the hydrologic response characteristics and provide insight into the hydrologic function of catchments (Sawicz et al., 2011), rather than assessing model performance solely on the discharge at the outlet.

This paper explores the suitability of three differing model concepts to be used for flood forecasting purposes in a basin located in Bogotá (Colombia). The aim of the research is to explore the performance of the models in order to identify the most appropriate modelling approach, given the characteristics of the study area. A lumped model (HECHMS Soil Moisture Accounting), a semi-distributed model (TOPMODEL) and a distributed model (TETIS) were used. In the case of the semi-distributed and distributed model, resolution was explored in order to identify the most suitable pixel size to be used. Finally, a comparison of precipitation input uncertainty and model performance is carried out in order to identify the importance of these in the modelling results, which constitutes relevant information for future improvement to the models.

The study area exhibits a high degree of complexity, since the upper basin corresponds to a páramo area (tropical high montane ecosystem), characterised by soils with a high water storage capacity and high conductivity with a hydrologic be-

haviour for which still major gaps in knowledge exist (Sevink, 2011; Reyes, 2014; Buytaert et al., 2005b, 2006a) and where the hydrometeorological data are scarce. Most modelling efforts in páramo areas have been carried out in micro-watersheds (Buytaert et al., 2004, 2006b, 2005b; Buytaert and Beven, 2011; Buytaert et al., 2007; Crespo et al., 2011) and have focused on advancing the understanding of hydrological processes and anthropogenic impacts. However, there is a relevant need to model larger páramo watersheds (Crespo et al., 2012), and advance in the challenge to produce forecasts for flood early warning to
5 downstream communities. Previous modelling efforts include the use of the AvSWAT model (Díaz-Granados et al., 2005), the use of the linear reservoir model to study land-use changes (Buytaert et al., 2004), a combination of linear reservoirs and TOPMODEL to assess the hydrological functioning of the páramo ecosystem (Buytaert and Beven, 2011) and the analysis of climate change impacts through the use of the WEAP model (Vergara et al., 2011).

2 Study Area

10 Páramos constitute the source of water for Bogotá, the capital city of Colombia. Water is supplied by three main páramo systems namely Chingaza, Sumapaz and Tibitoc (Empresa de Acueducto y Alcantarillado de Bogotá 2015). The Tunjuelo river basin (see Figure 1) with an area of approximately 380 km^2, is located in the south of the city of Bogotá. The upper part of the basin is a páramo area where two reservoirs (Chisaca and Regadera) with volumes of 3.3 Mm^3 and 6.7 Mm^3 operate to supply 1.2 m^3/s of water to the south of Bogotá. This area belongs to the Sumapaz páramo, which is the largest páramo of the
15 world (Torres, 2014). It faces threats such as burning, inappropriate cropping, extensive cattle raising, mining, afforestation with inappropriate species, among others (Torres, 2014). The middle basin corresponds to the transition from the rural area to the urban area of Bogotá (see Figure 1).

In 2006, a dry dam (Cantarrana Dam) was constructed in the middle basin for flood control purposes given the history of flooding of the Tunjuelo river (see Figure 1). The last significant flood occurred in 2002 causing the river to change its course,
20 flowing into two mining pits that currently act as inline reservoirs. In the urban area three retention basins are located upstream of the confluence of the Tunjuelo river with the Bogotá river.

The watershed has a unimodal precipitation regime in the upper part (rainy season April-November) that transforms into a bimodal regime in the lower basin, with rainy seasons in March-May and September-November. The average annual precipitation varies with the influence of the topography; from 600 mm in the North-West to 1500 mm in the upper basin (South-West)
25 (Bernal et al., 2007).

The geology of the watershed consists of sedimentary rocks of Cretaceous, tertiary and quaternary age (INGETEC, 2002). These sedimentary rocks form mountains up to 4000 m altitude, thus reaching some 1500 m above the level of the high altitude plain of Bogotá (Torres et al., 2005). The main soils correspond to inceptisols, andisoils and entisoils (characteristic of páramo areas).

30 The hydrological monitoring network installed in the basin is shown in Figure 1. Although tipping bucket telemetric rain gauges have been operating in the Tunjuelo river basin since the year 2000, the development of the network has been gradual, and only in 2008 the network extended to cover the upper watershed. Six discharge gauges were selected in this analysis, three

of which are located in the upper watershed. Rain gauges provide data each 10 minutes, while discharge gauges report each hour. Even though, there have been significant efforts in recent years to improve the monitoring of the basin, the monitoring network is still considered sparse.

Table 1 summarizes the hydrologic characteristics of the upper watershed of the Tunjuelo river. According to Sevink (2011) the hydrological processes in such páramo areas are fairly simple and dominated by two main pathways: (1) interflow through the upper litter layer, and (2) percolation through the soil layer (which is generally less than 1 m thick) down to the bedrock and subsurface flow parallel to the slope in a saturated zone just above the bedrock. As a result, simple models such as a set of two linear reservoirs already give satisfactory results. The characteristics of these reservoirs are determined by the flow velocities through the respective pathways. Buytaert et al. (2004) and Buytaert et al. (2005a) successfully used the linear reservoir model and the TOPMODEL to study the influence of different land use on the hydrological characteristics of páramo watersheds. Buytaert and Beven (2011) analysed the structure of 9 models to represent the páramo hydrology, finding that the addition of a slow parallel store to the original TOPMODEL (Beven and Kirkby, 1979) appears the most realistic representation of the system to date. However, Buytaert and Beven (2011) highlight that a correct estimation of peak flow remains a challenge.

3 Methods

The methodology is composed of three parts: model setup and calibration; performance analysis and diagnostics; and analysis of precipitation input uncertainty and comparison of models. Three model codes were selected; TETIS (Frances, 2012), HECHMSSMA (HEC HMS Soil Moisture Accounting) (USACE, 2000) and the TOPMODEL (Beven and Kirkby, 1979). These were chosen based on previous use and identified to be suitable in mountainous and páramo areas for the case of TETIS and TOPMODEL (Sevink, 2011), and on the convenience of the HECHMS software since it is widely used in Colombia. However, criteria such as the simplicity and low computational demand were also taken into account. Initial parameters were derived from existing soil data and topography and calibration was carried out using the Shuffled Complex Evolution automatic search algorithm (Duan et al., 1992). On the basis of the calibration results, a performance analysis and diagnosis of each model was carried out by using selected standard performance indices, as well as analysing how well the hydrological signatures due to different processes were represented. Finally, with the aim of analysing the impact of precipitation input uncertainty on the comparative performance of the models, these were driven by Gaussian simulated rainfall fields, and the resulting discharge ensembles were analysed using rank histograms. These were obtained through the ranking of the peak discharge of each ensemble member according to bins created with reference ensembles obtained from the other two models.

3.1 Modelling set up and calibration

3.1.1 Description of the models

Due to the availability of data in the area, the three models were run for the period 01Jul2008-31Dec2012. Data from 01Jun2008 to 01Jul2009 were used for model spin-up. In order to choose a time step for the models, the HECHMSSMA model was tested

with time steps of 1 hour and 10 minutes, finding no significant differences in performance. A time step of 1 hour was used for

65   all subsequent simulations.

A digital elevation model (DEM) of the catchments was generated from contour lines with intervals of 1, 5, 10 and 25 m (depending on the slope). The contours were processed to obtain a triangulated irregular network that was then transformed into a raster through linear interpolation. The DEM was subsequently used to delineate the sub-basins, extract morphometric parameters, and to calculate the topographic index and the channel length distribution as required by the different models.

70   Figure 2 shows the conceptual diagram of the models. A short description of each model is presented in the next paragraphs. For further details, the reader is referred to the literature cited.

TETIS is a conceptual distributed model. The estimation of runoff is based on a hydrological balance in each cell, assuming that the water is distributed into six interconnected storage tanks as shown in Figure 2-a. In the hills, surface flow is a combination of laminar flow and the flow occurring in a network of rills. The hydrologic processes that occur in the interrill areas

75   and in the rills are treated jointly, in such a way that a geomorphological characterization of these elements is not needed. In parallel, interflow and base flow are generated in the corresponding soil layers. Once interflow reaches a cell with a drainage area superior to a defined threshold area for interflow, it reaches the surface, adding to the surface runoff that flows in the surface drainage network. The same occurs when the base flow reaches a cell whose drainage area is superior to the threshold for base flow. From that point on the three flows concentrate in the channel. Surface flow is then routed through the drainage

80   network using the kinematic wave method coupled to the basin geomorphologic characteristics. The model requires the spatial estimation and calibration of the following parameters: the static storage, evapotranspiration (for this study the factor to calibrate evapotranspiration was not used), direct runoff velocity, kinematic wave velocity, infiltration rate, percolation rate, interflow velocity, base flow velocity and deep percolation rate (Frances, 2012).

TOPMODEL (Beven and Kirkby, 1979) is a semi-distributed conceptual model. Total runoff is calculated as the sum of

85   two components (see Figure 2-b); saturation excess overland flow from variable contributing areas, and subsurface flow from the saturated zone of the soil (Uhlenbrook et al., 1999). TOPMODEL uses four basic assumptions to relate down slope flow from a point to discharge at the catchment outlet: the dynamics of the saturated zone are approximated by successive steady state representations; the recharge rate to the water table is spatially homogeneous; the effective hydraulic gradient of the saturated zone is approximated by the local topographic surface gradient S ($\tan\beta$ is the notation most common in TOPMODEL

90   descriptions, where β is the local slope angle); and the effective down slope transmissivity T of a soil profile at a point is a function of the soil moisture deficit at that point (Beven, 2012). Flow is routed through a delay function, which represents the time spent in the channel system. The model requires the estimation of the following parameters: Initial subsurface flow per unit area, transmissivity, rate of decline of transmissivity in the soil profile, Initial root zone storage deficit, maximum root zone storage deficit, unsaturated zone time delay per unit storage deficit, and channel flow velocity inside catchment (Buytaert,

95   2015).

Conceptually, the HMSHMSSMA model divides the potential path of rainfall in a watershed into five tanks as shown in Figure 2-c (USACE, 2000). The model simulates the movement of water through the five tanks, which represent the storage of water on vegetation, on the soil surface, in the soil profile and in the groundwater layers. Given precipitation and potential

Hydrology and
Earth System
Sciences
Discussions

evapotranspiration (ET) the model computes basin surface runoff, groundwater flow, losses due to ET and deep percolation

100   over the entire basin (USACE, 2000). Twelve parameters are needed to model the hydrologic processes of interception, surface depression storage, infiltration, soil storage, percolation, and groundwater storage. The maximum depth of each storage zone, the percentage that each storage zone is filled at the beginning of a simulation, and the transfer rates, such as the maximum infiltration rate are required to simulate the movement of water through the storage zones (Fleming and Neary, 2004). HECHMS provides several options for routing, among them the kinematic wave, which was chosen for this study.

105   **3.1.2   Hydrometeorological forcing**

Ordinary Kriging (OK) and Kriging with external drift were previously tested for rainfall field generation at daily scale in the study area (Rogelis and Werner, 2012). OK was applied to a climatological variogram derived from pooled precipitation data. For Regression Kriging an average residual variogram was derived using the residuals from an average regression surface derived from a set of standardized storm data. The results of this analysis showed that the differences in performance between individual variogram interpolation with OK and with Kriging with external drift and pooled variogram interpolation are not

5   significant. Therefore, both methods were used to obtain hourly rainfall fields for 4.5 years (July 2008 – December 2012) using the daily climatological variograms and daily average residual variograms obtained by Rogelis and Werner (2012).

The preliminary analysis of the hourly rainfall fields showed that Kriging with external drift resulted in unrealistic intensities for the study area in most storm periods (>100 mm/hr), therefore this interpolation method was not considered further. In the case of OK, runoff coefficients in the headwater catchments of the study area, showed unrealistically high values larger than

10   1, indicating an underestimation of the precipitation volume. In OK, when all sampling points are beyond the range of the variogram, the precipitation estimate corresponds to the mean value. Given the short ranges that characterize the convective nature of the precipitation of the study area and the sparse distribution of sub-daily rainfall gauges, most values obtained through kriging equal the mean of the recorded precipitation, leading to a significant underestimation of precipitation. However, OK rainfall fields were used as input to the models in order to identify the impact of precipitation underestimation in the models.

15   A second time series of rainfall fields was obtained through inverse distance weighting interpolation (IDW). The runoff coefficients obtained from these rainfall fields were in the range 0.51 to 0.56, which correspond to more realistic results.

