Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 154 (2016) 1116 – 1123

Procedia Engineering

www.elsevier.com/locate/procedia

12th International Conference on Hydroinformatics, HIC 2016

Development of the HBV-TEC hydrological model

M. Mendez[a], L. Calvo-Valverde[b],*

[a]*Escuela de Ingenieria en Construccion, Instituto Tecnologico de Costa Rica, 159-7050, Cartago, Costa Rica*
[b]*Escuela de Ingenieria en Computacion, Instituto Tecnologico de Costa Rica, 159-7050, Cartago, Costa Rica*

Abstract

In this paper, the HBV-TEC hydrological model is presented. The development of this model version, aims to provide researchers and scholars with a stable and robust implementation of the HBV hydrological model based on the R programming language. To evaluate its performance, the HBV-TEC model was applied to three subcatchments of the Aguacaliente river catchment, an experimental catchment located in the province of Cartago, Costa Rica. Results suggest a satisfactory performance of the model for two subcatchments and an unsatisfactory performance of the remaining subcatchment; most of which could be attributed to unsufficient meteorological data along with a highly heterogeneous spatial rainfall-distribution.

Peer-review under responsibility of the organizing committee of HIC 2016

Keywords: calibration; HBV-TEC; hydrological model; PEST; R; validation

1. Introduction

Hydrological models have become essential tools to study the response of hydrological systems to various natural and anthropogenic forcings. Models are simplified description of a real world hydrological system and their level of complexity greatly depends on the model structure and the model objectives [1]. Hydrological models are mainly classified into two main groups, physically-based and conceptual models [2]. Conceptual models approximate key hydrological processes through simplified equations and lumped parameters that represent spatially averaged

---

* Corresponding author. Tel.: +506-2550-2425; fax: +506-2552-5333.
E-mail address: mamendez@itcr.ac.cr

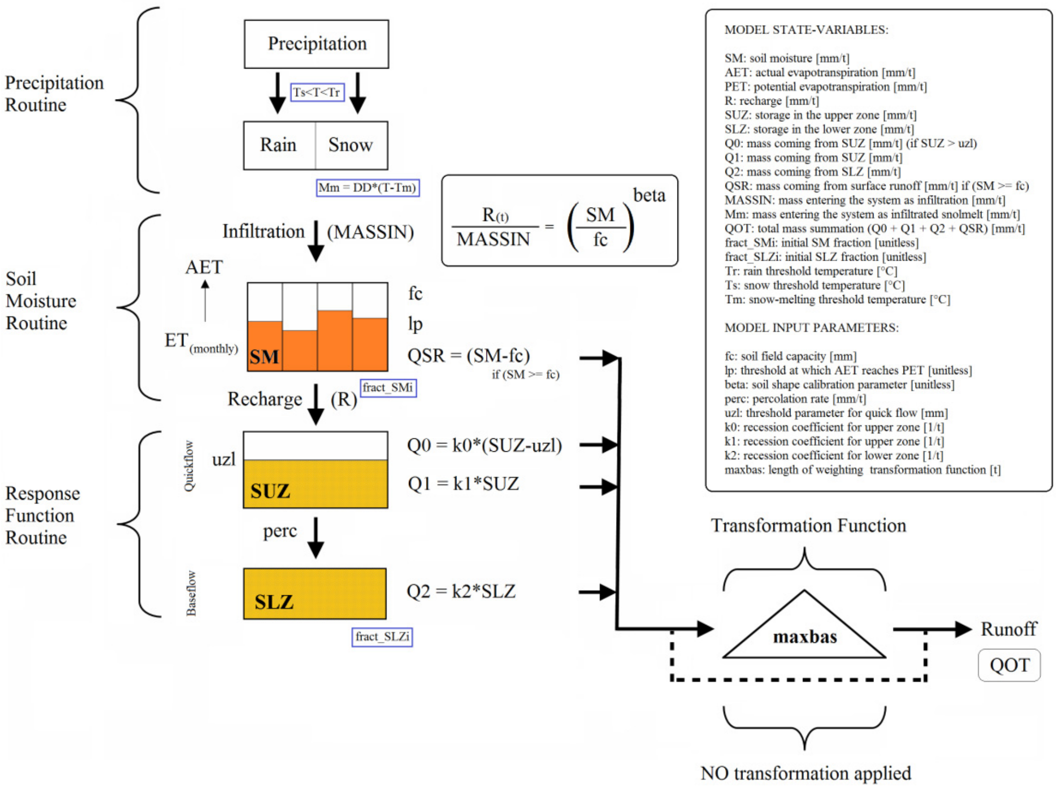
Peer-review under responsibility of the organizing committee of HIC 2016
doi:10.1016/j.proeng.2016.07.521

