**Hydrology and Earth System Sciences**

Open Access

# Assessment of the indirect calibration of a rainfall-runoff model for ungauged catchments in Flanders

N. De Vleeschouwer[1] and V. R. N. Pauwels[2]

[1]Laboratory of Hydrology and Water Management, Ghent University, Ghent, Belgium
[2]Department of Civil Engineering, Monash University, Clayton, Victoria, Australia

*Correspondence to:* N. De Vleeschouwer (niels.devleeschouwer@ugent.be)

**Abstract.** In this paper the potential of discharge-based indirect calibration of the probability-distributed model (PDM), a lumped rainfall-runoff (RR) model, is examined for six selected catchments in Flanders. The concept of indirect calibration indicates that one has to estimate the calibration data because the catchment is ungauged or scarcely gauged. A first case in which indirect calibration is applied is that of spatial gauging divergence: because no observed discharge records are available at the outlet of the ungauged catchment, the calibration is carried out based on a rescaled discharge time series of a very similar donor catchment. Both a calibration in the time domain and the frequency domain (also known as spectral domain) are carried out. Furthermore, the case of temporal gauging divergence is considered: limited (e.g. historical or very recent) discharge records are available at the outlet of the scarcely gauged catchment. Additionally, no time overlap exists between the forcing and discharge records. Therefore, only an indirect spectral calibration can be performed in this case. To conclude also the combination case of spatio-temporal gauging divergence is considered. In this last case only limited discharge records are available at the outlet of a donor catchment. Again the forcing and discharge records are not concomitant, which only makes feasible an indirect spectral calibration. For most catchments the modelled discharge time series is found to be acceptable in the considered cases. In the case of spatial gauging divergence, indirect temporal calibration results in a better model performance than indirect spectral calibration. Furthermore, indirect spectral calibration in the case of temporal gauging divergence leads to a better model performance than indirect spectral calibration in the case of spatial gauging diver-gence. Finally, the combination of spatial and temporal gauging divergence does not lead to a notably worse model performance compared to the case of spatial gauging divergence.

## 1 Introduction

The practical application of rainfall-runoff (RR) models requires a proper assignment of the parameter values, also known as the process of parametrisation or calibration (Duan et al., 1992). Ideally, this calibration process should be fed by in situ measurements or remote sensing data. Practical considerations, however, implicate an alternative strategy. In a classic calibration framework, the parameter values are adjusted until the match between the modelled and observed output (e.g. discharge) is found to be acceptable. In practice the conditions to perform an ordinary direct calibration are not always fulfilled. This implies an indirect calibration strategy (Montanari and Toth, 2007). In the past decade, the research concerning indirect calibration has gained attention in the hydrologic community through the Prediction in Ungauged Basins (PUB) initiative (Sivapalan et al., 2003) set up by the International Association of Hydrological Sciences (IAHS). In scarcely gauged regions, discharge records may lack entirely for the catchment of interest, and may only be available at the outlet of a nearby catchment. This situation will be indicated in this paper by the term "spatial gauging divergence". A technique that copes with this problem and that has already received a lot of attention in the literature is that of parameter regionalisation (Seibert, 1999; Merz and Bloschl, 2004; Parajka et al., 2005; Bardossy, 2007;

Oudin et al., 2008). According to this approach the model parameters of an ungauged catchment are estimated based on the model parameters of other gauged catchments and other relevant information, e.g. spatial proximity or physical landscape properties. Another occurring problem in many scarcely gauged catchments is that of non-concomitant forcing (e.g. precipitation) and discharge records. Consequently, the modelled discharge cannot be compared to the observations. Hereafter, this case will be indicated by the term "temporal gauging divergence". In the case of an indirect calibration approach, it can be expected that the resulting predictive power of the model will be lower than the predictive power obtained by an ordinary direct calibration. Therefore, the research question is whether an acceptable predictive power of the model can be obtained in a certain case of gauging divergence.

