The Hyperbolic Smoothing Approach for Automatic Calibration of Rainfall-Runoff Models

Adilson Elias Xavier, Otto Corrêa Rotunno Filho, Paulo Canedo de Magalhães

Abstract—This paper addresses the issue of automatic parameter estimation in conceptual rainfall-runoff (CRR) models. Due to threshold structures commonly occurring in CRR models, the associated mathematical optimization problems have the significant characteristic of being strongly non-differentiable. In order to face this enormous task, the resolution method proposed adopts a smoothing strategy using a special C∞ differentiable class function. The final estimation solution is obtained by solving a sequence of differentiable subproblems which gradually approach the original conceptual problem. The use of this technique, called Hyperbolic Smoothing Method (HSM), makes possible the application of the most powerful minimization algorithms, and also allows for the main difficulties presented by the original CRR problem to be overcome. A set of computational experiments is presented for the purpose of illustrating both the reliability and the efficiency of the proposed approach.

Keywords—Rainfall-runoff models, optimization procedure, automatic parameter calibration, hyperbolic smoothing method.

I. INTRODUCTION

Rainfall-runoff modeling continues to be a challenge for hydrologists. An unsolved problem is the definition of the parameters associated with the physical processes. The nature of the system operation is inferred from the input and output observations, achieved through a model-fitting process. Therefore, application of such models requires the identification of proper values for the parameters which govern the functions that describe the underlying physical system. Calibration is the stage in the simulation process where the parameters should be identified and estimated. Although the literature has devoted much attention to this stage, there is still a clear need for improving the analysis on parameter optimization and on detection of model structural identifiability problems in CRR models.

This paper focuses on the automatic parameter calibration technique, which uses a computerized mathematical optimization method to adjust the values of the unknown parameters based on changes in the values of a pre-specified objective function. These optimization algorithms search the parameter space for the extremum of an estimation criterion which measures the agreement between observed and simulated flows. The computational method introduces higher speed and less subjectivity into the calibration process.

Much research has been devoted to the development and improvement of optimization methods applied to rainfall-runoff models: [1]-[16], among others.

Despite all the efforts made, the process of automatic calibration still presents serious deficiencies. The major contribution of this work is to develop a general deterministic approach to solve the mathematical problem of calibrating parameters in rainfall-runoff models. It should be noted that the approach of the proposed hyperbolic smoothing is also applicable to other kind of hydrologic models as well as to other practical applications.

The present paper adopts a scheme, called HSM, applied with broad success for solving large nondifferentiable problems in general, such as for the min-max problem [17], for the covering of plane domains by circles [18], for covering of solid bodies by spheres [19], for the minimum sum-of-squares clustering illustrated in [20] and [21] and for multisource Fermat-Weber location [22].

Given the previous established framework, this paper is presented in the following sequence. First, it addresses the general difficulties observed in the optimization of CRR models, as previously pointed out by [5] and [13]. In Section III, a typical rainfall-runoff model is presented with a threshold-type component that produces discontinuities in the first derivatives. In Section IV, the smoothing technique introduced bypasses these difficulties. It must be emphasized that the overall hyperbolic smoothing approach preserves completely the conceptual model structure. A practical application is shown in Section V by using the SMAP model [23], which is a lumped rainfall-runoff model widely used in Brazil in the electric sector. Water resources play an essential role in the energy matrix of Brazil, which ranks among the top five countries in the world in terms of abundance of this kind of resource and where hydroelectric power corresponds presently to more than 60% [29]. In addition, the SMAP model, due to its simplicity, fits well in the framework of showing the applicability of the smoothing technique to the calibration of parameters in rainfall-runoff models. Subsequently, the resolution of the calibration problem and illustrative computational results are presented in Section VI and Section VII. Next, the hyperbolic smoothing approach performance is analyzed and the main conclusions are drawn.

II. CHARACTERISTICS OF CRR MODELS

CRR models are composed of reservoirs with limiting thresholds that are activated when the internal levels reach these...
limits. It is here that the derivative discontinuities originate [11]. Reference [13] presented five major characteristics of CRR models that generate difficulties in the calibration stage: regions of attraction, minor local optima, roughness, sensitivity and shape. All features were already considered under the framework described by [5]. However, it should be pointed out that regions of attraction and minor local optima are the most important features and they are related to the local minima challenge. Basically, [13] explained that the structure of multiple optima exists on at least two scales. At the large scale, there are broad regions of attraction into which the sequence of points generated by an optimization algorithm may converge. At the small scale, each major region of attraction contains a myriad of minor local optima.