characteristics in a hydrological system [3]. Conceptual models have several advantages over physically-based models, including more moderate input data requirements and reduced computational resources to setup, run and calibrate the model [4-5]. This presents an opportunity for constant innovation and development of conceptual models using modern and robust programming languages. At the same time, the use of programming languages enables users and programmers to take full advantage of specialized libraries and functions, which can then be incorporated into the models. This project seeks to produce a new and updated code of the HBV conceptual hydrological model based on the R programming language [6]. R is an open-source interpreted programming language mainly designed to perform statistical analysis, graphics and machine learning. R was chosen based on various criteria, including (a) availability of specialized libraries, (b) permanent support and development, (c) open-domain implementation, (d) cross-platform compatibility and (e) simple and flexible syntax coding.

2. Methodology

2.1. Model description

The HBV-TEC is a redesign of the HBV (*Hydrologiska Byråns Vattenbalansavdelning*) model [7-8] developed using the R programming language [6]. Similar to the original HBV version, the HBV-TEC is a semi-distributed conceptual rainfall-runoff model for continuous calculation of runoff (Fig.1). The basic concept is that discharge is related to storage through a conservation of mass equation and a transformation routine. The hydrologic response is easily modelled due to the use of lumped/semi-distributed data and a simplified conceptual representation of flow processes. The structure of the HBV-TEC model consists of routines for precipitation, soil moisture, response function and transformation. The model can be run using daily or hourly time-steps; input data are precipitation, air temperature and long-term estimates of monthly potential evapotranspiration.



MODEL STATE-VARIABLES:

SM: soil moisture [mm/t]
AET: actual evapotranspiration [mm/t]
PET: potential evapotranspiration [mm/t]
R: recharge [mm/t]
SUZ: storage in the upper zone [mm/t]
SLZ: storage in the lower zone [mm/t]
Q0: mass coming from SUZ [mm/t] (if SUZ > uzl)
Q1: mass coming from SUZ [mm/t]
Q2: mass coming from SLZ [mm/t]
QSR: mass coming from surface runoff [mm/t] if (SM >= fc)
MASSIN: mass entering the system as infiltration [mm/t]
Mm: mass entering the system as infiltrated snolmelt [mm/t]
QOT: total mass summation (Q0 + Q1 + Q2 + QSR) [mm/t]
fract_SMi: initial SM fraction [unitless]
fract_SLZi: initial SLZ fraction [unitless]
Tr: rain threshold temperature [°C]
Ts: snow threshold temperature [°C]
Tm: snow-melting threshold temperature [°C]

MODEL INPUT PARAMETERS:

fc: soil field capacity [mm]
lp: threshold at which AET reaches PET [unitless]
beta: soil shape calibration parameter [unitless]
perc: percolation rate [mm/t]
uzl: threshold parameter for quick flow [mm]
k0: recession coefficient for upper zone [1/t]
k1: recession coefficient for upper zone [1/t]
k2: recession coefficient for lower zone [1/t]
maxbas: length of weighting transformation function [t]

Fig. 1. The HBV-TEC hydrologicalmodel structure. Model input parameters and state variables are indicated.

The precipitation routine represents either rainfall or snowfall. In the case of rainfall, precipitation is directly used as infiltration input (*MASSIN*). In the case of snowfall, this can either be accumulated or melted into effective rainfall (and then infiltrated) depending on threshold temperatures *Tr*, *Ts* and *Tm* (rainfall, snowfall and melting temperatures correspondingly). The parameter degree-day (*DD*) indicates decrease of the water-content in the snowpack caused by 1°C above the freezing threshold per time-step. The soil moisture routine controls changes in soil moisture content (*SM*) of the catchment based on infiltration coming from the precipitation routine. These changes are controlled by parameters *fc* (soil field capacity), *lp* (threshold at which *AET* reaches *PET*) and *beta* (soil shape calibration parameter). *SM* is also compensated by actual evapotranspiration losses (*AET*). If soil moisture (*SM*) is greater or equal than soil field capacity (*fc*) in one time-step, a surface-runoff mass balance component (*QSR*) is estimated and directly routed to the transformation function. Infiltration that cannot be stored, routed or evapotranspirated, is used as input for the response function routine, as represented by parameter *R* (recharge). Model state-variable *fract_SMi* (initial *SM* fraction) can either be input by the user or calculated, based on a defined warming period.

