FLOW ROUTING WITH UNKNOWN RATING CURVES

József Szilágyi[1], Gábor Bálint[2], Balázs Gauzer[2], Péter Bartha[2]

[1]*Water Resources Research Centre, VITUKI, Hungary, on leave from the Conservation and Survey Division, University of Nebraska-Lincoln, Lincoln, USA, e-mail: jszilagyi1@unl.edu*
[2] *Water Resources Research Centre, VITUKI, Budapest, Hungary, e-mail: balint@vituki.hu*

Abstract: A discrete version of the Kalinin-Milyukov-Nash-cascade is formulated for operational forecasting of stream stages when no information of rating curves is available. Model performance is slightly reduced in comparison to flow routing results using accurate, single-valued stage-discharge relationships. However, when only inaccurate rating curves are available, the present approach may yield superior forecasts. Since in practice the accuracy of the employed rating curves, used to convert stage measurements into discharge values for flow routing, may somewhat be uncertain, application of the present technique is recommended for rating-curve verification. The method allows for stage predictions using physically-based flow routing in rivers where flow rates are unknown or the available rating curves are inaccurate. The technique can also be used without modification for streams with tributaries.
Keywords: stream-flow, flow routing, forecasting, hydrologic model, rating curve.

FLOW ROUTING MIT UNBEKANNTEN SCHLÜSSELKURVEN

Zusammenfassung: Es wurde eine diskrete Version der Kalinin-Milyukow-Nash-Kaskade für die operative Wassetrstandsvorhersage in solchen Fällen entwickelt, in welchen keine Information über die Schlüsselkurven zur Verfügung steht. Die Wirksamkeit dieses Vorhersagemodells ist in bißchen geringer, als in dem Falle, wenn eine genaue und eindeutige Schlüsselkurve verwendet werden kann, doch ist sie wesentlich höher, als wenn nur ungenaue Schlüsselkurven zur Verfügung stehen. Da es jedoch in der Praxis oft vorkommt, daß bei der Anwendung der Methode *flow routing* die zur Konversion zwischen Wasserstand und Abfluß herangezogenen Schlüsselkurven mehr oder weniger ungenau sind, kann die Anwendung der hier vorgeführten Methode auch zur Verifizierung der Schlüsselkurven empfohlen werden. Die Methode ermöglicht nämlich eine Wasserstands-Vorhersage über ein physisch begründetes *flow routing,* falls die Abflußwerte unbekannt oder die zur Verfügung stehenden Schlüsselkurven ungenau sind. Das vorgeschlagene Verfahren kann, ohne irgendeine Abänderung, auch für Flußstrecken mit Nebenflüssen verwendet werden.
Schlüsselworte: flow routing, Vorhersage, hydrologische Model, Schlüsselkurve.

1. Model description

The linear storage equation results if one assumes that the exponent (α) is the same in the functional relationships between outflow rate (*Q*) and stage as well as between water stored (*S*) in a channel reach and stage

$$Q(t) = c_1[H(t) + a]^\alpha$$
$$S(t) = c_2[H(t) + a]^\alpha \qquad (1)$$

where *H[L]* is the measured value of stage above or below datum, and $c_1[L^{3-\alpha}T^{-1}]$, $c_2[L^{3-\alpha}]$ and α *[L]*are constants. Dividing Eq. (1a) by (1b) yields

$$Q(t) = \frac{c_1}{c_2} S(t) = kS(t) \tag{2}$$

Inserting Eqs. (1a), (1b), and (2) into the lumped continuity equation

$$\dot{S}(t) = -Q(t) + u(t) = -kS(t) + u(t) \tag{3}$$

where u is inflow rate, results in

$$c_2 \alpha [H_2(t) + a]^{\alpha-1} \frac{dH_2(t)}{dt} = -\frac{c_1}{c_2} c_2 [H_2(t) + a]^{\alpha} + c_3 [H_1(t) + b]^{\beta} \tag{4}$$

where the subscripts 1 and 2 refer to the up- and downstream ends of the channel reach, and c_3 $[L^{3-\beta}T^{-1}]$, $b[L]$, and β are constants of the stage-discharge relationship of the upstream location. By rearranging Eq. (4) one obtains