A third time series was created in order to force the models with rainfall ensembles representing the uncertainty in precipitation inputs. This corresponds to an OK rainfall field time series, bias corrected using the IDW rainfall fields as reference time series. The bias correction was carried out through Distribution-Based Scaling - DBS (Yang et al., 2010). This method was ap-

20   plied to the mean precipitation over each sub-basin to generate the bias corrected input for HECHMSSMA and TOPMODEL. In the case of the TETIS input, the bias correction was carried out pixel by pixel to obtain a bias corrected rainfall field. The bias correction procedure modifies the mean value of the rainfall field while preserving the error variance.

Conditional Gaussian simulations were obtained with the same prediction model as used for OK to create an ensemble of 50 rainfall fields, under the assumption that the variance for the original and bias corrected OK rainfall is the same. Ensembles

25   were generated for 78 storms chosen in the period July 2009 – December 2012, which were the most significant in the basin in

this period. The rainfall field ensembles were used to force the models starting from initial conditions previously estimated in a continuous simulation using the bias corrected OK rainfall fields as an input.

Hourly potential evapotranspiration fields were calculated using the Pennman FAO equation (Allen et al., 2006). A crop factor of 0.42, as found by Buytaert et al. (2006a) in the paramos in Ecuador, was used for the areas with paramo vegetation.

30 This was considered constant during the year, and water stress was considered non-existent (Buytaert et al., 2006b). Daily evapotranspiration was first calculated and then a temporal distribution pattern was applied. The temporal distribution of reference evapotranspiration across the day was calculated using data of temperature, humidity, wind velocity and global solar radiation from seven hydrometeorological stations that collect data at 10 minute intervals.

3.1.3 Model Configuration and Calibration

35 Parameters were calibrated separately on a sub-basin level from upstream to down-stream in the three models. The parameters were calibrated against observed discharge measurements at the internal stations. Figure 3-a shows the sub-basins and the calibration points where discharge measurements are available. In order to not propagate upstream errors in the calibration process, observed discharges at upstream sub-basin outlets are used as inflow when calibrating downstream sub-basins.

The initial parameters for the three models were obtained from existing soil, land cover and topographical data of the basin.

40 These are shown in Figure 3-b and Figure 3-c. Calibration was performed by optimization of the Kling and Gupta efficiency (KGE) (Gupta et al., 2009) with the Shuffled Complex Evolution (SCE) automatic search algorithm (Duan et al., 1992). In the case of TETIS, the SCE algorithm is implemented in the software. For HECHMSSMA and TOPMODEL the SCEoptim function of the hydromad R package was used.

In the HECHMSSMA model, the Tunjuelo river basin was divided into sub-basins linked with channel reaches as shown in

45 Figure 3-a. The ARCGIS HEC-GEOHMS extension (Fleming and Doan, 2013) was used for basin delineation. The initial set of sub-basins was modified to take into account the hydrological stations and the flood control structures of the river, leading to a total of 13 sub-basins with areas ranging from 4 to 92 km^2. For the watersheds in the upper basin, the hydrological stations are located immediately upstream of the reservoirs, allowing the calibration of the entire watersheds contributing to the reservoirs.

50 Channel reach length and slope were determined using HEC-GEOHMS. The resolution of the DEM and absence of bathymetry prevented accurate extraction of channel cross-section information. A trapezoidal section was assumed in the middle and upper basin, with a constant Manning roughness coefficient of 0.04. In the lower basin, an average section was used according to the available bathymetry and a Manning coefficient of 0.035 was extracted from a calibrated hydrodynamic model available for the lower part of the basin.

55 All five tanks available in the HECHMSSMA model were used, with the Clark unit hydrograph applied as transformation method. The linear reservoir model was used for base flow estimation. With this configuration, the model has 16 parameters that require calibration in each sub-basin, as well as the initial condition of each of the five tanks. The assumption of negligible deep percolation, given the low permeability bedrock in the whole basin, reduces to 15 the number of parameters, while a warm up period eliminates the effect of initial conditions. The model parameters were first estimated based on the land cover, geology

60 and soil information and then a three-stage calibration was carried out. First a manual calibration of the three less sensitive parameters was carried out, subsequently, SCE was used to calibrate the 12 most sensitive parameters (see Table 3) and finally a manual recalibration was used to refine the 3 less sensitive parameters. The sensitivity analysis of the model showed that the canopy storage and surface storage are less sensitive than the other parameters, as well as the time of concentration.

For areas of the basin formed by two sub-basins with only one discharge station at the outlet, distributed precipitation forcing
65 is averaged over each sub-basin and identical model parameters are used for constituent sub-basins, thus optimizing a single parameter set.

In the case of TETIS and TOPMODEL, the effect of model resolution was explored. Parameters for the TETIS model were estimated using pixel sizes of 100, 250, 500 and 1000 meters. Smaller pixel sizes where not used due to excessive run model times. Pixel sizes of 25, 50, 100, 250, 500 and 1000 metres were used for the TOPMODEL.

70 Initial spatial distributed parameters for the TETIS model were estimated according to the land cover information, soils and geology (Puricelli, 2008). Grids with the chosen resolution were created for elevation, static storage, hydraulic conductivity of the soil, percolation, horizontal saturated conductivity, horizontal saturated conductivity of the substrate, deep percolation, surface flow velocity in the hills, slope, flow direction and flow accumulation. In order to create the grids, the R project software in combination with SAGA GIS and ArcGIS was used to process the following input data: DEM of the basin; soil
75 characteristics (sand, clay and gravel content, organic matter content, profile) according to the soil type as shown in Figure 3-b (IGAC, 2000); geology (INGETEC, 2002); and the land cover obtained from the classification of a LANDSAT Thematic Mapper 5 (TM5) image taken in 2001.

The behaviour of the water in the tanks of the model is described by equations that incorporate multiplicative correction factors for calibration purposes. The description of these correction factors is shown in Table 2.

80 TETIS uses the kinematic wave with hydraulic characteristics of the channels obtained from the geomorphological information of the watershed; this is the kinematic geomorphologic wave. An additional correction factor, FC_9 is used to correct the wave velocity. All correction factors were calibrated except for FC_2 to preserve the same input in all models. The maximum storage capacity of the gravitational tank (H3_max, see Figure 2-a) determines the return flow that produces saturation excess. H3_max cannot be calibrated automatically by the SCE algorithm that is hardwired in TETIS, and a manual procedure was
85 therefore carried out to estimate this parameter. Calibrations were carried out using maximum capacities of the gravitational storage of 10, 30, 50, 80, 100, 150 and 200% of Hu (maximum static storage capacity, see Figure 2-a) and a large value to completely avoid saturation excess. The sizes of the tanks are used by the TETIS model in millimetres. These variations of H3_max were tested in two model configurations: a) considering very low percolation (rock strata under the gravitational storage) therefore the aquifer tank is not used; and b) a percolation similar to the saturated conductivity, thus the aquifer tank is
90 used in the simulations. The tests showed that for the two configurations only one of the two subsurface storages dominated the response of the watersheds. Furthermore, variations in the maximum storage do not affect the KGE coefficient and have a marginal impact on the FDC. The modifications tested in the model do not affect the overland flow, with this being minimal in all cases. The best performance of the TETIS model, from the KGE and the FDC signatures, was obtained for a model with a

large capacity of the gravitational storage so no saturation excess is produced and considering a low permeability (rock strata

95 under the gravitational storage). This was used for the subsequent phases of the analysis.

The package TOPMODEL for R (Buytaert, 2015) was used to set up the models for the three headwater sub-basins of the páramo area. The TOPMODEL application was limited to only these three watersheds, since the response of the watershed downstream is mainly dominated by the routing of the reservoir releases (see Figure 1), with the páramo area being the main priority for flood forecasting. A DEM with the required resolution for each sub-basin was used as input and the functions of

100 the R package were used to obtain the topographic index distribution and the delay function. Table 4 shows the parameters that were calibrated.

3.2   Performance analysis and diagnostics

Model diagnosis is a process by which inferences are made about the representation of hydrological processes through targeted evaluation of the input-state-output response of the model (Yilmaz et al., 2008). In order to carry out a diagnosis of the models two approaches were followed: a) an analysis of the fluxes produced by each model (e.g. percolation, base flow, interflow etc) and b) the analysis of the flow duration curve (FDC) obtained from the simulated discharges at the calibration points.

Following Yilmaz et al. (2008), the flow duration curve (FDC) was used as a tool to summarize a catchment's ability to produce flow values of different magnitudes, and is therefore strongly sensitive to the vertical redistribution of soil moisture

5 within a basin, while being relatively insensitive to the timing of hydrologic events. Five signature measures based on the FDC were used as shown in Table 5. The approach partitions the FDC into three segments: (1) the high flow segment, which characterizes watershed response to large precipitation events; (2) the mid-flow segment, which characterizes watershed response to moderate size precipitation events as well as the medium-term baseflow relaxation response of the watershed; and, (3) the low flow segment, which characterizes the long-term sustainability of flow (Yilmaz et al., 2008; Pfannerstill et al., 2014).

10 ## 3.3   Analysis of precipitation input uncertainty and comparison of models

Bias corrected Gaussian simulations were used to produce a 50-member discharge ensemble for each model for the 78 chosen storms. For the models where pixel size was tested, only the best performing model resolution was used. The spread of the discharge ensembles was used as a metric of the sensitivity of the models to the variability of the precipitation. The interquartile range (IQR), the median absolute deviation averaged (MAD), and the range for all the chosen storms were calculated according

15 to Equation 1, Equation 2 and Equation 3 (Franz and Hogue, 2011).

$$\overline{IQR} = \frac{1}{n} \sum_{t=1}^{n} \left(q_{0.75}(t) - q_{0.25}(t) \right) \tag{1}$$

$$\overline{MAD} = \frac{1}{n} \sum_{t=1}^{n} median_i \left| x_i(t) - x_{med(t)} \right| \tag{2}$$

$$\overline{Range} = \frac{1}{n} \sum_{t=1}^{n} \left(x_{(1)}(t) - x_z(t) \right) \tag{3}$$

where $q_{0.75(t)}$ and $q_{0.25(t)}$ are the 75th and 25th percentiles of the ensemble, respectively; $x_i(t)$ represents the value of a variable in each ensemble member for timestep t; $x_{med}(t)$ is the ensemble median; $x_{(1)}(t)$ and $x_{(z)}(t)$ are the lowest and highest valued ensemble members, respectively; and n is the number of timesteps.

Furthermore, rank histograms were constructed to compare the discharge ensembles between models. For each ensemble member, the peak flow was ranked using as reference the ensemble of peaks of the other two models. The peak flows of the comparison ensembles are assigned to the intervals created with the ordered peaks of the reference ensembles. Thus, the shape of the resulting histogram provides information about the ensemble in comparison with the reference ensemble. If the histogram is uniform the two ensembles are similar, if the histogram is skewed to the right the comparison ensemble tends to higher values than the reference ensemble and the opposite if it is skewed to the left.

Once the frequency of the peak discharges of the comparison ensemble has been determined according to bins created with the reference ensembles, all the rank histograms are pooled obtaining the frequency of the ensemble peaks of each model according to the ordered ensemble peaks of the other two models.

4 Results

4.1 Model calibration

4.1.1 KGE for HECHMSSMA, TOPMODEL and TETIS

The first two columns for each sub-basin in Table 6 show the optimum KGE values obtained from calibration using as forcing for the models the OK and IDW rainfall fields. The third column (OKbc) shows the KGE obtained from the simulations with models configured with the parameters obtained from calibration with IDW rainfall fields but using the bias corrected OK rainfall fields as input precipitation. In the case of the TOPMODEL, the KGE for OKbc was obtained for grid sizes of 500 m and smaller, due to the drop in performance for larger grid sizes.

There is an increase in performance when using IDW rainfall fields in comparison with OK rainfall fields for the Mugroso and Curubital sub-basins. For the Chisacá sub-basin the increase in performance occurs for the HECHMSSMA model and for the TETIS model with pixel sizes smaller than 500 m. In the case of the sub-basins located in the middle and lower basin the differences are less significant. The use of OKbc as forcing for the models produces minor reductions of efficiencies when compared with the best efficiency obtained for IDW precipitation and a pixel size of 500 m in the case of HECHMSSMA and TETIS; and very similar values in the case of the TOPMODEL.

The calibration results can be grouped into poor performance ($0.5 > KGE > 0$), intermediate ($0.75 > KGE > 0.5$) and good performance ($KGE > 0.75$) (Thiemig, 2014). According to this classification the headwater catchments located in the paramo area (Chisaca, Mugroso and Curubital sub-basins), exhibit maximum efficiency values in the range of intermediate performance

(see Table 6). The maximum efficiency values are similar for all the three models, with HECHMSSMA and TETIS reaching the
highest values compared to TOPMODEL. Regarding the watersheds downstream of the páramo area, the results are dominated
50 by the discharge from the reservoirs, and therefore depend mainly on the routing of the measured hydrograph. KGE values are
in the range of intermediate to good performance.