The response function routine which receives recharge (*R*) from the soil moisture routing, is comprised of two reservoirs; *SUZ* (storage in the upper zone) and *SLZ* (storage in the lower zone), which represent quickflow and baseflow mass balance components respectively. Outflow from *SUZ* leaves the upper reservoir through three possible mass balance paths; *Q0*, which describes a linear discharge relationship controlled by recession coefficient *k0* (if a threshold *uzl* is exceeded); *Q1*, also a linear discharge relationship controlled by recession coefficient *k1* (if a threshold *uzl* is not exceeded); and a constant parameter *perc* (percolation rate). Outflow from the baseflow reservoir (*SLZ*) is determined by percolation coming from *SUZ* reservoir and *k2*, which controls the linear discharge relationship described by *Q2*. Model state-variable *fract_SLZi* (initial SLZ fraction) can either be input by the user or calculated, based on a defined warming period. *QOT*, which represents the summation of contributing mass balance components (*QSR* + *Q0* + *Q1* + *Q2*) is finally routed through the transformation function. A weighting parameter (*maxbas*), controls the base in an triangular weighting function, which represents the duration (*length*) of the transformation routine itself. If the user determines that no transformation is to be applied, total runoff would be equal to *QOT*. Relevant R packages used in the development of HBV-TEC include: *base*, *DescTools*, *doParallel*, *dplyr*, *foreach*, *ggplot2*, *grid*, *gridExtra*, *gstat*, *lubridate*, *MASS*, *pastecs*, *plyr*, *RColorBrewer*, *reshape* and *visreg*.

2.2. Optimization process and model performance

The non-linear parameter estimation and optimization package PEST [9] is presently being used for parameter optimization and sensitivity analysis of the HBV-TEC model, other optimization methods are currently being developed. PEST uses the gradient-based Gauss-Marquardt Levenberg (GML) algorithm which searches for optimum values of model parameters by minimizing the deviations between field observations and modelled values. For this study, a total of 9 HBV-TEC model parameters (*perc*, *uzl*, *k0*, *k1*, *k2*, *maxbas*, *fc*, *lp* and *beta*) were included in the optimization process. The selection of these parameters followed the next reasons: (a) to minimize compensating error, the equifinality problem needed to be reduced by constraining the parameter space [10]; and (b) only parameters having a direct influence on runoff generation were to be considered. The chosen parameters control the total volume and shape of the hydrographs and are associated with the response, routing and soil moisture routines of the HBV-TEC model. Parameter optimization ranges were selected based on recommended literature values. [7-8]. Table 1 shows the various objective functions used to evaluate the performance of the HBV-TEC model. The correlation coefficient (R^2) was also included in the analysis. The NS_{eff} efficiency is a normalized statistic that determines the relative magnitude of the residual variance compared to the observed data variance. The $LNNS_{eff}$ efficiency is aimed to reduce NS_{eff} sensitivity to extreme values (mainly peak flows) through the logarithmic transformation of the data sets. The APB is a measure of the timing difference between the observations and the modeled values and it is usually used in conjunction with the PBIAS, *which is* commonly used t*o quantify water balance errors*. A mismatch in timing could be indicated by observed and simulated series where the PBIAS value is small and the APB is large. For the purposes of this study, values of $NS_{eff} \geqslant 0.80$, $LNNS_{eff} \geqslant 0.80$, PBIAS $\leqslant \pm 5\%$ and APB $\leqslant 20\%$, were considered satisfactory.

Table 1. Objective functions used to evaluate the performance of the HBV-TEC model.