$$\frac{dH_2(t)}{dt} = -\frac{c_1}{c_2 \alpha}[H_2(t) + a] + \frac{c_3}{c_2 \alpha} \frac{[H_1(t) + b]^{\beta}}{[H_2(t) + a]^{\alpha-1}} \tag{5}$$

which shows that in general the future outflow rate of the reach is determined by a certain combination of in- and outflow rates through the last term of the righ-hand-side of the equation. However, by assuming that both exponents are unity, Eq. (5) simplifies into

$$\frac{dH_2(t)}{dt} = -\frac{c_1}{c_2} H_2(t) + \frac{c_3}{c_2} H_1(t) + c_4 = -kH_2(t) + cH_1(t) + c_4 \tag{6a}$$

where $c=c_3/c_2[T^{-1}]$, and $c_4[LT^{-1}]$ are just another constants. The constant multiplier of H_1 and an additional constant value in Eq. (6a) are of no concern because linearity of the equation assures that the output of the reach is proportional to any constant multiplier in the input values, and the presence of a constant input means only an additional constant value in the output values. Because of the arbitrary reference points in the stage measurements of differing locations, routed upstream stage values have to be scaled up or down any way to match the measured downstream stage values, thus the presence of a constant multiplier (and an extra constant) in the input stage values means only an additional multiplication in the scaling process, consequently they can be chosen arbitrarily. This way Eq. (6a) can be written as

$$\frac{dH_2(t)}{dt} = -kH_2(t) + H_1(t) \tag{6b}$$

which now is of the same form as Eq. (3) of the Kalinin-Milyukov-Nash-cascade (Kalinin and Milyukov, 1957; Nash, 1957). The reason that the required scaling is not a linear function eventually stems from the general nonlinear shape of the rating curves while in the derivation of Eq. (6b) linear rating curves were employed. The required scaling of routed to observed stage values can be achieved by the application of a polynomial curve fitting in the form of

$$\widehat{H_2^{sc}}(t) = p_1 \widehat{H_2}^{m}(t) + p_2 \widehat{H_2}^{m-1}(t) + \ldots + p_m \widehat{H_2}(t) + p_{m+1} \tag{7}$$

where $\hat{H}_2^{SC}$ is the scaled, $\hat{H}_2$ is the original model estimate of the downstream stage value, and the $p_{i-s}[L^{i-m}]$ are the constant coefficients of the polynomial of a predefined order m.

With these considerations the solution of the KMN-cascade model can be applied. Szilágyi (2003) derived the solution for a sample-data system which implies that the stage measurements are available only at discrete time intervals (Δt) with an assumed linear change in the values between consecutive discrete samples. Applying the solution to Eq. (6b) over *n* serially connected subreaches one obtains

$$\mathbf{H}(t+\Delta t) = \mathbf{\Phi}(\Delta t)\mathbf{H}(t) + \Gamma_1(\Delta t)H_1(t+\Delta t) + \Gamma_2(\Delta t)H_1(t) \tag{8}$$

where the vector ***H*** comprises of the modeled stage values of the *n* subreaches, the $\Phi(\Delta t)$ state-transition matrix, and the $\Gamma_1(\Delta t)$ and $\Gamma_2(\Delta t)$ input-transition vectors are defined as (Szilágyi, 2003)

$$\mathbf{\Phi}(\Delta t) = \begin{bmatrix} e^{-\Delta tk} & 0 & 0 & \cdots & 0 \\ \Delta tk e^{-\Delta tk} & e^{-\Delta tk} & 0 & \cdots & 0 \\ \frac{(\Delta tk)^2}{2!}e^{-\Delta tk} & \Delta tk e^{-\Delta tk} & e^{-\Delta tk} & 0 & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \frac{(\Delta tk)^{n-1}}{(n-1)!}e^{-\Delta tk} & \frac{(\Delta tk)^{n-2}}{(n-2)!}e^{-\Delta tk} & \cdots & \Delta tk e^{-\Delta tk} & e^{-\Delta tk} \end{bmatrix} \tag{9}$$