An interesting technique useful for parameter estimation in ungauged and scarcely gauged catchments is spectral calibration (Montanari and Toth, 2007; Winsemius et al., 2009; Quets et al., 2010; Castiglioni et al., 2010; Pauwels and De Lannoy, 2011). In the ordinary form the spectral properties (e.g. the spectral density $S$) of both the observed and modelled output are matched instead of the time series themselves. In order to obtain those properties, one has to perform a transformation of the time series to the frequency domain. In the aforementioned cases of spatial and temporal gauging divergence, it is impossible to carry out a direct spectral calibration because observed outputs are missing in the calibration period for the catchment under consideration. Consequently the spectral properties of the non-observed discharge response need to be estimated. Montanari and Toth (2007) first illustrated the opportunities of indirect spectral calibration in hydrological modelling using a maximum likelihood estimator proposed by Whittle (1953). Under the condition of stationarity, the spectral densities of two observed time series separated in time have a higher degree of agreement than the observations in the time domain. This demonstrates the possibility of obtaining a proper estimate of the spectral density of a time series based on non-concomitant records. Furthermore, it is possible to carry out the calibration in absence of discharge records at the outlet of the considered catchment. The spectral density estimates can then be based on discharge time series in nearby catchments.

In this paper indirect calibration is applied to the probability-distributed model (PDM) (Moore, 2007) for six catchments in Flanders. By alternately considering these catchments gauged and ungauged, indirect calibration can be compared to direct calibration in terms of the predictive power of the RR model. Both the cases of spatial and temporal gauging divergence are examined. Additionally, the combination of temporal and spatial gauging divergence is considered.

## 2 Spectral properties: mathematical background

The spectral densities $S$ $[\text{L}^6\,\text{T}^{-2}]$ of a discharge time series without missing records can be approximated by calculating the periodogram $c^2$ $[\text{L}^6\,\text{T}^{-2}]$, which requires a transformation of the time series to the frequency domain. The discharge time series $Q(t)$ $[\text{L}^3\,\text{T}^{-1}]$ consisting of $D$ equally long time steps ($t \in [1, \ldots, D]$) can be written as a Fourier series (Thibos, 2003):

$$Q(t) = \sum_{k=0}^{N} \Psi(k) \left[ a(k)\cos\left(\frac{2\pi k}{D}(t-1)\right) \right.$$
$$\left. + b(k)\sin\left(\frac{2\pi k}{D}(t-1)\right) \right]. \tag{1}$$

$k\,[-] \in [0, \ldots, N]$ is the harmonic number. For $k > 0$ this variable determines the wavelength $\lambda\,[\text{T}]$ of the terms through the relationship $\lambda = \frac{L}{k}$, $L\,[\text{T}]$ being the length of the time series. If the time series consists of an even number of time steps, the highest harmonic $N = \frac{D}{2}$ (Shannon, 1984) and $\Psi(k)\,[-] \in [\frac{1}{2}, 1, \ldots, 1, \frac{1}{2}]$. If this is not the case, then $N = \frac{D-1}{2}$ and $\Psi(k) \in [\frac{1}{2}, 1, \ldots, 1]$. $a(k)$ and $b(k)$ are referred to as Fourier coefficients. The periodogram is calculated as $c^2(k) = \Psi(k^2)(a^2(k) + b^2(k))$. As a consequence the spectral densities for $k > 0$ are related to the amplitudes of the contributing goniometric functions. The spectral density for $k = 0$, however, is an estimate of the mean of the process. The spectral densities in function of the harmonic number $k$ are also called the density spectrum.

Since discharge time series usually contain record gaps, it is often not possible to perform this computationally efficient approximation of the discharge density spectrum. Therefore, the latter has to be estimated through the Wiener–Khinchin relationship (Papoulis, 1965; Brown and Hwang, 1992):

$$S(k) = \mathcal{F}[R(\tau)]. \tag{2}$$

$\mathcal{F}$ stands for the Fourier transformation. $R(\tau)\,[\text{L}^6\,\text{T}^{-2}]$, known as the correlation function in signal processing disciplines, is calculated as follows (Papoulis, 1965; Brown and Hwang, 1992):

$$R(\tau) = E[Q(t)Q(t-\tau)]. \tag{3}$$

Herein $\tau\,[\text{T}]$ stands for the temporal lag. For the calculation, stationarity of the time series is assumed. It could also be chosen to use the more commonly known autocorrelation function:

$$R'(\tau) = E\left[\frac{(Q(t) - \overline{Q})(Q(t-\tau) - \overline{Q})}{\sigma_Q^2}\right]. \tag{4}$$