Abbr.	Objective function	Equation
NS_{eff}	The Nash and Sutcliffe efficiency criterion (expressed as fraction)	$NS_{eff} = 1 - \frac{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_i^{\mathrm{mod}}\right)^2}{\sum_{i=1}^{n}\left(Q_i^{obs} - \overline{Q}_i^{obs}\right)^2}$
$LNNS_{eff}$	The Nash and Sutcliffe efficiency with logarithmic values (expressed as fraction)	$LNNS_{eff} = 1 - \frac{\sum_{i=1}^{n}\left(LNQ_i^{obs} - LNQ_i^{\mathrm{mod}}\right)^2}{\sum_{i=1}^{n}\left(LNQ_i^{obs} - \overline{LNQ}_i^{obs}\right)^2}$
PBIAS	The Percent Bias (expressed as percentage)	$PBIAS = \frac{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_i^{\mathrm{mod}}\right)\cdot 100}{\sum_{i=1}^{n}\left(Q_i^{obs}\right)}$
APB	The Absolute Percent Bias (APB) (expressed as percentage)	$APB = \frac{\sum_{i=1}^{n}\left\lvert Q_i^{obs} - Q_i^{\mathrm{mod}}\right\rvert}{\sum_{i=1}^{n}\left(Q_i^{obs}\right)}$

where i is the timestep, n is the total number of time-steps, Q is the discharge and subscripts *obs* and *mod* refer to observed and modeled correspondingly.

2.3. Case study and data sources

The Aguacaliente River catchment (179.82 km^2) is located in the province of Cartago, central Costa Rica (Fig.2). The topography is mountainous with elevations ranging from 2902 to 1175 m. The mean annual rainfall range of the area is between 1270 and 1530 mm and the mean annual temperature range is between 16.1 and 25.2 ºC [11].

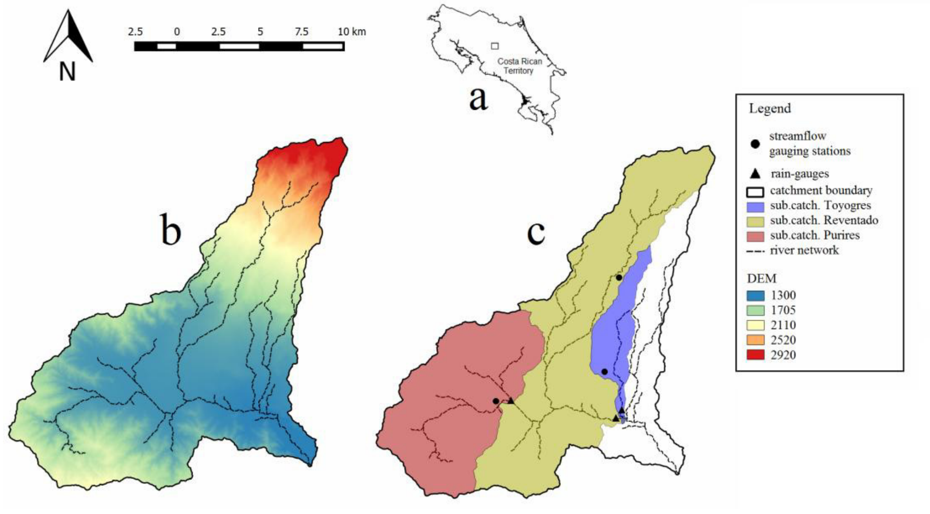


Fig. 2. (a) Position of the Aguacaliente river catchment in Costa Rica; (b) Catchment boundary, river network, WorldView digital elevation model (DEM), (c) Subcatchments, rain-gauges and streamflow gauging stations.

Sparse historical records, mainly concerning streamflow data, motivated the recent instrumentation of the catchment. This included the installation of 3 rain-gauges and 3 streamflow gauging stations. For the purpose of this study, and based on the location of the streamflow gauging stations; the Aguacaliente catchment was discretized into 3 subcatchments. This includes the rivers: Reventado (82.38 km^2 up to gauging station), Purires (55.93 km^2) and Toyogres (12.38 km^2). Rainfall and temperature were calculated from daily measurements at the 3 available rain-gauges using a powered-3 Inverse distance weighting (IDW) interpolation method for a period of 675 days (2013-2015). In consequence, 300 days was used for calibration, 225 days were used for validation and 150 days were excluded as part of the warming period. The catchment boundary was delineated using a WorldView 1m digital elevation model (DEM) resampled to 10 m. All spatial data processing was executed using the R programming language along with specialized R packages (mainly *gstat*, *sp* and *raster*). The monthly long-term mean potential evapotranspiration records were calculated using the Penman-Monteith method.