$$\Gamma_1(\Delta t) = \begin{bmatrix} \frac{1}{k}\frac{\Gamma(1,\Delta tk)}{\Gamma(1)}\left[1 + \frac{e^{-\Delta tk}}{\Gamma(1,\Delta tk)} - \frac{1}{\Delta tk}\right] \\ \frac{1}{k}\frac{\Gamma(2,\Delta tk)}{\Gamma(2)}\left[1 + \frac{(\Delta tk)e^{-\Delta tk}}{\Gamma(2,\Delta tk)} - \frac{2}{\Delta tk}\right] \\ \vdots \\ \frac{1}{k}\frac{\Gamma(n,\Delta tk)}{\Gamma(n)}\left[1 + \frac{(\Delta tk)^{n-1}e^{-\Delta tk}}{\Gamma(n,\Delta tk)} - \frac{n}{\Delta tk}\right] \end{bmatrix} \tag{10}$$

and

$$\Gamma_2(\Delta t) = \begin{bmatrix} \frac{1}{k}\frac{\Gamma(1,\Delta tk)}{\Gamma(1)}\left[\frac{1}{\Delta tk} - \frac{e^{-\Delta tk}}{\Gamma(1,\Delta tk)}\right] \\ \frac{1}{k}\frac{\Gamma(2,\Delta tk)}{\Gamma(2)}\left[\frac{2}{\Delta tk} - \frac{(\Delta tk)e^{-\Delta tk}}{\Gamma(2,\Delta tk)}\right] \\ \vdots \\ \frac{1}{k}\frac{\Gamma(n,\Delta tk)}{\Gamma(n)}\left[\frac{n}{\Delta tk} - \frac{(\Delta tk)^{n-1}e^{-\Delta tk}}{\Gamma(n,\Delta tk)}\right] \end{bmatrix} \tag{11}$$

The output equation now becomes

$$\widehat{H_2}(t) = [0, 0, \ldots, 1]\begin{bmatrix} H_1(t) \\ \vdots \\ H_n(t) \end{bmatrix} \tag{12}$$

the term on the left-hand-side being the input to Eq. (7). For channel reaches with tributaries, stages are routed separately between up- and downstream stations on the main channel and the upstream station of each tributary and the downstream station of the main channel due to linearity of the KMN-cascade, before inserting the $\hat{H}_{2j}(t)$ $(j=1,\ldots,T+1)$, where T is the number of tributaries within the reach) values into Eq. (7). Then the p_i $(i=1,\ldots,m)$ coefficients of the polynomial become vector-valued.