III. ORIGIN OF DISCONTINUITIES

As quoted in [12] and [15], the central difficulty in the calibration of CRR models lies precisely in the discontinuities of the derivatives. The origin of these discontinuities must be analyzed. Reference [11] defined the possible modes of operation of a typical rainfall-runoff model and developed their arguments based on the threshold structures commonly employed in this kind of models. This typical structure of threshold values that appears in rainfall-runoff models leads to the multiplicity of possible ways of model operation, represented in the program code by IF structures. These different modes of operation cause the discontinuities in the derivatives of the model functions that represent a theoretical restriction to the applicability of first order and second order derivative-based techniques.

IV. SMOOTHING TECHNIQUE USED (HSM)

As previously highlighted, the response surface of CRR models contains derivative discontinuities which correspond to the presence of thresholds in the corresponding reservoirs. In order to overcome this difficulty, a natural idea is to use a smoothing approach [24]. Reference [25] suggested two possible alternative approximations to a threshold behavior in a real-time estimation and forecasting model of river flows. However, such suggestions were not used in calibration procedures of rainfall-estimation and forecasting model of river flows. As [12] states, the proposal made by [25] of replacing discontinuities with smooth S-shaped jumps would introduce perturbations in the derivatives and not completely solve the problem of non-smoothness. On the other hand, the present paper examines a continuously differentiable function that properly approximates the function \( R_t(x_t - M) \) of the rainfall-runoff model. The following function was adopted:

\[
\phi(x_t, M, d) = \frac{1}{2} \left[ x_t - M + \left( (x_t - M)^2 + 4d^2 \right)^{1/2} \right]
\]  

(1)

The function \( \phi(x_t, M, d) \) presents the following properties:

a) \( \phi(x_t, M, d) \) is asymptotically tangent to the straight lines \( \tau_1 \) \( (x_t, M) = 0 \) and \( \tau_2 \) \( (x_t, M) = x_t - M \);  
b) \( \lim_{d \to 0} \phi(x_t, M, d) = 0 \), if \( x_t \leq M \); \( \lim_{d \to 0} \phi(x_t, M, d) = x_t - M \), if \( x_t > M \);  
c) \( \phi(x_t, M, d) \) is continuous and continuously differentiable in the variables \( x_t \) and \( M \);  
d) \( \phi(x_t, M, d) \) is convex in \( x_t \) and \( M \) (increasing in \( x_t \) and decreasing in \( M \)).

Property (a) indicates function \( \phi(x_t, M, d) \) as a good smoothing for \( R_t(x_t, M) \). Property (b) shows that the difference between \( \phi(x_t, M, d) \) and \( R_t(x_t, M) \) may be made as little as desired. The variable \( d \) introduced in the model represents the maximum deviation between the functions \( \phi(x_t, M, d) \) and \( R_t(x_t, M) \). Taking that into consideration, full control can be imposed over its value, in such a way that the difference between the smoothed model and the original model can be established at any desired level. Therefore, it is possible to guarantee that the model’s overall structure remains intact for hydrological purposes. Property (c), which refers to first and second differentiability, will allow for the use of the optimization algorithms that are known to be most powerful. Property (d) implies the preservation of fortuitous convex structures.

V. DESCRIPTION OF AN APPLICATION

The smoothing technique specified (HSM) was implemented in the SMAP (Soil Moisture Accounting Procedure) model [23], which is a deterministic CRR model with a structure analogous to other models found in the literature. The model simulates the land phase of the hydrological cycle by means of three linear reservoirs that represent respectively, the surface runoff, the soil surface zone and the groundwater flow.

The flow calculated by the model at each instant is the result of the sum of the parts contributed by the surface and underground reservoirs. The part corresponding to the rain that infiltrates is determined according to an equation of the US Soil Conservation Service [26]. The variables are constantly updated based on precipitation (RAIN) and potential evapotranspiration (EVPT) data and considering the mass conservation principle. The objective function (FO) considers the minimization of the sum of squares of differences between simulated (QGER) and observed streamflows (QOBS).

The model, in its more simplified version, has six parameters: initial losses caused by vegetable retention and soil depressions (ABSII); surface recession coefficient (KSUP); soil saturation level (NSAT); field capacity (NPER); top soil reservoir recession coefficient (KPER); groundwater recession coefficient (KSUB). The SMAP model’s equations are given below:

\[
PEFE_t = \max(0, RAIN_t - ABSI)  
\]

\[
QRES_t = \frac{PEFE_t^2}{(PEFE_t + NSAT - LSOI_{t-1})}  
\]

\[
QSOIL_t = RAIN_t - QRES_t  
\]

\[
QINF_t = \max(0, QSOIL_t - EVPT_t)  
\]

\[
PEV1_t = QSOIL_t - QINF_t  
\]
For the validation of the proposed methodology, it was decided to use series of synthetic flows generated by the model itself from daily rain and potential evapotranspiration series for the Fartura river watershed, Sao Paulo state, Brazil. In this way, it was possible to isolate, from the analysis, the problems of the natural errors of input data and the natural imperfections of the hydrological model; that is, any divergencies between the series of flows generated and observed are exclusively due to the conceptual model structure.