3. Results and discussion

3.1. Model performance

In the light of a significantly recent, short and spatially sparse input dataset, HBV-TEC model performance varies noticeably among all subcatchments for both calibration and validation periods (Table 2). Regarding calibration, the NS_{eff} efficiency, reaches a satisfactory value for the Purires subcatchment (0.906), barely satisfactory for the Toyogres subcatchment (0.780) and unsatisfactory for the Reventado subcatchment (0.612). $LNNS_{eff}$ and R^2 follow a similar pattern, with their most acceptable values describing the Purires subcatchment and the least satisfactory values describing the Reventado subcatchment. Most of this variance could be attributed to insufficient meteorological data along with a highly heterogeneous spatio-temporal rainfall-distribution. Nonetheless, it seems that the *Linda-Vista* rain-gauge carries sufficient information to properly describe most of the spatio-temporal rainfall-distribution variation of the Purires subcatchment. A similar situation could be happening with the *Llano-Grande* rain-gauge and the Toyogres subcatchment. In the case of the Reventado subcatchment however, 3 rain-gauges seem to be insufficient to capture the rainfall spatio-temporal distribution of the area (82.38 km^2); particularly on the northern side of the catchment, which exhibits elevations over 2900 m and lacks instrumentation (Fig. 2). Despite the differences in model efficiency shown by NS_{eff}, $LNNS_{eff}$, and R^2; errors in water balance, as stated by PBIAS, stay well below the 5% threshold regardless of the subcacthment. APB also shows a similar trend, with values below the adopted 20% threshold. This suggests that the HBV-TEC model is capable of accounting for a satisfactory water balance even on the grounds of scarce input data. Conversely, model performance is generally poorer at the validation stage, as revealed by the outcome of NS_{eff}, $LNNS_{eff}$ and R^2 (Table 2). This is particularly true for the Purires and the Toyogres subcatchments but so much for the Reventado subcatchment.

Table 2. Daily time-steps performance of the HBV-TEC hydrological model for three subcatchments of the Aguacaliente river catchment as evaluated by various objective functions during a period of 675 days (2013-2015).

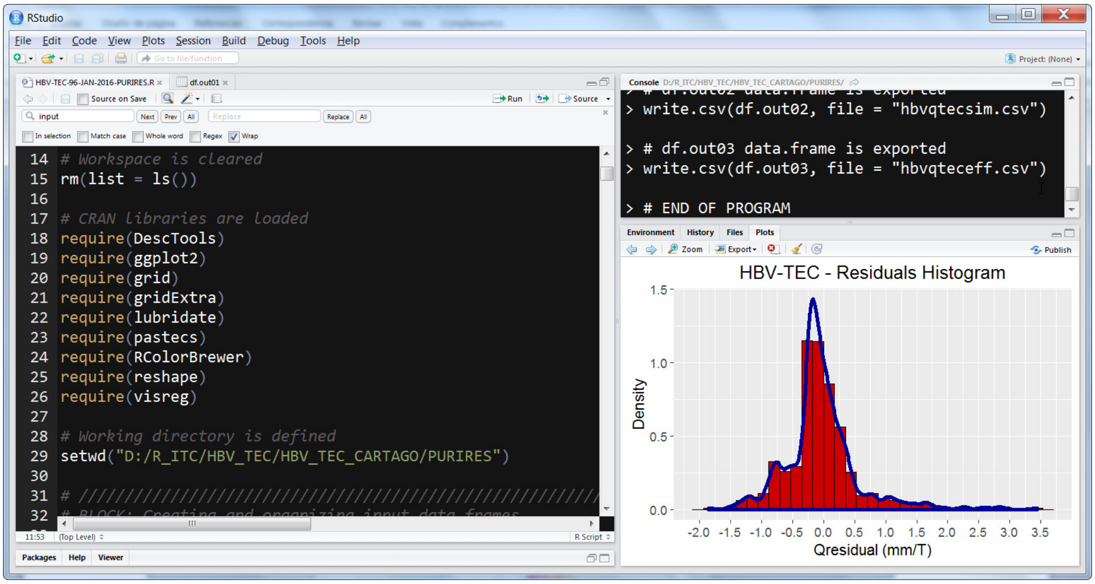
	Subcatchment					
	Reventado		Purires		Toyogres	
Objective.Function	Calibration	Validation	Calibration	Validation	Calibration	Validation
NSeff	0.612	0.580	0.906	0.763	0.780	0.616
LNNSeff	0.640	0.430	0.875	0.637	0.804	0.681
PBIAS	1.270	3.434	2.383	4.240	0.930	1.359
APB	16.201	17.609	16.120	21.132	9.720	14.068
R^2	0.693	0.547	0.908	0.716	0.787	0.638