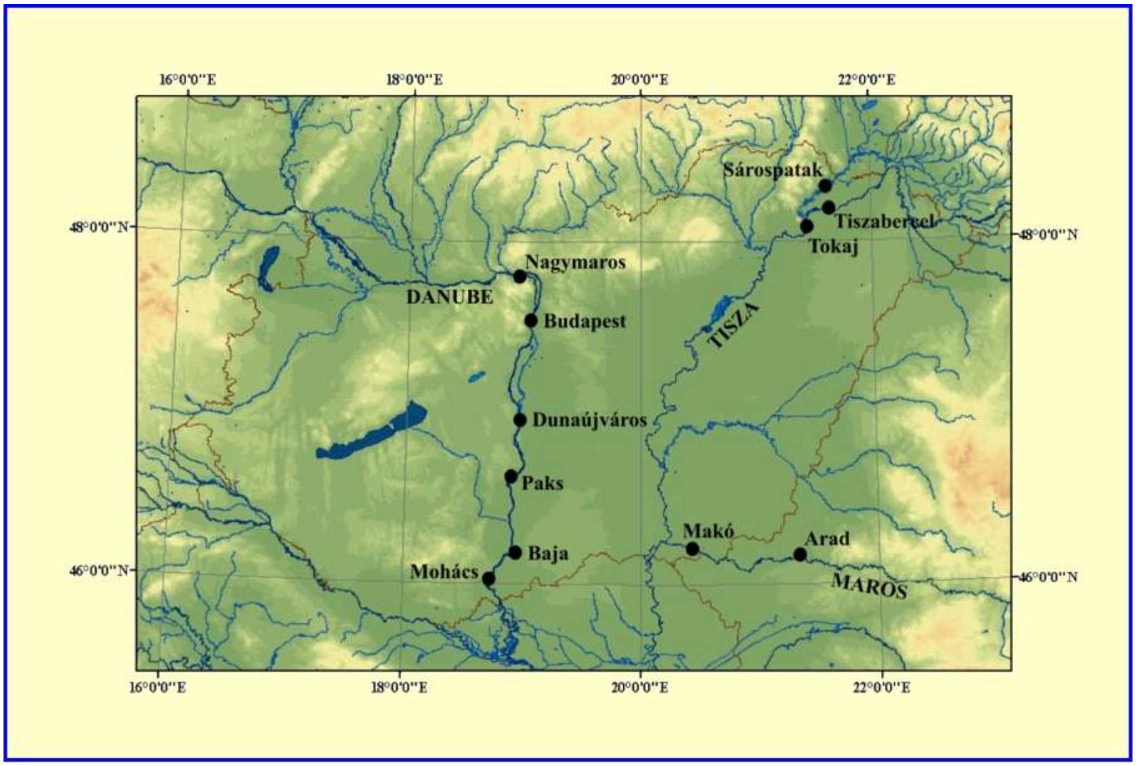

Figure 1. Map of Central Danube Basin

2. Model application and conclusions

The above model was tested on three rivers in Hungary: the Danube, its tributary, the Tisza River, and a tributary of the Tisza, the Maros River. See Table 1 for a list of gaging stations with corresponding drainage areas and mean channel slopes and see Figure 1 for a map of the mentioned area.

Table 1. Stream reaches used in the study with corresponding reach lengths L[km], average channel slopes I[%], as well as drainage areas D[km^2] belonging to the downstream stations.

	L (km)	D (km^2)	I (%)
Nagymaros - Budapest	48.1	184.893	0.0071
Budapest - Dunaújváros	65.9	188.273	0.0090

Budapest - Paks	115.2	189.092	0.0091
Paks – Baja	52.6	208.282	0.0065
Baja - Mohács	31.8	209.064	0.0058
Tiszabercel - Tokaj	25.9	49.849	0.0096
Sárospatak - Tokaj	37.1	13.000*	0.0114
Arad – Makó	72.7	30.149	0.0057

*Drainage area of the tributary (Bodrog) above the confluence

Model results were compared with that of an operative, real-time hydrological forecasting version of the KMN-cascade using actual rating-curve-derived discharge values. The operational model uses a time-step of Δt=12 hours and employs a multilinear approach (Becker and Kundzewicz, 1987; Szolgay, 1991) where discharges are routed through parallel cascades of linear storages representing low-, and mean-flow channel as well as flood conditions over the floodplain, thus creating a nonlinear model. The operative model has 3 x 2 = 6 (n and k values for each three cascades) parameters, plus a one-step autoregressive coefficient for prediction error updating while the proposed model has two, plus one autoregressive (*ar*), parameters and is run with a time-step of Δt=24 hours. To assure identical input values for model performance comparisons, the new model uses forecasted stage values of a lead-time of 24 hours, calculated by the operative model for the upstream stations. Both models were run in a continuous error-updating mode, which means that each forecast value is modified by a certain percentage (given by the value of the autoregressive parameter) of the previous day's model error prior to error updating.

Parameters of the proposed model were optimized with two years of data from the period January 1, 2000 - December 31, 2001. Model results, using the optimized parameter values, were compared with operative model outputs for the period January 1, 2002 - September 18, 2003. Model performance was assessed by two statistics: the mean root-square error (σ) and a Nash-Sutcliffe-type efficiency coefficient (*NSC*) which is defined as

$$NSC = 100\left(1 - \frac{\sum_i \left(\hat{H}_i - H_i\right)^2}{\sum_i \left(H_{i-1} - H_i\right)^2}\right)[\%]$$

(Interval číslování eq3)

where $\hat{H}_i$ is the predicted, and H_i the observed stage value on day i. The closer is the *NSC* value to 100% the better are the predictions. Note that the *NSC* value may be negative when the forecasts are worse than the naive prediction (see denominator), which takes the stage value of the actual day as the one-day forecast. Table 2 lists the optimized model parameter values.