4.2 Comparison of water balance fluxes

The total volumes of the fluxes in millimetres from each model and for IDW and OK rainfall fields are shown in Figure 4. The
results obtained by driving the models with OKbc rainfall fields are not shown since they are similar to the results obtained
55 from IDW rainfall fields. Only the results for the headwater watersheds in the páramo area are shown, since the release of the
reservoirs dominates the outflow discharge of the watersheds downstream.

Results show that the input rainfall obtained through OK is lower in comparison with the precipitation obtained from IDW,
leading to models where there is no actual evapotranspiration, which highlights the underestimation of precipitation values by
the OK interpolation. This underestimation significantly affects the performance of the models, not only in terms of efficiency
60 as shown in Table 6 but also affects the models' ability to properly simulate hydrological processes.

The results of the HECHMSSMA model show a dominance of the groundwater flow from the groundwater layer 2 for
Chisaca and Mugroso and overland flow for Curubital for OK rainfall fields and dominance of the groundwater flow from the
groundwater layer 2 for Mugroso and Curubital and overland flow for the case of Chisaca when IDW rainfall fields are used.

In the case of TOPMODEL, the response of all the models is dominated by subsurface flow with a small contribution
65 of overland flow. When OK rainfall fields are used the response in the three watersheds does not vary significantly and the
evapotranspiration is negligible. When the precipitation is calculated with IDW, the evapotranspiration is significant in the
water balance with approximately equal proportions in the Mugroso and Curubital sub-basins and approximately twice in the
Chisaca sub-basin.

The TETIS model has a similar behaviour for both the IDW rainfall fields and the OK rainfall fields with the latter being
70 lower in volumes. The dominant process in the headwater páramo catchments is interflow, this is the release from tank 4
(gravitational storage) in Figure 2-a. In the case of the IDW rainfall fields overland flow is negligible. In contrast, this flux is
observable when OK fields are used, albeit in a small volume. The increase in pixel size influences the proportion of subsurface
flow and evaporation. The most severe changes are observed for a pixel size of 1000 m. These are due to a significant change
in the drainage area and stream network caused by the coarse grid. In the case of the Chisaca sub-basin, for pixel sizes higher
75 than 500 m, the drainage area duplicates increasing the precipitation input, which causes a significant increase in the storage
when IDW rainfall fields are used and an increase in evapotranspiration in the case of OK rainfall fields to compensate for the
increase in drainage area, due to changes in connectivity. This resolution is, thus, considered to be too coarse for these small
watersheds.

Considering only pixel sizes up to 500 m, when input precipitation obtained from IDW is used, in the case of Mugroso and
80 Curubital watersheds, actual evapotranspiration reduces with pixel size, while in the case of Chisaca the evapotranspiration
accumulation remains constant.

4.3 Signature measures from the flow duration curve (FDC)

Due to the significant underestimation of precipitation with the OK interpolation, the results in this section will refer only to the precipitation obtained through IDW interpolation.

85 Figure 5 shows the flow duration curves from the HECHMSSMA, the TOPMODEL and TETIS the models. Only the results for the pixel size of 25 m are shown for TOPMODEL since the curves for the other sizes are similar. For the case of TETIS only the models up to a pixel size of 500 m were considered due to the significant deterioration of the representation of the drainage network occurring when larger pixel sizes are used.

In the case of TOPMODEL the lowest overall biases for all the watersheds are found for a pixel size of 25 meters, as well as 90 the highest KGE. This pixel size was therefore used for subsequent analysis.

The TETIS model better represents the high flow portion of the duration curve (discharges equalled or exceeded less than 20 % of the time) exhibiting the lowest bias values (FHV) in the case of the Mugroso and Curubital watersheds (see Table 7). For the Chisaca watershed, the TOPMODEL has a significantly better performance than the other models in this part of the FDC exhibiting the lowest bias values.

95 The middle portion of the FDC (flows equalled or exceeded between 20 and 70 % of the time, see vertical lines in Figure 5) is better represented (lowest FMS, see Table 7) by the TOPMODEL in the Chisaca and Curubital and by the TETIS model in the case of the Mugroso watershed.

The TETIS model exhibits the highest biases (FLV) for the lowest flows (flows equalled or exceeded more than 70 % of the time), while the lowest biases correspond to the TOPMODEL.

100 In terms of grid size, the lowest overall biases in the TETIS model are obtained for a grid size of 100 m for Mugroso and Curubital and for a grid size of 500 m for the Chisaca sub-basins (see Table 7). However, the lowest biases in the high flow segment of the FDC for Mugroso and Curubital correspond to the grid size of 500 m. A grid size of 500 m was chosen for the subsequent analysis due to its good representation of high flows, and insignificant reduction of performance in the middle and low segment of the FDC. Models until this grid size also had a shorter computation time and higher KGE values.

105 ## 4.4 Rainfall ensemble analysis, input precipitation uncertainty

The results in Table 6 , show that the bias corrected OK rainfall fields provide a very similar response of the models to that found with the IDW rainfall fields. Given the good performance of the bias correction, the Gaussian simulations were produced by applying the climatological variogram used in the OK interpolation, and then bias corrected using the corresponding mapping functions, and used as input to the models to test their sensitivity to variability in precipitation.

Table 8 shows the IQR, MAD and range for the ensemble discharge of the 78 storms selected in the period of analysis. In all three watersheds, the metrics calculated for all storms have similar values in each watershed, with the highest values consistently corresponding to the TOPMODEL, except for the MAD and IQR for the Chisaca sub-basin, where the highest 5 values correspond to the HECHMSSMA model. The smallest ensemble spreads are found in the Mugroso sub-basin, while the

highest are found in the Chisaca sub-basin. The average values for the peak discharge of each storm shows that the TOPMODEL is clearly the most sensitive model to variations in rainfall input, exhibiting the largest IQR, MAD and Range at the peak.

4.5   Comparison of model ensembles

Figure 6 shows the rank histograms for the head watersheds in the páramo area comparing the discharge ensembles of the models. The comparison of the ensembles obtained from HECHMSSMA and TOPMODEL (first column in Figure 6) shows that the members of the TOPMODEL ensemble have mostly higher values than the HECHMSSMA ensemble. The comparison of TETIS and TOPMODEL shows equally that the members of the TOPMODEL have mostly higher values than the TETIS ensemble. In the case of TETIS and HECHMSSMA, the rank histogram shows less difference between the two ensembles with an approximately uniform distribution for the Mugroso and Curubital watersheds. For the Chisaca watershed, the rank histogram shows underdispersion meaning that most values of the HECHMSSMA model are larger or smaller than the TETIS ensemble.

5   Discussion

5.1   Model calibration and performance

5.1.1   Water balance fluxes and hydrometeorological forcing

The precipitation and evapotranspiration data are considered as the main source of uncertainty in the models (Buytaer et al. 2005). Precipitation data in the páramo area are subject to errors inherent to the significant difficulties in the measurement process and high spatial rainfall variability (Buytaert et al., 2006b). Wind speeds at high altitude may be high and a smaller or larger portion of the rain may be blown over the rain gauge (Sevink, 2011). Furthermore, fog is highly difficult to quantify (Bruijnzeel, 2001; Tobón and Gil - Morales, 2007), and this may add an unknown quantity of water, especially where patches of arbustive species are present (Buytaert et al., 2006b).

Evapotranspiration is influenced by the particularly low evaporation characteristics of the vegetation. Tobón and Gil - Morales (2007) found that during those fog events that do not produce dripping onto the floor, there are no net inputs to the ecosystems, and the contribution of fog to the catchment water yield can be only through their control over forest transpiration.

Even though the approximation of evapotranspiration by using the Penman FAO equation is considered appropriate for paramo areas by some authors (Sevink, 2011), difficulties in the reliable estimation of humidity under foggy conditions (Sevink, 2011) may introduce significant errors. In addition, fog is not only thought to induce an extra input of water into the ecosystem but also to suppress evaporation (Sevink, 2011). Buytaert et al. (2006a) highlights the limited validity of the Penman FAO equation under the unusual meteorological conditions of the páramo.

Besides the impossibility to include fog interception given the lack of data, the estimation of the rainfall field has shown to be highly challenging. Different interpolation methods lead to significantly different precipitation volumes, strongly influencing the efficiency and performance of the models. Ordinary kriging using a daily climatological variogram produces lower KGE

values than IDW; this is mainly due to the underestimation of precipitation volumes in the case of the former. IDW seems to produce more realistic precipitation values.

The comparison of precipitation volumes with the observed discharge accumulated volumes is shown in Figure 7. The
40 precipitation volumes obtained from OK are less than the observed runoff in the three paramo watersheds indicating an underestimation of the precipitation. The accumulated rainfall is about 1000 mm lower than the observed runoff in the Mugroso and Curubital watersheds, while in the Chisaca watershed this difference reduces to 200 mm. The impact of this difference in the performance of the models is reflected in a reduction in the actual evapotranspiration in the three models when forced with the OK rainfall field. TOPMODEL and TETIS reduce the actual evapotranspiration to almost 0 through the reduction
45 of the model tank that represents the root zone storage in the former, and the interception and static storage in the latter. In the case of HECHMSSMA, the model does not completely reduce the evapotranspiration, which then leads to a significant underestimation of discharge compared with the other two models.

The accumulated IDW precipitation is approximately equal for the three watersheds. Given the similarities in terms of soils, land cover and geology actual evapotranspiration is expected to be approximately the same in the three watersheds. However,
50 the accumulated runoff for the Chisaca watershed is lower (approximately 1000 mm lower than in the other watersheds) which leads to a resulting actual evapotranspiration that is higher than in the other watersheds. This suggests a relative overestimation of precipitation (real precipitation lower than the precipitation in Mugroso and Curubital watersheds) for the Chisaca watershed that produces an increase in the actual evapotranspiration to balance outputs in the models. This behaviour of the Chisaca watershed suggests that the available precipitation data is not representative of the precipitation occurring in this watershed.

55 According to Buytaert et al. (2006a) literature values of calculated actual evapotranspiration for grass páramo range from 0.8 to about 1.5mm/day. The only two models in or close to that observed range are the TOPMODEL and TETIS forced with IDW rainfall fields with values of 0.82 and 0.89 mm/day respectively for Chisaca, 0.78 and 0.73 for Mugroso and 0.5 and 0.86 for Curubital. In the case of the other models the actual evapotranspiration is highly underestimated in comparison with observed values reported in literature. These results show that realistic ranges of actual evapotranspiration are only obtained
60 in the Chisaca watershed and in the Curubital watershed with TETIS, suggesting that the precipitation volume estimated with IDW is low mainly for Curubital and Mugroso.

5.1.2 Pixel size and flux variation for the TOPMODEL and TETIS

In the following sections the discussion will focus only on the results obtained from the calibration using IDW rainfall fields, due to the underestimation obtained when OK rainfall fields are used.
65 TOPMODEL and TETIS models with pixel sizes larger than 500 m produce similar, and in some cases better KGE than in the case of finer grids. However, the drainage network of the watersheds cannot be correctly represented with these pixel sizes in the watersheds in the paramo areas. The most notorious case corresponds to the Chisaca and Mugroso watersheds where the distance between the two main streams is less than 1000 m in some reaches, which leads to accumulation grids that cannot correctly represent the stream network. The similar KGE are due to adjustment of the model parameters without

70 correctly representing the hydrological behaviour of the watersheds. The results of these coarse models will also not be taken into account further.

In TOPMODEL the KGE values vary in a maximum range of only +/- 0.03, reflecting very similar efficiencies regardless of the pixel size. Increasing the grid size of the DEM increases mean values of the topographic index (Deginet, 2008). The mean topographic indices for the three watersheds increase when the grid size increases going from values close to 6.5 to

75 10.4 for the coarser resolution (1000 m), increasing more significantly for pixel sizes larger than 100 m. This is due to the greater upslope contributing area and smaller slope (Wu et al., 2007). This behaviour in the topographic index is consistent with previous studies (Deginet, 2008; Bruneau et al., 1995). Wu et al. (2007) found that the smoothing effect of grid size increase may result in deteriorated topographic index distributions at coarse resolutions. However, this can be moderated by parameter calibration, as found in the results shown in Table 6. Despite the change in the topographic index distributions,

80 fairly similar efficiencies can be preserved by the compensation effect of the calibration parameters, mainly transmissivity. The increase in grid size produces an increase in saturated areas that results in the increase of overland flow when the same calibration parameters are kept (Deginet, 2008). This behaviour is explained by the disappearance of the smaller values and increase of the mean values of the topographic index (Deginet, 2008). Hence, the adjustment of the transmissivity to higher values allows to obtain almost identical model efficiencies (Franchini et al., 1996; Saulnier et al., 1997; Wu et al., 2007). The

85 increase in transmissivity is larger for pixel sizes larger than 100 m, in correspondence with the increase in the topographic index. This increase in transmissivity keeps the overland flow proportion fairly similar for the three watersheds. The calibrated transmissivity values for pixes sizes up to 500 m range between 0.32 and 16.5 m2/h. The lowest values are consistent with the transmissitivy values found by Buytaert et al. (2005a) for páramos in Ecuador and the highest values are still in the range of transmissivity values found in other applications of the TOPMODEL (Beven, 1997).