The objective function chosen was the minimization of the sum of the squares of the differences between the observed flows and those generated by the model. The record length for the input was 5 years and the time step 1 day, making a total of 1826 observations, which means a problem with 31042 variables.

The five derivative discontinuities existing in the SMAP process of calibration. Moreover, exact analytical first derivatives of the objective function with respect to parameters can be trivially calculated through elementary calculus rules. Calibration thus consisted of resolving a completely differentiable nonlinear programming problem. Nevertheless, it is desirable to include constraints in the space of parameters to place them always within the domains of physical meaning, and also to avoid the occurrence of unpleasant divergent numerical sequences. Thus, the recession coefficients (KSUP, KPER, KSUB) were put into the interval [0,1] and for the other parameters lower bounds were set to zero and specific safe upper bounds were defined. It was therefore a question of solving a constrained nonlinear programming problem.

The solution to the constrained optimization problem was performed by using the hyperbolic penalty method [27], which resorts to the solution of a sequence of unconstrained penalized subproblems. Each subproblem was solved by using the routine VA13C from Harwell Library, which implements a Quasi-Newton method with BFGS updating, described in many sources, such as in [28]. Similar to other penalty methods, the sequence of penalized subproblems is generated by the sequence of smoothed CRR models and therefore a single C∞ differentiable path replaces the operational paths derived from the sequence of the five IF’s. Moreover, for the objective function with respect to parameters can be trivially calculated through elementary calculus rules. Calibration thus consisted of resolving a completely differentiable nonlinear programming problem. Nevertheless, it is desirable to include constraints in the space of parameters to place them always within the domains of physical meaning, and also to avoid the occurrence of unpleasant divergent numerical sequences. Thus, the recession coefficients (KSUP, KPER, KSUB) were put into the interval [0,1] and for the other parameters lower bounds were set to zero and specific safe upper bounds were defined. It was therefore a question of solving a constrained nonlinear programming problem.

The solution to the constrained optimization problem was performed by using the hyperbolic penalty method [27], which resorts to the solution of a sequence of unconstrained penalized subproblems. Each subproblem was solved by using the routine VA13C from Harwell Library, which implements a Quasi-Newton method with BFGS updating, described in many sources, such as in [28]. Similar to other penalty methods, the sequence of penalized subproblems is generated by the controlled variation of an external penalty parameter in a convergent sequence to zero. In the above description of the hyperbolic smoothing technique, it was shown that this also is a function of an external smoothing parameter, d, which provokes a concrete alteration of the original CRR model. This difficulty is definitively overcome through linear coupling of the smoothing function to the penalty parameter that goes to zero. In this way, the sequence of values of the smoothing parameter will also converge to zero and the sequence of smoothed CRR models will converge to the original model. Conceptual model structure is maintained throughout.

VII. RESULTS FOR PARAMETERS CALIBRATION

In order to demonstrate the performance of the proposed hyperbolic smoothing technique, it is presented the computational results for the SMAP model with six-parameter...
calibration. The calibration sample was established based on a known solution with the parameters assuming the following values: \( \text{ABSI} = 5.0; \text{KSUP} = 0.70; \text{NSAT} = 300.0; \text{NPER} = 90.0; \text{KPER} = 0.008 \) and \( \text{KSUB} = 0.95 \). Six calibration runs were performed, where the complete parameter set was perturbed 10%, 20%, 30%, 40%, 50% and 75%. A brief synthesis of the results is given below.

First, in Table I, a typical sequence of points generated by the HSM method in solving the 10% deviation is shown. The corresponding columns respectively, present; the name of the parameter, the initial point, the known solution, the initial and final points obtained by the variation of the smoothing parameter \( d \), from an initial value 1000 until 0.001. The last but one row presents the objective function values, while the last row depicts the small number of function and gradient evaluations.

It was identified a linear convergence of the sequence of points generated by the algorithm to the true solution with the controlled linear decreasing of parameter \( d \) (1000, 100, 10, 1, 0.1, 0.01, 0.001).