Table 2. Optimized model parameter values for different stream reaches.

	K [1/day]	n	ar
Nagymaros - Budapest	11	4	0.2
Budapest - Dunaújváros	6.8	4	0.2
Budapest - Paks	3.9	4	0.6
Paks - Baja	3.2	2	0.8
Baja - Mohács	2.7	1	0.7
Tiszabercel -Tokaj	8.5	1	0.9

Sárospatak – Tokaj*	1.5	2	
Arad – Makó	14.5	17	1.0

* The tributary (Bodrog) of the Tisza

Optimization of the *n*, *k* and *ar* values of the proposed model was carried out by a systematic trial-and-error search where trial values of the parameters were chosen from ever-decreasing predefined ranges of the parameters with ever-increasing corresponding resolution terminating at a chosen set resolution. Parameters of the nonlinear regression equation (Eq. [7]) were obtained using the Matlab function "Nlinfit" for the multivariate case, and the function "Polyfit" for the univariate case, both by prescribing a 3-rd order polynomial.

Application of assumed linear rating curves instead of more realistic measured ones causes the curvature of the best-fit polynomial at large values. Such a systematic error, however, can be easily corrected via Eq. (7). Note that the first few forecast values may be off mark, since modeling starts with an arbitrary zero initial value of the **H** vector. Consequently, the first four forecast values were left out from all subsequent analysis. Table 3 lists the performance statistics of the one-day model predictions for the two distinct periods.

Based on Table 3 it can be stated that the proposed model has stable optimized parameter values since model performance deteriorates only slightly between the two periods. During the verification period there happened to be a major, but a relatively short-term (several days) water release through a dam of the Tisza downstream of Tokaj which contributed to a large drop in model efficiencies between the periods.

In general, physically-based models are expected to have more stable parameters in time than so-called black-box models (Szöllősi-Nagy, 1989) and also more accurate forecasts with increasing lead-times (Szöllősi-Nagy, 1989; Szilágyi, 1992). Because the former give some insight (may that be very simplified) into the physical processes involved, temporal changes in parameter values can often be linked to changes in the channel or floodplain conditions, such as conveyance. Also, model transferability of physically-based models between gaged and ungaged basins is typically better than that of black-box models (Nash and Sutcliffe, 1970) simply because model parameter values can be linked to measurable basin properties. In our case, the ratio of optimized values of *n* and *k* yields the mean travel time of flow propagation for the given reach. Since this latter is a function of channel properties mainly, initial guesses of the *n* and *k* values for a new, ungaged stream can be obtained by using such information only.

Overall, model performance of the proposed model is very similar to that of the operative model. For certain stations (Budapest, Baja, and Makó) the operative model produces more accurate predictions than the proposed model. This is what would normally be expected, since the operative model uses extra information (i.e. known rating curves) for flow routing.

Table 3. Performance statistics of the one-day ahead stage predictions.

Model performance statistics (standard deviation, Nash and Sutcliffe criterion) of the one-day ahead stage forecasts

Legend
Better
Worse

Optimization period

	σ (cm)		NSC (%)	
Station / Model	Proposed	Operative	Proposed	Operative
Budapest	5.95	5.67	94.21	94.75
Dunaújváros	6.58	8.42	92.15	87.14
Paks	5.08	7.46	92.67	91.96
Baja	6.92	5.68	91.75	94.43
Mohács	5.28	5.49	94.34	93.9
Tokaj	6.23	8.53	78.87	60.34
Makó	12.02	11.85	66.79	67.72

Verification period

	σ (cm)		NSC (%)	
Station / Model	Proposed	Operative	Proposed	Operative
Budapest	8.11	7.83	91.66	92.23
Dunaújváros	8.59	9.88	89.13	85.75
Paks	6.07	9.46	95.7	89.55
Baja	7.69	7.87	91.81	91.45
Mohács	6.16	6.72	93.79	92.61
Tokaj	9.72	17.57	44.76	0
Makó	9.36	10.85	64.01	51.49