90 In the TETIS model the variation of pixel size produces only minor changes of +/-0.03 in the KGE. However, a pixel size of 500 m is an optimum in terms of KGE in the case of Mugroso and Curubital. In all three watersheds, the lateral conductivity of the soil increases with pixel size. This is the main parameter used by the model to compensate for variations in pixel size. The discharge coefficient (α), that multiplies the storage in tank 3 (H3) to obtain its outflow (interflow) (see Figure 2-a), is a function of the horizontal saturated conductivity, the pixel size and the time interval. This mathematical relation explains its

95 scale dependency. The values of the discharge coefficient in the paramo watersheds range from 0.72 to 0.89 implying high outflows from tank 3.

The comparison of the calibration results of the three watersheds in terms of the behaviour of each tank (see Figure 2) can be summarized as follows:

a) Static storage: this storage corresponds to water that can be evaporated from surface depressions, vegetation and water

100 retained in the soil through capillary forces (Frances, 2012). The correction factors that multiply the capacity of the storage correspond to values higher than 1 for Chisacá, but remain approximately constant for pixel sizes from 100 to 500. In the case of Mugroso, correction factors increase with pixel size but remain low, reaching 0.12 for a pixel size of 500 m. For Curubital correction factors slowly increase with pixel size to reach a value close to 1 for a pixel size of 500 m. Due to the close connection of this storage with the evapotranspiration process, these results may be due to a lack of representativity of

105 the precipitation data obtained from the station located most to the south-west of the study area (see Figure 1), that may lead to relative underestimation of precipitation in the Chisaca watershed, forcing the model to compensate by increasing the capacity of the tank to increase evapotranspiration losses in comparison with the other watersheds (see Figure 4). b) Superficial storage: The calibrated hydraulic conductivities are high in comparison with the rainfall intensities of the páramo area. For a pixel of 500 m calibrated hydraulic conductivities range between 82-135 mm/h, which correspond to high values in comparison with a range of 10-60 mm/h found in other páramo areas (Crespo et al., 2009). These high conductivity values result in no infiltration excess occurring in the model, and the water moving to the gravitational storage, which is consistent with the characteristics of the páramo described in Table 1. c) Gravitational storage: the calibrated percolation is very low, which means almost no

5 water is going to the aquifer storage. This explains why the flow is dominated by the outflow from this tank (see Figure 4). The fluxes of the model are dominated by the discharge from the tank number 3, which can be interpreted as the discharge from shallow soil above impervious strata. No saturation excess flow is produced; therefore the model does not simulate any rapid response/overland flow of the watershed. The behaviour of the gravitational storage is coherent with the hydrological behaviour of páramo watersheds described in Table 1. d) Aquifer storage: due to the very low permeability of the rock underneath the soil

10 layer, the storage in this tank is negligible, as well as the outflow.

5.1.3 HECHMSSMA calibration results and fluxes

The KGE values obtained from the HECHMSSMA are similar to the values obtained from the calibration of the TETIS model. In terms of fluxes of the models (see Figure 4), these are similar for the Mugroso and Curubital watersheds for IDW rainfall fields (dominance of subsurface flow). Conversely, the response of the Chisacá watershed is dominated by overland flow

15 generated through infiltration excess. This is due to a low soil infiltration in the calibrated model, producing a response dominated by overland flow. This representation of processes in the model is not consistent with the other two models, or with the perceptual model of the watershed, implying the inability of the HECHMSSMA model to adequately represent the hydrology of the Chisacá watershed, given the available data. From the previous analysis, a relative overestimation of precipitation was detected in the Chisaca watershed, which suggested a lack of representativity of the measured precipitation in this watershed.

20 This difference in hydrometeorological forcing may be the cause for different hydrological processes calibrated to represent the watershed response.

The response of the Mugroso and Curubital rivers is similar. Saturated conductivities larger than the rainfall intensity prevent the occurrence of overland flow. The soil percolation rates are high, therefore infiltrated water moves rapidly to the first ground water layer. The percolation rate from the first groundwater layer to the second is high, therefore water moves quickly to the

25 second groundwater layer. This rapid percolation to the second groundwater layer inhibits outflow from the first groundwater layer; therefore the subsurface response is dominated by the outflow from the second groundwater layer. This behaviour is consistent with the dominance of subsurface flow, characteristic of páramo areas described in Table 1.

5.1.4 Flow duration curve and signatures

The sensitivity to variations in pixel size is negligible in the case of TOPMODEL where the transmissivity parameter com-
pensates changes in grid size, reaching similar KGE values and producing very similar FDC with almost the same signatures,
with the finer pixel model showing the smallest biases. Conversely, the TETIS models are significantly sensitive to changes
in pixel size mainly, for low discharges (equalled or exceeded more than 70% of the time). Furthermore, the TETIS model
exhibits the poorest performance for low discharges. This is due to the rapid outflow from the storage representing the sub-
surface flow, which fails to represent the slow water release of the soil of the paramo areas. The same behaviour is observed
in the HECHMSSMA model for the Mugroso and Curubital sub-basins. However, in the Chisaca sub-basin that model better
represents low discharges when compared to TETIS, since the response of the model is dominated by infiltration excess, and
the subsurface flow is modelled through the water release from the second underground storage with a large routing coefficient.
In the TETIS models and in the HECHMSSMA models, the subsurface flow is represented by only one storage, despite having
the possibility to use two. In both models, the water flows rapidly to the deeper storage that controls the response.

In general, the TETIS models overestimate discharges for large discharges (equalled or exceeded less than 20% of the time)
and underestimates for lower discharges. The HECHMSSMA model has a similar behaviour. The TOPMODEL overestimates
low discharges in the case of Mugroso and Curubital and slightly underestimates them in the case of the Chisaca watertheshed.
For high discharges, TOPMODEL has a good representation of the FDC in the case of Chisaca and Curubital sub-basin, and a
slight underestimation in the case of Mugroso.

Given the FDC results and the KGE, TOPMODEL appears to be the most realistic model of the three models tested in this
analysis. This is supported by the assumptions of TOPMODEL that seem to be able to adequately represent the main charac-
teristics of the paramo soils response (Buytaert and Beven, 2011), with the hydrologic response dominated by the topography
and no infiltration excess overland flow; and nonlinear transmissivity profile. In agreement with other studies carried out in the
páramo area (Buytaert and Beven, 2011) the assumption of an exponential function of the storage deficit seems to provide a
good representation of the processes in these watersheds.

According to the results, higher performance metrics such as the KGE do not necessarily mean a better representation
of hydrological processes, and therefore, they are not an indication of realism of the model, which is necessary for flood
forecasting reliability (Kavetski and Fenicia, 2011). However, the use of signatures and analysis of model fluxes provides a
means to compare model structures in terms of their abilities and limitations to reproduce the dominant hydrological processes,
and to gain insight into the characteristics of a model that make it more suitable than others. Consistency, defined by Euser
et al. (2013) as the ability of a model to reproduce several hydrological signatures with the same parameter set is a criterion
that provides the means to assess the reliability. Furthermore, the comparison in terms of process representation is crucial
to interpret the effects of using different model structures (McMillan et al., 2011). The correspondence between catchment
structure and model structure was identifiable in this study, which provides understanding about the watersheds behaviour.

5.2 Comparison of discharge ensembles

The analysis of the discharge ensemble spread in the models shows a higher sensitivity of TOPMODEL to variation in the rainfall. Increases in the precipitation cause a significant increase of the peak discharges of the storms since the precipitation over saturated areas immediately contributes to overland flow. For the TETIS and HECHMSSMA models precipitation infiltrates and flows as subsurface flow through the underground tanks, which reduces the increase in peak flow in comparison to TOPMODEL. This means that, TOPMODEL is the most sensitive model to rainfall variability, albeit the most realistic.

The Chisaca sub-basin shows the larger ensemble spread metrics, with this being the sub-basin with the most unreliable precipitation input. Due to the apparent relative overestimation of precipitation in this watershed, the parameters of all models adjust to increase evapotranspiration and reduce the outflow discharge. Therefore when increases in precipitation occur in the rainfall ensembles, the increases in peak flows are larger than in the other models, where the balance between the fluxes seems more realistic.

6 Conclusions

A distributed model (TETIS), a semi-distributed model (TOPMODEL) and a lumped model (HEC HMS soil moisture accounting) were used to simulate the discharges of a tropical high mountain basin characterized by soils with high water storage capacity and high conductivity. The performance analysis and diagnostic applied allowed insight in the representatively and appropriateness of the models. The comparison of models, through performance measures combined with analysis of fluxes and flow duration curve signatures, provided a means to assess the abilities and limitations of the models. This analysis allows insight into the models process representation, providing the information needed to identify a model structure that is more suitable than the others in terms of how realistically relevant hydrological processes are simulated.

Different model structures were shown to have similar performance according to the King and Gupta efficiency (KGE) value, however their ability to reproduce hydrological processes varies. The ability to reproduce hydrological processes is also influenced by inputs errors. Overestimation and underestimation of precipitation can produce a change in the dominant hydrological processes simulated by the models, with some models more sensitive to these errors than others. In the study area, the use of a climatological variogram with ordinary kriging to interpolate hourly rainfall fields proved to result in underestimation of rainfall, significantly affecting the performance of the models. Due to the complex spatio-temporal variability of precipitation, the simpler approach, using Inverse Distance Weighting (IDW) was found to be the most appropriate.

The use of varying pixel sizes in the semi-distributed and distributed model, showed that a first and determinant criteria for upper limits in pixel size is the ability of the grid to appropriately reproduce the drainage characteristics of a basin. Furthermore, variations in the pixel size are compensated by selected parameters in each model, in order to reach approximately the same performance for all grid sizes. In the case of TOPMODEL the compensation is achieved though variations in the transmisitivy, for TETIS the compensation is manly achieved through variations in the lateral conductivity of the soil.

Despite the compensation of parameters, an optimum grid size could be identified in the TETIS and TOPMODEL through the use of the FDC signatures, through which the slight variations in representation of processes could be identified according to pixel size. These optimum grid sizes are 500 meters for TETIS and 25 meters for TOPMODEL.

The behaviour of TETIS and HECHMSSMA models for the páramo is similar in terms of the water flow in the underground tanks. Only one of the two underground tanks available is used due to high conductivity values that produce a rapid flow towards the deeper tank. In the case of TETIS the tests with several configurations of the model showed that a model consisting of a tank representing the soil layer over an impervious rock layer (aquifer storage) performs best. This is consistent with the perceptual model of the hydrology of the watershed. In the case of HECHMSSMA one of the two tanks representing interflow dominates the response. However, saturation excess is not modelled by any of these two models, thus the flow is exclusively dominated by the release of one underground tank. With this configuration, none of these models has the ability to reproduce the slow water release in the low flow portion of the FDC. This is due to a rapid flow of water from the dominating underground tank in response to the high conductivities that are obtained from calibration. For these models, even if a relatively good representation of high discharges can be achieved, low flows cannot be modelled appropriately.

TOPMODEL appears to be the most realistic model for the páramo of the models tested in this analysis, although it is more sensitive to rainfall fields variability. This model is able to reproduce the slow water release from the soil layer over the rock stratum that is one of the main characteristics of the páramo soil. The signatures obtained from the flow duration curves show that this is the model that more closely reproduces all ranges of discharge in the three páramo sub-basins. Besides providing more reliability, TOPMODEL demands low computational resources and short run times. These aspects support that TOPMODEL is the preferred choice from a flood early warning perspective.

Hydrology and
Earth System
Sciences
Open Access
Discussions
EGU

References

20   Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Evapotranspiración del cultivo Guías para la determinación de los requerimientos de
     agua de los cultivos, in: Estudio FAO Riego y Drenaje, p. 323, 2006.

Basher, R.: Global early warning systems for natural hazards: systematic and people-centred., Philosophical transactions. Series A, Mathe-
     matical, physical, and engineering sciences, 364, 2167–82, http://www.ncbi.nlm.nih.gov/pubmed/16844654, 2006.

Bernal, G., Rosero, M., Cadena, M., Montealegre, J., and Sanabria, F.: Estudio de la Caracterización Climática de Bogotá y cuenca alta del

25   Río Tunjuelo, Tech. rep., Instituto de Hidrología, Meteorología y Estudios Ambientales IDEAM - Fondo de Prevención y Atención de
     Emergencias FOPAE, Bogotá, 2007.