These results imply a lower agreement between observed and modelled values during the validation period, which might be partly caused by such a short validation period (225 days). Establishing sufficiently long and homogeneous input datasets for both calibration and validation of hydrological models is imperative [11]. Nevertheless, this is a real challenge in developing countries, and Costa Rica is not the exception. In this case, lower model performance in validation goes beyond the HBV-TEC model structure and is more related to the quality, quantity and extension of input data.

3.2. Model results and presentation

The main reason that motivated the rewriting of the HBV code, was to take advantage of the data analysis and high-level visualization capabilities available in R. As R is an interpreted programming language, scripts must be run using an *interpreter*, which in this case is the R console. However, the authors recommend running HBV-TEC under *RStudio*-IDE [13], as it significantly facilitates interaction with the model (Fig. 3).

The results of the HBV-TEC model are both displayed on the screen and written to output files; which include either ASCII format time-series (mainly CSV) or PNG raster images. As an example, a monthly modelled-residuals boxplot is presented for the Purires subcatchment for the period 2013-2015 (Fig. 4(a)). As shown, modelled residuals are more spread-out for the months of June, September and October, probably to more extreme rainfall events. Additionally, a normalized mass-balance summary graph is also presented for the Purires subcatchment (Fig. 4(b)). This graph shows the various contributions of the general mass balance (mm) for the entire period as a fraction of *outMASSIN* (mass entering the systems as rainfall). In this case, *outQOBS* (observed mass) represents a 0.618 fraction of *outMASSIN*, whereas *outQSIM* (modelled mass) represents a 0.575 fraction of *outMASSIN*.



```
# Workspace is cleared
rm(list = ls())

# CRAN libraries are loaded
require(DescTools)
require(ggplot2)
require(grid)
require(gridExtra)
require(lubridate)
require(pastecs)
require(RColorBrewer)
require(reshape)
require(visreg)

# Working directory is defined
setwd("D:/R_ITC/HBV_TEC/HBV_TEC_CARTAGO/PURIRES")

# ///////////////////////////////////////////////////
# BLOCK: Creating and organizing input data frames
```

```
> write.csv(df.out02, file = "hbvqtecsim.csv")

> # df.out03 data.frame is exported
> write.csv(df.out03, file = "hbvqteceff.csv")

> # END OF PROGRAM
```

Fig. 3. Running the HBV-TEC hydrological model under *RStudio*-IDE.

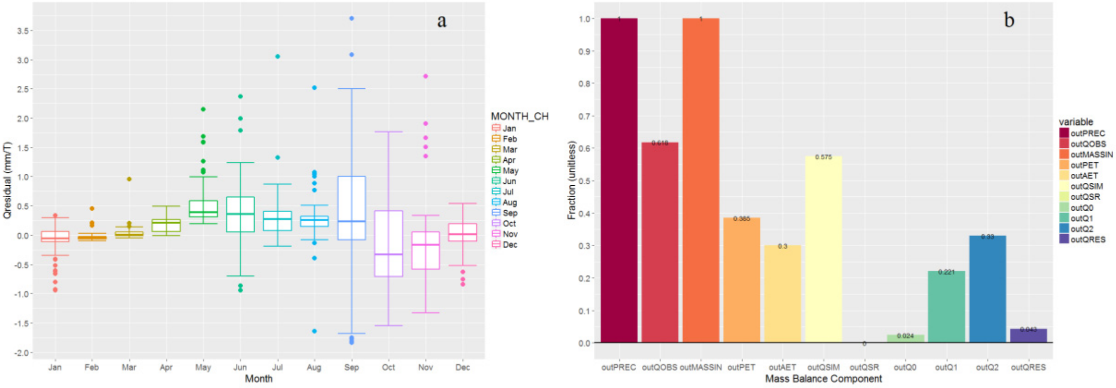


Fig. 4. (a) IDW input precipitation boxplot (by month) and (b) normalized mass-balance summary for the Purires subcatchment.