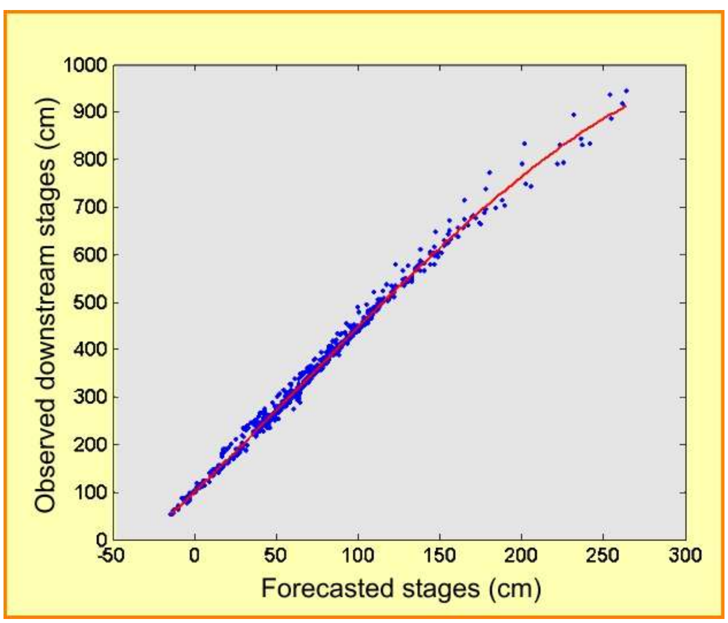

Figure 2a. Observed stages at Baja of the verification period versus unscaled one-day forecasts

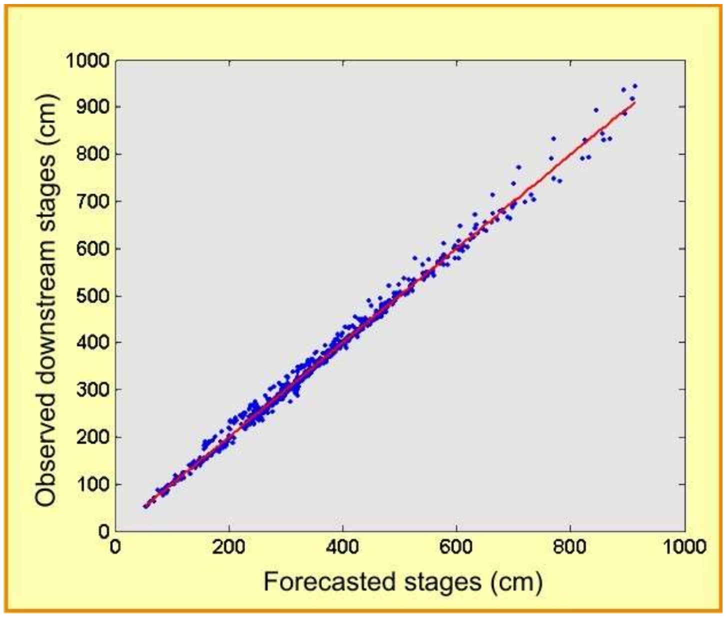

Figure 2b. Observed stages at Baja of the verification period versus scaled one-day forecasts without error updating