Beven, K.: TOPMODEL : A CRITIQUE, Hydrological Processes, 11, 1069–1085, http://hydraulics.lecture.ub.ac.id/files/2012/10/
     TOPMODEL-A-critique.pdf, 1997.

Beven, K.: Hydrological Similarity and Distribution Function Rainfall-Runoff models, in: Rainfall-Runoff Modelling, pp. 397–448, John

30   Wiley & Sons, Ltd, http://dx.doi.org/10.1002/9781119951001.refs, 2012.

Beven, K. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction., http://www3.interscience.wiley.
     com/journal/112736150/abstract, 1992.

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de
     zone d'appel variable de l'hydrologie du bassin versant, Hydrological Sciences Bulletin, 24, 43–69, http://www.tandfonline.com/doi/pdf/

35   10.1080/02626667909491834, 1979.

Bruijnzeel, L.: Hydrology of tropical montane cloud forests: a reassessment, Land use and water resources research, 1, 1–18,
     http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Hydrology+of+tropical+montane+cloud+forests+:+A+Reassessment#
     1\$\delimiter"026E30F\$nhttp://www.bvsde.paho.org/bvsacd/cd56/hydrology/cap3-4.pdf, 2001.

Bruneau, P., Gascuelodoux, C., Robin, P., Merot, P., and Beven, K.: Sensitivity to space and time resolution of a hydrological model using

40   digital elevation data, Hydrological Processes, 9, 69–81, <GotoISI>://WOS:A1995QL75000006, 1995.

Buytaert, W.: Topmodel, http://cran.r-project.org/web/packages/topmodel/, 2015.

Buytaert, W. and Beven, K.: Models as multiple working hypotheses: hydrological simulation of tropical alpine wetlands, Hydrological
     Processes, 25, 1784–1799, http://doi.wiley.com/10.1002/hyp.7936, 2011.

Buytaert, W., De Bièvre, B., Wyseure, G., and Deckers, J.: The use of the linear reservoir concept to quantify the impact of changes in land

45   use on the hydrology of catchments in the Andes Use of linear reservoir concept to quantify the impact of changes in land use on the
     hydrology of catchments in the Andes, Hydrology and Earth System Sciences, 8, 108–114, 2004.

Buytaert, W., Célleri, R., De Biévre, B., Deckers, J., and Wyseure, G.: Modelando el comportamiento hidrológico de microcuencas de
     páramo en el Sur del Ecuador con TOP MODEL, http://www.uazuay.edu.ec/publicaciones/revista3132.pdf#page=149http://paramo.cc.ic.
     ac.uk/pubs/ES/topmodel.pdf, 2005a.

50   Buytaert, W., Wyseure, G., De Bièvre, B., and Deckers, J.: The effect of land-use changes on the hydrological behaviour of Histic Andosols
     in south Ecuador, Hydrological Processes, 19, 3985–3997, http://doi.wiley.com/10.1002/hyp.5867, 2005b.

Buytaert, W., Célleri, R., De Bièvre, B., Cisneros, F., Wyseure, G., Deckers, J., and Hofstede, R.: Human impact on the hydrology of the
     Andean páramos, Earth-Science Reviews, 79, 53–72, http://linkinghub.elsevier.com/retrieve/pii/S0012825206000808, 2006a.

Buytaert, W., Iñiguez, V., and Celleri, R.: Analysis of the water balance of small páramo catchments in south Ecuador, Environmental Role

55   of . . . , http://www.springerlink.com/index/KW365276Q2Q12200.pdf, 2006b.

Buytaert, W., Iñiguez, V., and Bièvre, B. D.: The effects of afforestation and cultivation on water yield in the Andean páramo, Forest Ecology and Management, 251, 22–30, http://linkinghub.elsevier.com/retrieve/pii/S0378112707004641, 2007.

Buytaert, W., Cuesta-Camacho, F., and Tobón, C.: Potential impacts of climate change on the environmental services of humid tropical alpine regions, Global Ecology and Biogeography, 20, 19–33, http://doi.wiley.com/10.1111/j.1466-8238.2010.00585.x, 2011.

60 Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. a., Vrugt, J. a., Gupta, H. V., Wagener, T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, Water Resources Research, 44, n/a–n/a, http://doi.wiley.com/10.1029/2007WR006735, 2008.

Crespo, P., Celleri, R., Buytaert, W., Feyen, J. A. N., Iñiguez, V., Borja, P., Bievre, B. D. E., and Cuenca, U. D.: Land use change impacts on the hydrology of wet Andean páramo ecosystems, 2009.

65 Crespo, P., Feyen, J., Buytaert, W., Célleri, R., Frede, H.-G., Ramírez, M., and Breuer, L.: Development of a conceptual model of the hydrologic response of tropical Andean micro-catchments in Southern Ecuador, Hydrology and Earth System Sciences Discussions, 9, 2475–2510, http://www.hydrol-earth-syst-sci-discuss.net/9/2475/2012/, 2012.

Crespo, P. J., Feyen, J., Buytaert, W., Bücker, A., Breuer, L., Frede, H.-G., and Ramírez, M.: Identifying controls of the rainfall–runoff response of small catchments in the tropical Andes (Ecuador), Journal of Hydrology, 407, 164–174, http://linkinghub.elsevier.com/retrieve/pii/S002216941100477X, 2011.

70 Deginet, M. D.: Land surface representation for regional rainfall-runoff modelling, upper Blue Nile basin, Ethiopia, Master of science, International Institute for Geo-Information Science and Earth Observation, http://www.itc.nl/library/papers_2008/msc/wrem/moges.pdf, 2008.

Díaz-granados, M., Céspedes, D., Tamayo, A., Clavijo, W., and Saénz, J.: SIG en el Estudio de Cuencas de Páramo, in: Congreso Mundial de Páramos, edited by Jaramillo, C. A., Uribe, C., Hincapié, F., Rodríguez, J., and Durán, C., pp. 698–704, Ministerio del Medio Ambiente, Bogotá, 2002.

75

Díaz-Granados, M., González, J., and López, T.: Páramos: Hidrosistemas Sensibles, Revista de Ingeniería Universidad de Los Andes, http://revistaing.uniandes.edu.co/pdf/A822.pdf, 2005.

Duan, Q.-Y. Y., Sorooshian, S., and Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models, http://www.agu.org/pubs/crossref/1992/91WR02985.shtml, 1992.

80

Efstratiadis, A. and Koutsoyiannis, D.: One decade of multi-objective calibration approaches in hydrological modelling: a review, Hydrological Sciences Journal, 55, 58–78, http://www.tandfonline.com/doi/abs/10.1080/02626660903526292, 2010.

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., and Savenije, H. H. G.: A framework to assess the realism of model structures using hydrological signatures, Hydrology and Earth System Sciences, 17, 1893–1912, http://www.hydrol-earth-syst-sci.net/17/1893/2013/, 2013.

85

Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: A comparison of alternative multiobjective calibration strategies for hydrological modeling, Water Resources Research, 43, 1–16, 2007.

Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Understanding catchment behavior through stepwise model concept improvement, Water Resources Research, 44, 1–13, 2008.

90 Fenicia, F., Kavetski, D., and Savenije, H. H. G.: Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, Water Resources Research, 47, 1–13, http://doi.wiley.com/10.1029/2010WR010174, 2011.

Fleming, M. and Doan, J.: HEC-GeoHMS Geospatial Hydrologic Modeling Estension, Tech. rep., US Army Corps of Engineers, Davis, 2013.

Hydrology and
Earth System
Sciences
Discussions

Fleming, M. and Neary, V.: Continuous hydrologic modeling study with the hydrologic modeling system, Journal of Hydrologic Engineering,
95   9, http://ascelibrary.org/doi/pdf/10.1061/(ASCE)1084-0699(2004)9:3(175), 2004.

Frances, F.: Descripción del modelo conceptual distribuido de simulacion hidrológica TETIS v.8, Tech. rep., Universidad Politécnica de
      Valencia, Valencia, 2012.

Franchini, M., Wendling, J., Obled, C., and Todini, E.: Physical interpretation and sensitivity analysis of the TOPMODEL, Journal of
      Hydrology, 175, 293–338, 1996.

100   Franz, K. J. and Hogue, T. S.: Evaluating uncertainty estimates in hydrologic models: borrowing measures from the forecast verification
      community, Hydrology and Earth System Sciences, 15, 3367–3382, http://www.hydrol-earth-syst-sci.net/15/3367/2011/, 2011.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance crite-
      ria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, http://linkinghub.elsevier.com/retrieve/pii/
      S0022169409004843, 2009.

Hofstede, R. G. M., Chilito, E. J. P., and Sandovals, E. M.: Vegetative structure, microclimate, and leaf growth of a páramo tussock grass
      species, in undisturbed, burned and grazed conditions, Vegetatio, 119, 53–65, http://www.springerlink.com/index/H0Q8375G81316U27.
      pdf, 1995.

5   IGAC: Estudio General de Suelos y Zonificacion de Tierras del Departamento de Cundinamarca, Tech. rep., Instituto Colombiano Agustín
      Codazzi, Bogotá, 2000.

INGETEC: Diseños para la construcción de las obras de control de crecientes de la cuenca del Río Tunjuelo, Tech. rep., Empresa de
      Acueducto y Alcantarillado de Bogotá, Bogotá, 2002.

Kavetski, D. and Fenicia, F.: Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights,
10   Water Resources Research, 47, 1–19, 2011.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to ad-
      vance the science of hydrology, Water Resources Research, 42, 1–5, http://seismo.berkeley.edu/{~}kirchner/reprints/
      2006{_}77{_}right{_}answers{_}right{_}reasons.pdf, 2006.

McMillan, H. K., Clark, M. P., Bowden, W. B., Duncan, M., and Woods, R. a.: Hydrological field data from a modeller's perspective: Part 1.
15   Diagnostic tests for model structure, Hydrological Processes, 25, 511–522, http://doi.wiley.com/10.1002/hyp.7841, 2011.

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I—A discussion of principles, Journal of hydrology, 0,
      http://www.sciencedirect.com/science/article/pii/0022169470902556, 1970.

Pfannerstill, M., Guse, B., and Fohrer, N.: Smart low flow signature metrics for an improved overall performance evaluation of hydrological
      models, Journal of Hydrology, 510, 447–458, http://linkinghub.elsevier.com/retrieve/pii/S0022169413009414, 2014.

20   Pokhrel, P., Yilmaz, K. K., and Gupta, H. V.: Multiple-criteria calibration of a distributed watershed model using spatial regularization and
      response signatures, Journal of Hydrology, 418-419, 49–60, http://linkinghub.elsevier.com/retrieve/pii/S0022169408005970, 2012.

Puricelli, M.: Estimación y distribución de parámetros del suelo para la modelación hidrólogica., Ph.D. thesis, Universidad Politécnica de
      Valencia, http://www.riunet.upv.es/handle/10251/2904, 2008.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The
25   challenge of identifying input and structural errors, Water Resources Research, 46, 1–22, 2010.

Reusser, D. E.: Combining smart model diagnostics and effective data collection for snow catchments, Dissertation, University of Postdam,
      2010.

Reyes, O.: Utilización de modelos hidrológicos para la determinación de cuencas en ecosistemas de páramo, Revista Ambiental Agua, Aire y Suelo, pp. 56–65, http://revistas.unipamplona.edu.co/ojs_viceinves/index.php/RA/article/view/432, 2014.

30 Rogelis, M. C. and Werner, M. G. F.: Spatial Interpolation for Real-Time Rainfall Field Estimation in Areas with Complex Topography, Journal of Hydrometeorology, 14, 85–104, http://dx.doi.org/10.1175/JHM-D-11-0150.1, 2012.

Saulnier, G.-M., Obled, C., and Beven, K. J.: Analytical compensation between DTM grid resolution and effective values of saturated hydraulic conductivity within the TOPMODEL framework., 11, 1331–1346, http://www3.interscience.wiley.com/journal/11467/abstract, 1997.

35 Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. a., and Carrillo, G.: Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrology and Earth System Sciences, 15, 2895–2911, http://www.hydrol-earth-syst-sci.net/15/2895/2011/, 2011.

Sene, K.: Flood Warning, Forecasting and Emergency Response, Springer, http://link.springer.com/content/pdf/10.1007/978-3-540-77853-0.pdf, 2008.

40 Sevink, J.: Páramo Andino Project Hydrology workshop in Merida , Venezuela, 2011.

Thiemig, V.: The development of pan-African food forecasting and the exploration of satellite-based precipitation estimates, Ph.D. thesis, http://www.narcis.nl/publication/RecordID/oai%3Adspace.library.uu.nl%3A1874%2F301673, 2014.