Furthermore, a modelled vs. observed runoff-plot for the Purires subcatchment (Fig. 5(a)), along with a daily input precipitation plot (Fig. 5(b)), display the observed values based on the proportion of their corresponding residuals; the more intense red and bigger dots, represent observations with a higher modelled residual.

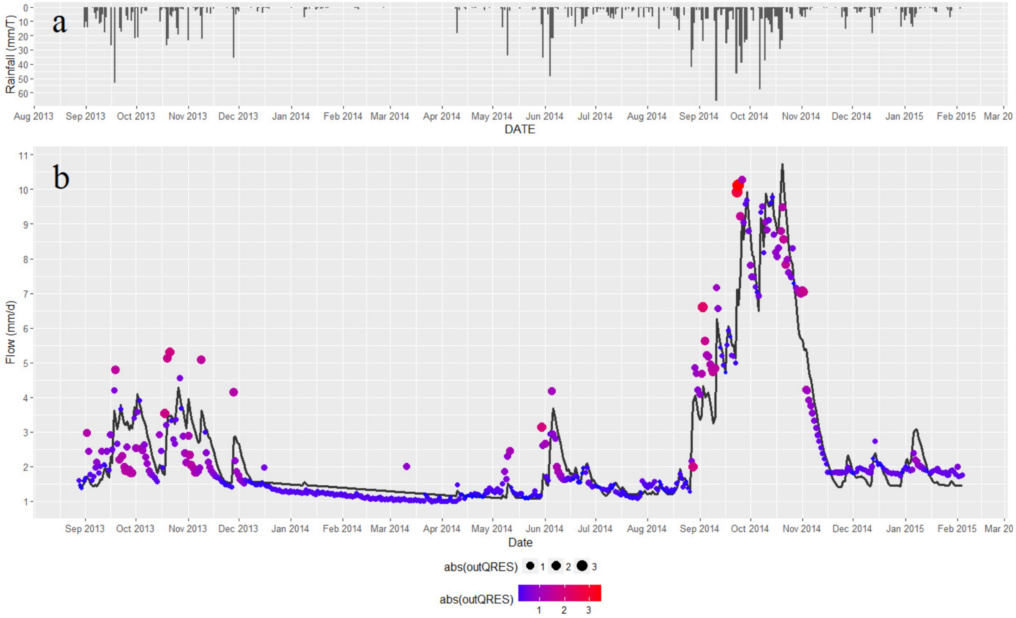


Fig. 5. (a) Daily input precipitation and (b) modelled and observed runoff for the Purires subcatchment.

Finally, a grid plot (Fig. 6) displays the HBV-TEC PEST optimal parameters, along with their 95% confidence intervals (CI) for all subcatchments. As observed, confidence intervals for various model parameters (e.g. *fc*, *k0* and *uzl*) clearly overlap, which could suggest parameter interdependence or correlation.