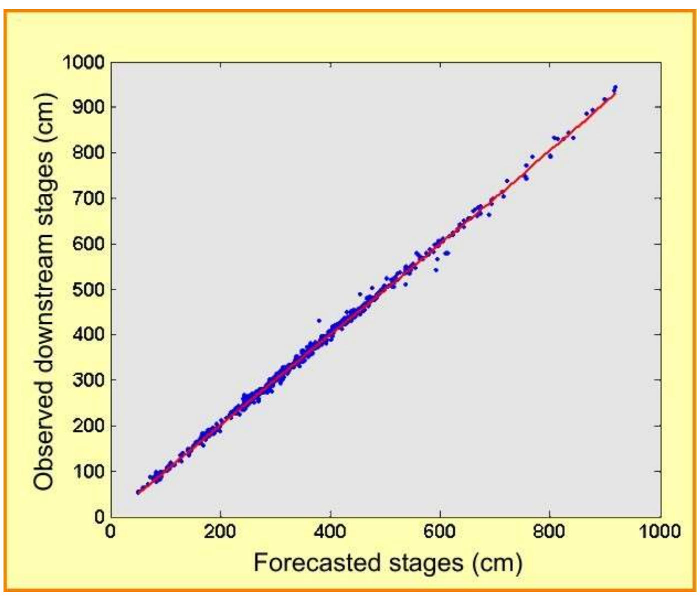

Figure 2c. Observed stages at Baja of the verification period versus scaled one-day forecasts with error updating

Scaling is based on historical data and corrects for the systematic difference between assumed and real (but unknown) rating curves.

One plausible explanation of why the proposed model may perform better than the operative one for other stations (Dunaújváros, Paks, and Tokaj) can be that for those stations the rating curves may not be accurate enough or they may be outdated, i.e. they do not reflect correctly the channel and flow conditions of the modeled periods. Suboptimal parameter values (which could stem from a higher number of parameters to be optimized, i.e. 7 as opposed to 3) in case of the operative model might also explain its underperformance, but it is unlikely knowing that parameter values of the operative model are updated each day using information from the previous 90 days (Szilágyi, 1992). Here it should be emphasized that the proposed model is not meant for replacing models that use measured rating-curve information. Whenever reliable rating curves are available, a flow-rate formulation, i.e. Eq. (3), should always be preferred over a stage formulation, Eq. (6b). However, an additional (on top of flow rates) flow routing using stages only, can detect inadequacies in the data required by the former. Naturally, when no information of rating curves is available, the proposed model (or its variant, such as the multilinear formulation) may easily be a proper candidate of a physically-based model to apply. Szilágyi (2004) provides an exhaustive list of the advantages of applying a state-space approach of flow routing over a numerical solution of the linear kinematic or diffusion wave equations beyond the already-mentioned properties that flow routing is a lumped parameter approach while the kinematic and diffusion wave equations are distributed ones.

3. References

Becker, A., Kundzewicz, Z. W. (1987): *Nonlinear flood routing with multilinear models*. **Water Resour. Res.**, 23(6), 1,043-1,048.

Kalinin, G. P., Milyukov, P. I. (1957): *O raschete neustanovivshegosya dvizheniya vody v otkrytykh ruslakh (On the computation of unsteady flow in open channels)*. **Met. i Gydrologia Zhurnal**, 10, 10-18 ,Leningrad

Nash, J. E. (1957): *The form of instantaneous unit hydrograph*. **Int. Assoc. Sci. Hydrol. Publ.,** 45(3), 114-121.

Nash, J. E., Sutcliffe, J. V. (1970): *River flow forecasting through conceptual models Part I - A discussion of principles.* **J. Hydrol.**, 10, 282-290.

Szilágyi, J. (1992): *Stage forecasting by an adaptive, stochastic model* (In Hungarian: Vízállások előrejelzése adaptív, sztochasztikus modellel), **Vizügyi Közlemények**, 74(1), 91-104.

Szilágyi, J. (2003): *State-space discretization of the Kalinin-Milyukov-Nash cascade in a sample-data system framework for streamflow forecasting*. **J. Hydrol. Eng.**, 8(6), 339-347.

Szilágyi, J. (2004): *Accounting for stream-aquifer interactions in the state-space discretization of the KMN-cascade for streamflow forecasting*. **J. Hydrol. Eng.**, 9(2), 135-143.

Szolgay, J. (1991): *Prediction of river runoff changes due to hydropower development on the Danube at Gabcikovo*. **IAHS Publ.** 201, 209-218.

Szöllősi-Nagy, A. (1989): ***Real-Time Streamflow Forecasting using Dynamically Structured Deterministic-Stochastic Models*** (In Hungarian: A Mederbeli Lefolyás Real-Time Előrejelzése Dinamikus Strukturális-Sztochasztikus Modellekkel), VITUKI, Budapest.