Tobón, C. and Gil - Morales, E. G.: Capacidad de interceptación de la niebla por la vegetación de los páramos andinos, Avances en Recursos Hidraulicos, 15, 35–46, 2007.

45 Torres, M.: Efecto del Uso del Suelo en la Capacidad de Almacenamiento Hídrico en el Páramo de Sumapaz - Colombia, Rev. Fac. Nal. Agr. . . . , 67, 7189–7200, 2014.

Torres, V., Vandenberghe, J., and Hooghiemstra, H.: An environmental reconstruction of the sediment infill of the Bogotá basin (Colombia) during the last 3 million years from abiotic and biotic proxies, Palaeogeography, Palaeoclimatology, Palaeoecology, 226, 127–148, http://linkinghub.elsevier.com/retrieve/pii/S0031018205002592, 2005.

50 Uhlenbrook, S., Seibert, J., and Leibundgut, C.: Multi-criterial validation of TOPMODEL in a mountainous catchment Abstract :, 1620, http://www.seibert-space.com/hydro/pdf/Guentner_etal_1999_HP_multi_criteria_TOP.pdf, 1999.

USACE: Hydrologic Modeling System HEC-HMS, Tech. rep., U.S. Army Corps, Davis, 2000.

Vergara, W., Deeb, A., and Leino, I.: Assessment of the impacts of climate change on mountain hydrology: development of a methodology through a case study in the Andes of Peru, Tech. rep., The World Bank, Washington, D.C., 2011.

55 Wu, S., Li, J., and Huang, G. H.: Modeling the effects of elevation data resolution on the performance of topography-based watershed runoff simulation, Environmental Modelling and Software, 22, 1250–1260, 2007.

Yang, W., Andréasson, J., Phil Graham, L., Olsson, J., Rosberg, J., and Wetterhall, F.: Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies, Hydrology Research, 41, 211, http://www.iwaponline.com/nh/041/nh0410211.htm, 2010.

60 Yilmaz, K. K., Gupta, H. V., and Wagener, T.: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, Water Resources Research, 44, n/a–n/a, http://doi.wiley.com/10.1029/2007WR006716, 2008.

Zhang, X., Hörmann, G., Gao, J., and Fohrer, N.: Structural uncertainty assessment in a discharge simulation model, Hydrological Sciences Journal, 56, 854–869, 2011.

Hydrology and
Earth System
Sciences

Open Access

EGU

Discussions

Table 1. Hydrologic characteristics of the páramo area in the Tunjuelo river basin

Component	Description
Forcing data	Horizontal precipitation, fog and mist play an important role in the water balance (Díaz-granados et al., 2002). There are, however, no measurements available for the study area. Rainfall events in the páramo are typically of high frequency and low intensity. In combination with strong winds and a very rough topography (rain shading) this results in high spatial rainfall variability and large errors in precipitation registration (Buytaert et al., 2004). Actual evapotranspiration is low due to the presence of xerophytic plants (Buytaert et al., 2006b), low temperature, high frequency of fog, cloud cover and high relative humidity (Reyes, 2014; Buytaert and Beven, 2011; Buytaert et al., 2011). Literature values of actual evapotranspiration range from 0.8 mm/day to about 1.5 mm/day (Hofstede et al., 1995; Buytaert et al., 2004).
Vegetation	Because of a predominance of grass species, water storage in the vegetation layer is minimal (Buytaert et al., 2005b). However, natural páramo vegetation play an important role in the water cycle with a hydrologic behaviour that is as yet poorly understood (Buytaert et al., 2006a).
Soils	The soils in the páramo area correspond mostly to inceptisols, although andisols and entisols are present. These characteristic páramo soils have a high content of organic matter, high porosity, a large hydraulic conductivity (Buytaert et al., 2006a). Infiltration capacities between 15 and 150 mm/h, and water retention capacities up to 90 vol% in saturated conditions (Buytaert and Beven, 2011). Soils are relatively shallow (about 50 cm). The soils effectively regulate water producing a slow hydrologic response caused by the combination of a high water storage capacity and high conductivity (Buytaert et al., 2005a). Thus the soil acts as a buffering reservoir, and turns the variable rainfall into a continuous water discharge (Buytaert et al., 2004). Changes in soil moisture storage over time are relatively small (Buytaert et al., 2007). There is an abundance of hydrologically disconnected areas because of the irregular topography, which gives rise to a large number of lakes and swamps (Buytaert and Beven, 2011).
Soils	Due to the steep topography, no permanent water table exists, except in local depressions where flows accumulate and permanent saturation occurs. As a result, no significant groundwater is present, and water flow is restricted to overland flow and subsurface flow in the soil layer above the bedrock (Buytaert et al., 2007). Rainfall intensities are commonly lower than infiltration rates (Buytaert et al., 2006b). Thus, infiltration excess overland flow (Hortonian flow) is virtually non-existent. The hydrological regime is dominated by a slow flow response (Buytaert et al., 2007). Vertical infiltration through the soil is dominant during the beginning of rainfall events, and dependent on the antecedent soil moisture conditions. By contrast, during low intensity rainfall events, preferential flow is dominant between the organic horizon and the underlying mineral horizon or the bedrock. Saturation excess surface flow is only observed during long rainfall events; otherwise near sub-surface lateral flow in the organic layer occurs during peaks (Crespo et al., 2009). On the other hand, surface roughness and local depressions are important in delaying surface runoff (Buytaert et al., 2006a).
Base flow	Base flow is relatively constant during the year (Buytaert et al., 2004), due to the climate, topography and soils (Buytaert et al., 2007). Thus, the hydrological regime of the natural catchment is dominated by a slow base flow response (Buytaert et al., 2007).
Deep percolation	Subsurface groundwater is nearly absent because of the presence of impermeable bedrock (Buytaert et al., 2005b) and the lack of a groundwater storage system. Due to mountainous terrain and the impenetrable bedrock, deep percolation is negligible, and the major hydrological processes occur in the soil layer (Buytaert et al., 2004).

Table 2. Correction factors of the TETIS model

Correction Factor	Parameter corrected by the factor
FC_1	Static storage
FC_2	Evapotranspiration
FC_3	Hydraulic conductivity of the soil
FC_4	Surface flow velocity in the hills
FC_5	Percolation
FC_6	Horizontal saturated conductivity
FC_7	Deep percolation
FC_8	Horizontal saturated conductivity of the substrate
FC_9	Wave velocity

Table 3. 12 most sensitive HECHMSSMA calibration parameters

Parameter	Description
MaxSoilInfil	Soil maximum infiltration
MaxSoilStore	Maximum volume of the soil storage
TenStore	Tension storage
ClarkSC	Storage coefficient for the Clark's unit hydrograph
MaxSoilPerc	Maximum Soil Percolation
MaxGWStore1	Maximum Soil Percolation
MaxGWStore2	Maximum ground waters storage capacity in ground water layer 1
RoutGWStore1	Maximum ground waters storage capacity in ground water layer 2
RoutGWStore2	Groundwater flow routing coefficient in ground water layer 1
MaxPercGw1	Groundwater flow routing coefficient in ground water layer 2
MaxSoilPerc	Maximum percolation rate in ground water layer 1
MaxPercGw2	Maximum percolation rate in ground water layer 2
RoutLR12	Routing coefficient for linear reservoir 1 for baseflow
RoutLR22	Routing coefficient for linear reservoir 2 for baseflow

Table 4. Calibration parameters of the TOPMODEL

Correction Factor	Parameter corrected by the factor
lnTe	Logarithm of the areal average of the transmissitivy
m	Model parameter controlling the rate of decline of transmissivity in the soil profile
Sr0	Initial root zone storage deficit
Srmax	Maximum root zone storage deficit
td	Unsaturated zone time delay per unit storage deficit
vr	Channel flow velocity inside catchment

Table 5. Signature measures from the FDC (Yilmaz et al., 2008; Pokhrel et al., 2012). Q_S and Q_O correspond to simulated and observed flows. The sub indices indicate: m1 and m2 are 0.2 and 0.7 flow exceedance probabilities; h=1,2,...H are the flow indices for flows with exceedance probabilities lower than 0.2; l=1,2,...,L is the index of the flow value located within the flow-flow segment of the FDC (0.7-1.0 flow exceedance probabilities); QS_{med} corresponds to the median value of the simulated flows and QO_{med} to the median value of the observed flows.