### Table I

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial Value (10%)</th>
<th>True Value</th>
<th>( d = 1000 )</th>
<th>( d = 0.001 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSI</td>
<td>4.50000</td>
<td>5.00000</td>
<td>4.73176</td>
<td>5.00000</td>
</tr>
<tr>
<td>KSUP</td>
<td>0.63000</td>
<td>0.70000</td>
<td>0.68397</td>
<td>0.70000</td>
</tr>
<tr>
<td>NSAT</td>
<td>270.000</td>
<td>300.000</td>
<td>311.227</td>
<td>300.000</td>
</tr>
<tr>
<td>NPER</td>
<td>81.0000</td>
<td>90.0000</td>
<td>97.0440</td>
<td>90.0000</td>
</tr>
<tr>
<td>KPER</td>
<td>0.00720</td>
<td>0.00800</td>
<td>0.00742</td>
<td>0.00800</td>
</tr>
<tr>
<td>KSUB</td>
<td>0.85500</td>
<td>0.95000</td>
<td>0.95460</td>
<td>0.95000</td>
</tr>
<tr>
<td>Objective</td>
<td>0.29E4</td>
<td>0.00000</td>
<td>0.49E2</td>
<td>0.15E-7</td>
</tr>
<tr>
<td>Function/Gradient</td>
<td></td>
<td></td>
<td>135</td>
<td>51</td>
</tr>
</tbody>
</table>

Table II presents the solutions produced by HSM after 7 similar iterations where the parameter set was perturbed 10%, 30%, and 50%. For each level of perturbation, it is only shown the last solution produced with smoothing parameter \( d \) equal to 0.001.

### Table II

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Value</th>
<th>Solution (10%)</th>
<th>Solution (30%)</th>
<th>Solution (50%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSI</td>
<td>5.00000</td>
<td>5.00000</td>
<td>5.00000</td>
<td>5.00000</td>
</tr>
<tr>
<td>KSUP</td>
<td>0.70000</td>
<td>0.70000</td>
<td>0.70000</td>
<td>0.70000</td>
</tr>
<tr>
<td>NSAT</td>
<td>300.000</td>
<td>300.000</td>
<td>300.000</td>
<td>300.000</td>
</tr>
<tr>
<td>NPER</td>
<td>90.0000</td>
<td>90.0000</td>
<td>90.0000</td>
<td>90.0000</td>
</tr>
<tr>
<td>KPER</td>
<td>0.00800</td>
<td>0.00800</td>
<td>0.00800</td>
<td>0.00800</td>
</tr>
<tr>
<td>KSUB</td>
<td>0.95000</td>
<td>0.95000</td>
<td>0.95000</td>
<td>0.95000</td>
</tr>
<tr>
<td>Objective</td>
<td>0.00000</td>
<td>0.15E-7</td>
<td>0.15E-7</td>
<td>0.15E-7</td>
</tr>
</tbody>
</table>

Based on the analysis of the results presented in Tables I and II, first, it is possible to verify the almost perfect performance of the HSM. The convergence rate of the six parameters and of the objective function presents a linear behavior with the smoothing parameter \( d \).

### VIII. THE SMOOTHING PROCEDURE

Many authors [5]-[15] have reported that due to the effect of the discontinuities of the derivatives, the high degree of non-linearity, the presence of extensive valleys with low declivity resulting from the interaction between the parameters, and the existence of local minima, automatic calibration of CRR models becomes an extremely difficult task. The alternative considered in this paper was to resort to smoothing, in order to improve the structural characteristics of the problem and allow the use of more robust optimization algorithms and, by including a set of constraints, maintain the physical meaning of the parameters. The almost perfect convergence for the solution, whatever the initial point, in contrast to the experience registered in the literature, suggests the appropriateness of the smoothing approach and of the other procedures implemented.

The reason for the success seems to result from the total or partial elimination of the difficulties of automatic calibration of CRR models referred to in Section II, following the remarks made by [5] and [13].

### IX. CONCLUSION

The authors have wishful expectation that the techniques presented herein may just as successfully be used in other CRR with different structural functions, as well as in other conceptual models in different environmental areas such as groundwater modeling and atmospheric sciences. However, it must be remembered that the SMAP calibration problem is a global optimization problem with several local minima. Therefore, HSM does not offer the guarantee of obtaining a global optimum point, although it has been possible to observe that the present computational implementation produced such global points.

Summing up, the option taken in this paper was not to attempt to discover an appropriate algorithm to perform calibration of CRR models, but rather to transform the model by preserving its conceptual structure so as to provide the necessary conditions to enable the existing robust algorithms to work properly.

### ACKNOWLEDGMENT

The authors would like to express their gratitude towards the Federal University of Rio de Janeiro (UFRJ) for the support. The authors are thankful and recognize that this study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES)- Finance Code 001. The authors are also grateful to the Brazilian Ministry of Science, Technology, Innovation and Communication (MCTIC), through the National Council for Scientific and Technological Development (CNPq) and the Financier of Studies and Projects (FINEP), and to the Fundação Carlos Chagas de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) for the means and support of the development of this research work.

### REFERENCES

[4] Ibbitt, R. P., and T. O'Donnell, Fitting methods for conceptual catchment,