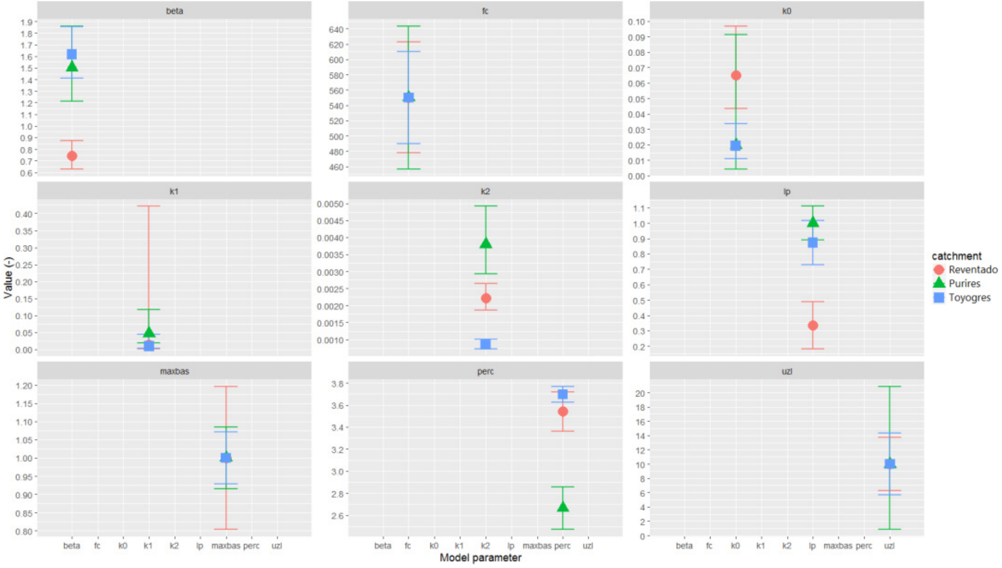


Fig. 6. HBV-TEC PEST optimal parameters plus 95% CI for the Reventado, Purires and Toyogres subcatchments.

4. Conclusions and recommendations

This study presented the development and assessment of the HBV-TEC hydrological model. The following conclusions can be drawn:

1. The combination of a HBV code and the R programming language forms a flexible and robust framework that enables the user to incorporate the data analysis and high-level visualization capabilities available in R into the hydrological model routines.
2. The HBV-TEC is an open-source code, which enable the user to have full access and control over the water balance solution included in the model.
3. The HBV-TEC proved to be capable of accounting for a satisfactory water balance even on the grounds of scarce input data.
4. Assessment of the HBV-TEC model over three subcatchments of Aguacaliente river produced considerably dissimilar results in terms of model performance and efficiency; which seems to be mostly related to the quality, quantity and extension of input data rather than to the model structure itself.

Future development of the HBV-TEC model will focus on: (a) further validation over additional catchments aiming to verify the efficiency of the code and identify potential improvements, (b) incorporation of selected regionalization methods for simulation in ungauged basins, (c) addition of advanced parameter-optimization and sensitivity analysis methods besides PEST, (d) incorporation of direct relevant geoprocessing operations based on dedicated R spatial-packages and (e) development of a fully distributed version of the code to be compared against present lumped and semi-distributed approaches.

Acknowledgements

This research was supported by Vicerrectoría de Investigación & Extensión, Instituto Tecnológico de Costa Rica (TEC). The authors are grateful to all the research assistants who participated in fieldwork and the administrative personnel who supported the logistics of this project.

References

[1] H. Moradkhani, S. Sorooshian, General review of rainfall-runoff modeling: model calibration, data assimilation, and uncertainty analysis, Hydrol Model Water Cycle. 63 (2008) 1–24.

[2] T. Cornelissen, B. Diekkrüger, S. Giertz, A comparison of hydrological models for assessing the impact of land use and climate change on discharge in a tropical catchment, J Hydrol Eng 498(2013) 221–236.

[3] D. Duan, S. Sorooshian, V.K. Gupta, Effective and efficient global optimization for conceptual rainfall runoff models, Water Resour Res. 24 (1992) 1163–1173.

[4] J. Seibert, M.J.P. Vis, Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, Hydrol Earth Syst Sci 16 (2012) 3315–3325.

[5] A. Aghakouchak, E. Habib, Application of a conceptual hydrologic model in teaching hydrologic processes, Int. J. Engng Ed 26(2010) 963–973.

[6] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2015. URL http://www.R-project.org/

[7] S. Bergström, The HBV model. In: Singh VP (ed) Computer Models of Watershed Hydrology, Water Resources Publications, Highlands Ranch, Colorado, USA, 1995, pp. 443–476.

[8] G. Lindström, B. Johansson, M. Persson, M. Gardelin, S. Bergström, Development and test of the distributed HBV-96 hydrological model, J. Hydrol. 201(1997), 272–288.

[9] J. Doherty, PEST: model independent parameter estimation: User's Manual, Watermark Numerical Computing, Brisbane, Australia, 2005.

[10] M. Mendez, J.A. Araya, L.D. Sanchez, Automated parameter optimization of a water distribution system, J HYDROINFORM. 15 (2013) 71–85.

[11] J.A. Retana, R. Villalobos, Eventos extremos meteorológicos en Linda Vista valle del Guarco, Cartago. Instituto Meteorologico Nacional, San Jose, Costa Rica, 2006.

[12] A. Efstratiadis, D. Koutsoyiannis, One decade of multi-objective calibration approaches in hydrological modelling: a review, Hydrol Sci J 55(2010) 58–78.

[13] RStudio Team, RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL, 2015. URL http://www.rstudio.com/.