| Signature		Description
$BiasFMS = \dfrac{[log\,(QS_{m1}) - log\,(QS_{m2})] - [log\,(QO_{m1}) - log\,(QO_{m2})]}{[log\,(QO_{m1}) - log\,(QO_{m2})]} \times 100$	(4)	Quantifies the % difference in the mid-segment slope of the FDC. Positive values imply that the slope of the middle portion of the simulated FDC is higher than the slope of the observed FDC.
$BiasFHV = \dfrac{\sum_{h=1}^{H} (QS_h - QO_h)}{\sum_{h=1}^{H} QO_h} \times 100$	(5)	Quantifies % volume bias in the highest 20% of the flows
$BiasFLV = -1 \cdot \dfrac{\sum_{l=1}^{L} [log\,(QS_l) - log\,(QS_L)] - \sum_{l=1}^{L} [log\,(QO_l) - log\,(QO_L)]}{\sum_{l=1}^{L} [log\,(QO_l) - log\,(QO)]} \times 100$	(6)	Quantifies the % volume bias in the lowest 30% of the flows
$BiasFMM = \dfrac{log\,(QS_{med}) - log\,(QO_{med})}{log\,(QO_{med})} \times 100$	(7)	Quantifies the % difference in the median flow
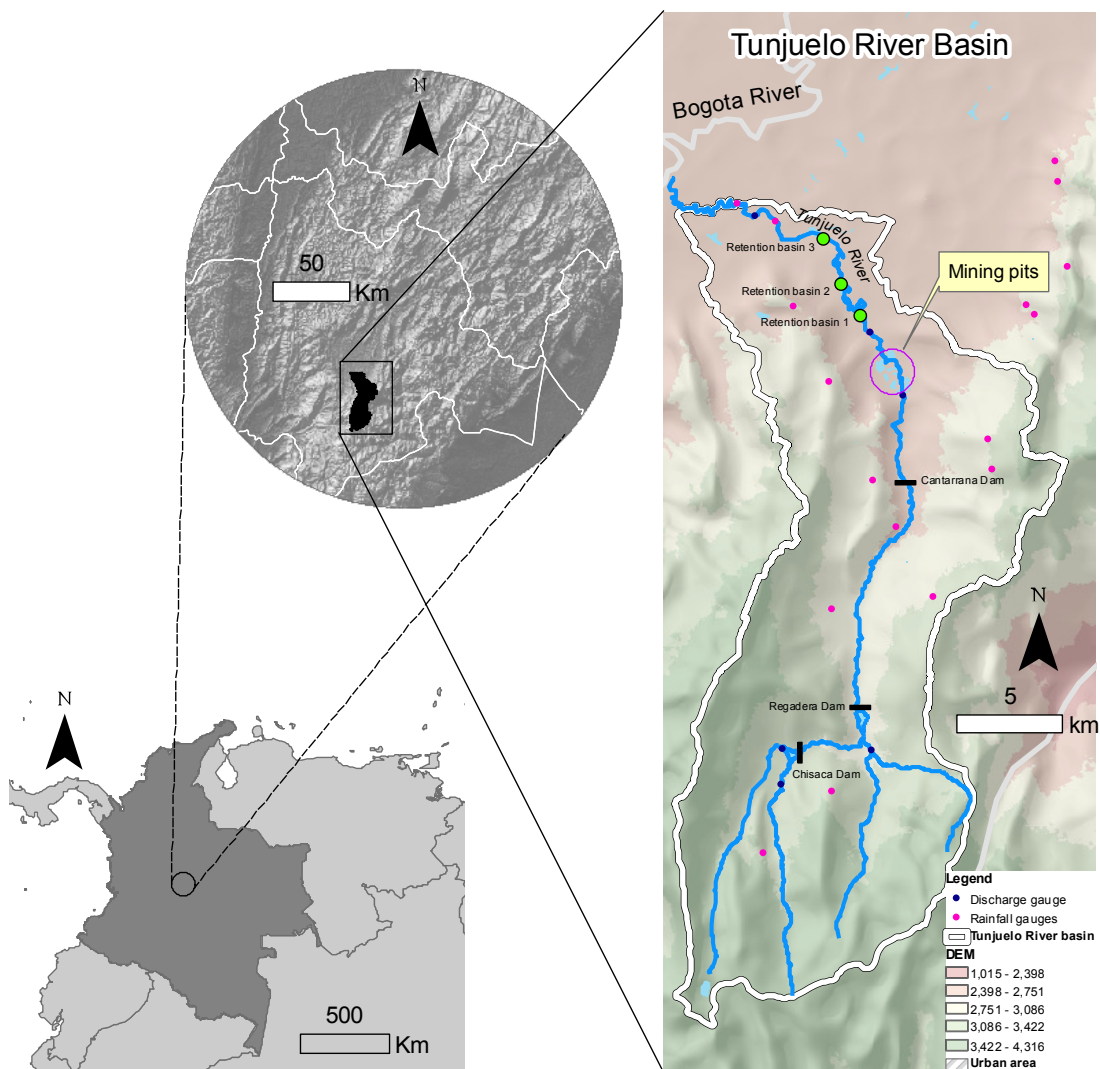
Figure 1. Study area
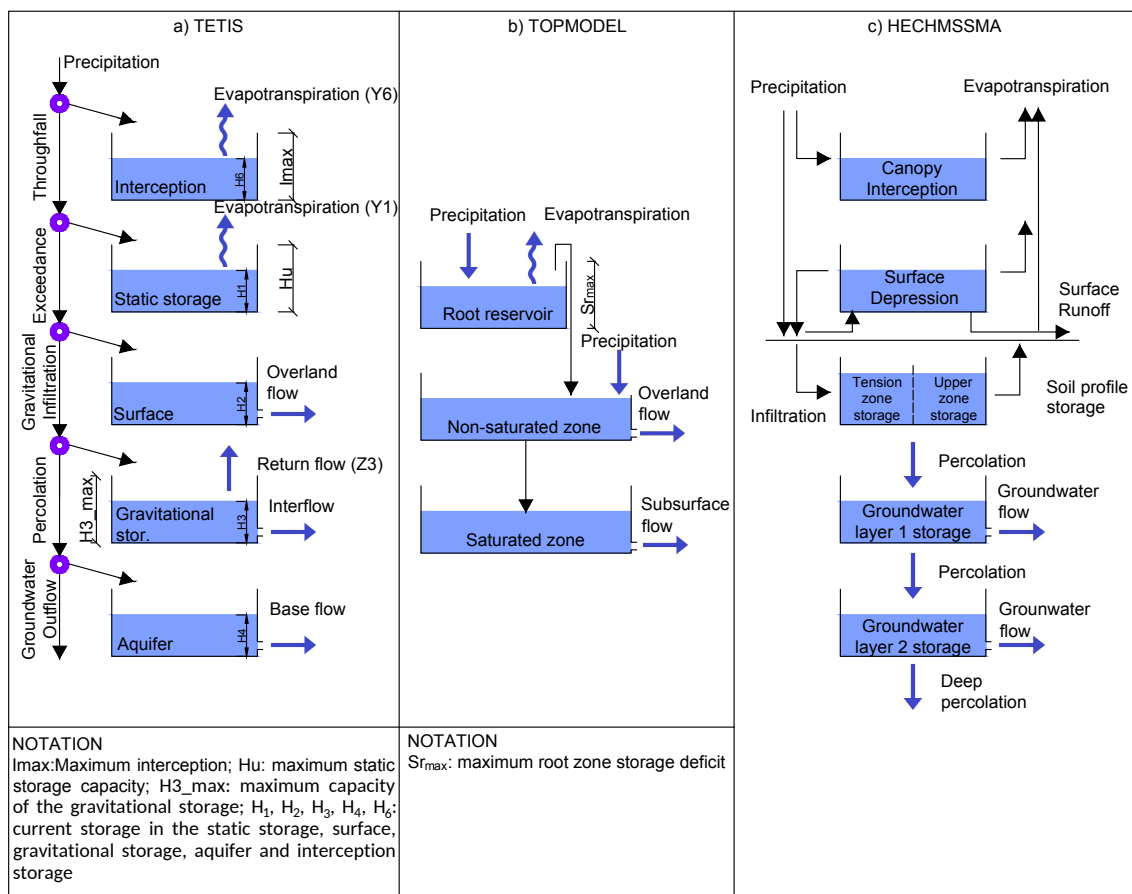
Figure 2. a) Conceptual tanks at cell level in TETIS, b) conceptual tanks TOPMODEL, c) conceptual tanks HECHMS SMA model
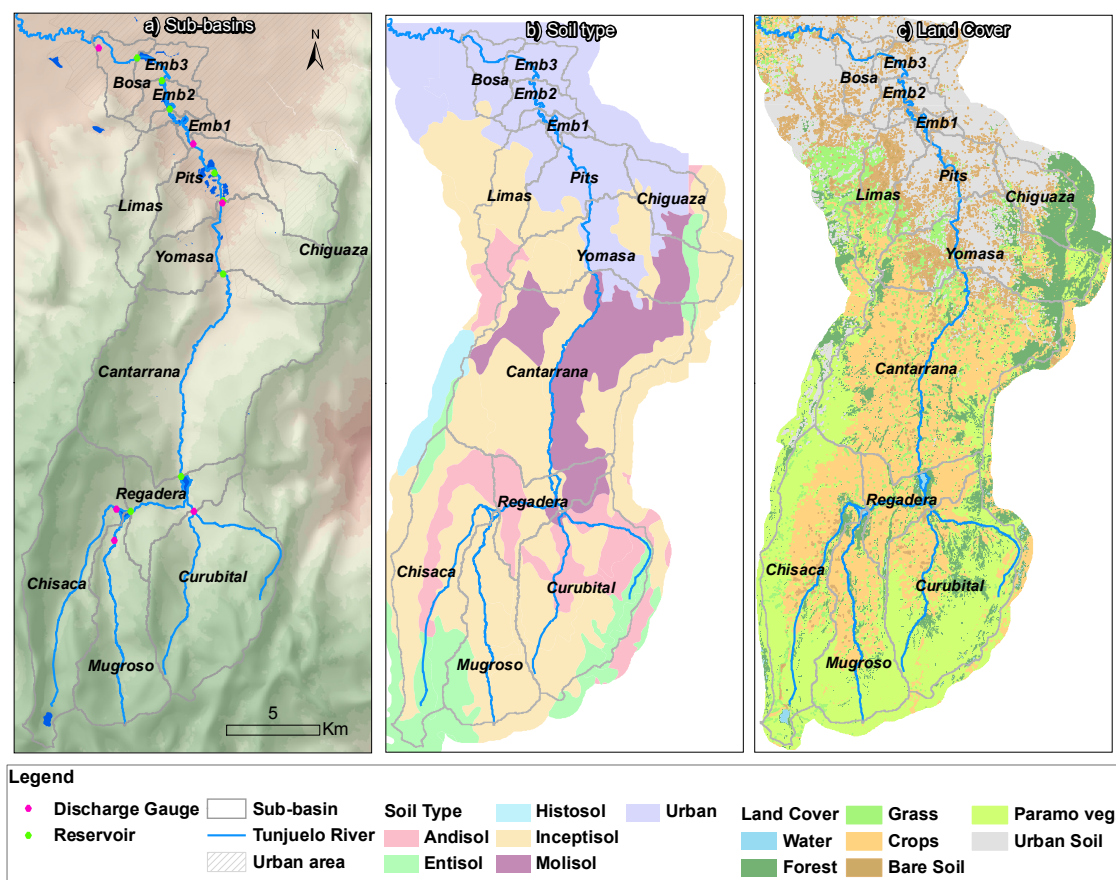
Figure 3. a) Sub-basin division; b) Soil types (source: (IGAC, 2000)); c) Land cover
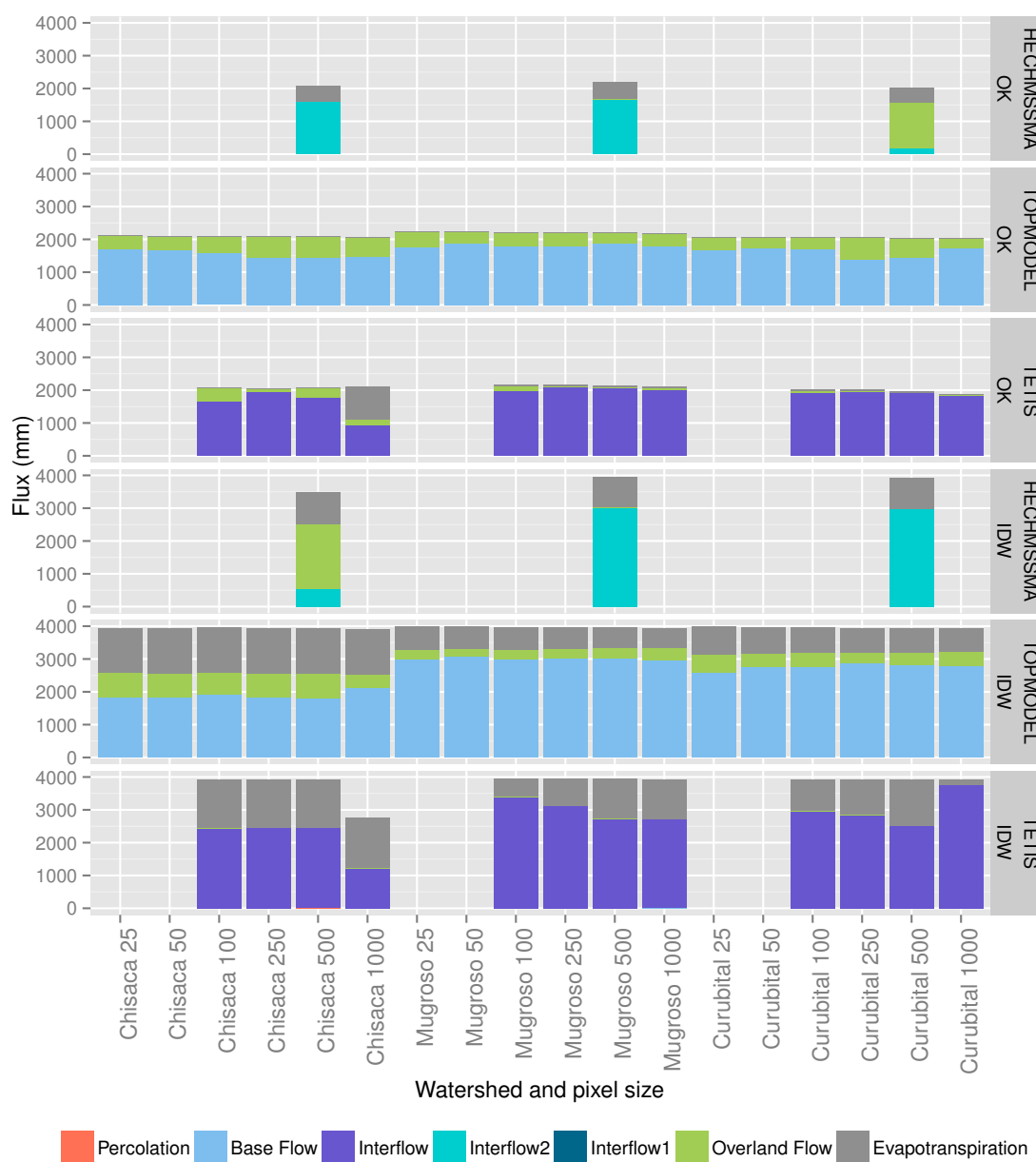
Figure 4. Fluxes obtained from the models
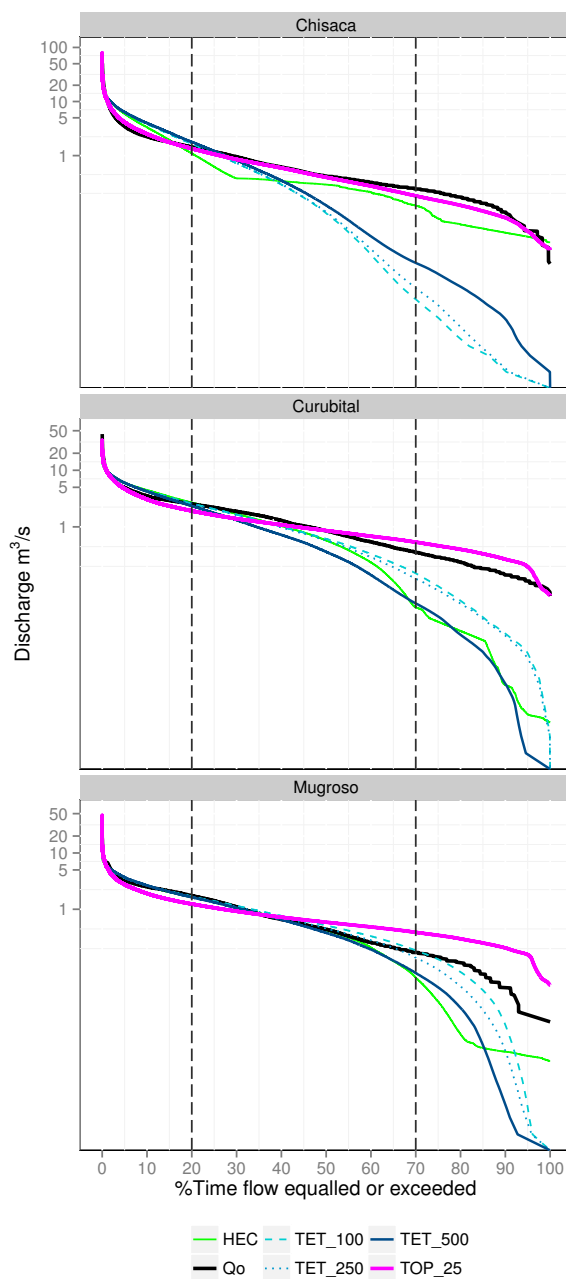
Figure 5. Flow duration curves and signature measures
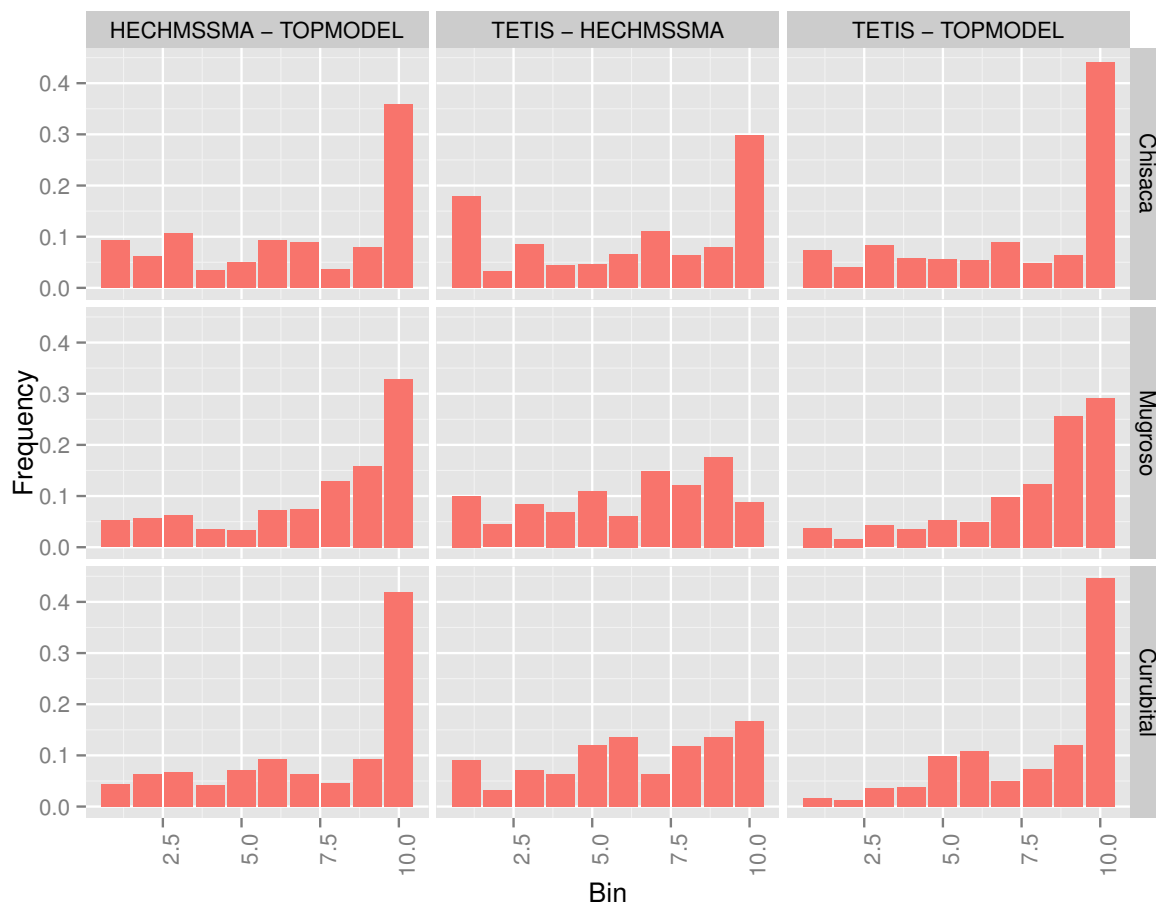
Figure 6. Rank histograms for the head watersheds in the páramo area for the three models. The bins were reduced to 10 for better visualization. The first model in the title corresponds to the reference ensemble, e.g. in HECHMSSMA - TOPMODEL the reference ensemble corresponds to the results of the HECHMSSMA model and the ensemble to be compared is the obtained from TOPMODEL simulations.

Figure 7. Water balances Chisaca, Curubital and Mugroso watersheds

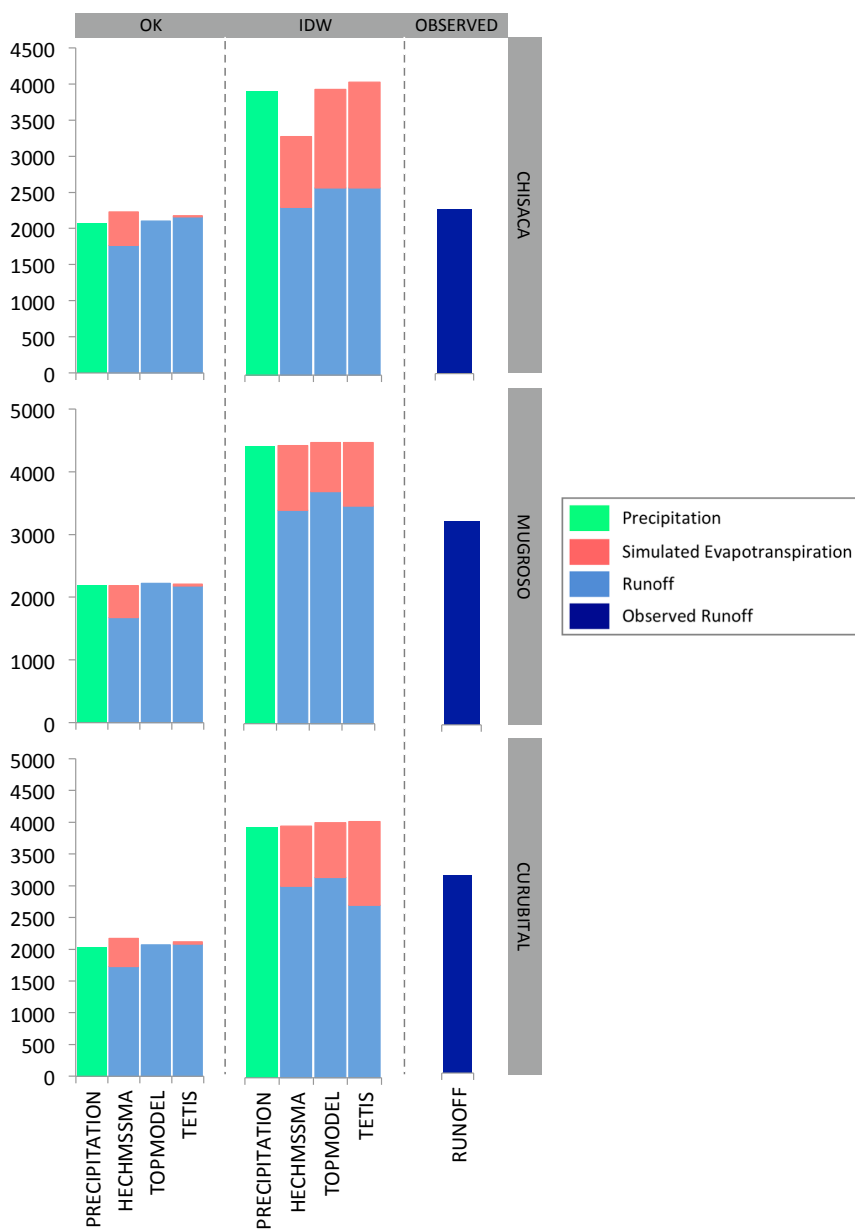| Model	Cantarrana			Chisaca			Curubital			Independencia			Mugroso			SnBenito		
	OK	IDW	OKbc	OK	IDW	OKbc	OK	IDW	OKbc	OK	IDW	OKbc	OK	IDW	OKbc	OK	IDW	OKbc
TETIS																		
100	0.85	0.81		0.57	0.60		0.44	0.63		0.84	0.84		0.43	0.64		0.64	0.75	
250	0.85	0.80		0.56	0.59		0.45	0.64		0.84	0.83		0.46	0.64		0.64	0.70	
500	0.85	0.81	0.78	0.59	0.58	0.57	0.45	0.65	0.59	0.84	0.84	0.79	0.51	0.67	0.66	0.67	0.75	0.75
1000	0.84	0.80		0.64	-0.24		0.25	0.59		0.84	0.84		0.51	0.67		0.79	0.84	
2000	0.85	0.78		0.62	-0.22		0.41	0.63		0.85	0.84		0.50	0.69		0.90	0.90	
Topmodel																		
25				0.54	0.57	0.59	0.43	0.62	0.59				0.46	0.63	0.65			
50				0.57	0.57	0.58	0.43	0.62	0.59				0.45	0.62	0.65			
100				0.58	0.57	0.58	0.43	0.61	0.58				0.45	0.62	0.65			
250				0.58	0.57	0.57	0.42	0.61	0.58				0.45	0.62	0.64			
500				0.57	0.55	0.56	0.41	0.61	0.58				0.44	0.60	0.63			
1000				0.51	0.50		0.42	0.61					0.42	0.58				
HEC-HMS																		
SMA	0.78	0.74	0.73	0.44	0.58	0.54	0.34	0.65	0.56	0.88	0.91	0.91	0.38	0.65	0.6	0.60	0.67	0.66

Table 6. Kling and Gupta coefficient obtained from calibration

Model	FHV	FLV	FMS	FMM
Chisaca				
TETIS 100	31.62	274.11	270.92	181.95
TETIS 250	32.68	261.76	245.22	175.84
TETIS 500	34.96	186.09	188.52	144.07
HECHMSSMA	22.28	34.20	24.33	41.95
TOPMODEL 25	5.06	10.70	12.72	6.24
TOPMODEL 50	4.01	11.31	13.45	6.89
TOPMODEL 100	3.65	11.51	14.39	9.37
TOPMODEL 250	2.61	11.32	13.34	9.77
TOPMODEL 500	−1.68	17.93	17.90	11.06
Mugroso				
TETIS 100	−2.05	54.77	−4.83	−19.12
TETIS 250	−2.54	75.47	7.33	−4.85
TETIS 500	−1.77	113.77	33.25	17.80
HECHMSSMA	−6.43	76.04	41.61	4.98
TOPMODEL 25	−20.16	−47.42	−50.10	−35.34
TOPMODEL 50	−20.57	−48.36	−50.66	−37.36
TOPMODEL 100	−20.57	−47.90	−50.50	−36.12
TOPMODEL 250	−20.72	−48.60	−50.80	−37.78
TOPMODEL 500	−21.57	−49.92	−52.02	−41.47
Curubital				
TETIS 100	4.98	127.12	45.72	207.02
TETIS 250	5.71	135.51	56.64	244.07
TETIS 500	2.67	247.47	103.26	533.48
HECHMASMA	5.87	213.19	118.50	245.71
TOPMODEL 25	−10.33	−27.66	−34.88	−5.34
TOPMODEL 50	−9.91	−30.33	−36.54	−34.56
TOPMODEL 100	−10.05	−32.26	−37.53	−46.09
TOPMODEL 250	−10.10	−32.60	−37.65	−54.08
TOPMODEL 500	−9.93	−32.30	−37.93	−51.87

Table 7. Flow duration curve signature measures

Hydrology and
Earth System
Sciences

Open Access

EGU

Discussions

Table 8. IQR, MAD and range of ensemble discharges for HECHMSSMA, TOPMODEL and TETIS

	IQR	MAD	Range	IQR_{peak}	MAD_{peak}	$Range_{peak}$
Curubital HMS	3.70	1.89	13.45	6.36	3.31	22.94
Curubital TET	3.58	1.78	14.75	7.19	3.60	28.27
Curubital TOP	4.16	2.00	16.74	15.03	7.27	68.47
Mugroso HMS	1.39	0.66	5.75	2.20	1.06	9.11
Mugroso TET	1.62	0.69	7.06	2.78	1.24	11.87
Mugroso TOP	1.77	0.79	8.62	6.17	2.74	36.59
Chisaca HMS	4.19	2.17	17.56	10.02	5.50	40.94
Chisaca TET	3.27	1.72	16.66	8.16	4.29	40.75
Chisaca TOP	3.80	1.85	17.86	16.00	8.06	76.31