$\overline{Q}\,[\text{L}^3\,\text{T}^{-1}]$ and $\sigma_Q\,[\text{L}^3\,\text{T}^{-1}]$ represent respectively the mean and standard deviation of the time series $Q(t)\,[\text{L}^3\,\text{T}^{-1}]$. As a consequence of subtracting the mean of the time series in the
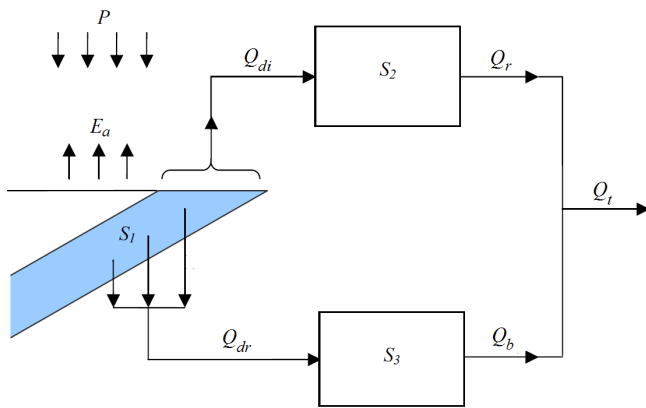
**Fig. 1.** General model structure of the PDM (Moore, 2007).

**Table 1.** Overview of the PDM parameters with indication of the lower and upper boundaries for catchments in Flanders.

| Parameter | Units | Lower boundary | Upper boundary |
|---|---|---|---|
| $c_{max}$ | [L] | 160 | 5000 |
| $c_{min}$ | [L] | 0 | 300 |
| $b$ | [–] | 0.1 | 2 |
| $b_e$ | [–] | 1 | 2 |
| $k_1$ | [T] | 0.9 | 40 |
| $k_2$ | [T] | 0.1 | 15 |
| $k_b$ | [$L^{-2}T^{-1}$] | 0 | 5000 |
| $k_g$ | [T] | 700 | 25 000 |
| $S_t$ | [L] | 0 | 150 |
| $b_g$ | [–] | 1 | 1 |
| tdly | [T] | 0 | 10 |
| $q_c$ | [$LT^{-1}$] | −4.08 | 0.03 |

numerator, the first spectral density will become zero. Because this is considered as a potential information loss, in the following calibration experiments a preference is given to Eq. (3). Calculating the entire correlation function ($\tau \in [0, 1, \ldots, D-1]$) becomes computationally intensive in the case of long time series. For this reason lower values are chosen for the maximum lag $\tau_{max}$ [T] in this paper. Consequently, the resulting variables are only approximations of the spectral density estimates, calculated with the full correlation function.

## 3 Model description

In this study the PDM, a lumped RR model, is used to simulate the discharge response in the considered catchments. The model basically consists of three storages to represent the water flow paths (see Fig. 1). The probabilistic distributed soil moisture storage $S_1$ [L] receives the net precipitation input $(P - aET)$ [$LT^{-1}$], $P$ [$LT^{-1}$] and aET [$LT^{-1}$] respectively being the gross precipitation and actual evapotranspiration. Based on the concept of Dunnian runoff (Dunne and Black, 1970), the net precipitation is partitioned into direct runoff $Q_{di}$ [$LT^{-1}$] and drainage $Q_{dr}$ [$LT^{-1}$]. The former is converted to surface runoff $Q_r$ [$LT^{-1}$] through a fast surface storage (cascade of two linear reservoirs), the latter to base flow $Q_b$ [$LT^{-1}$] through a slow subsurface storage. The sum of surface runoff and base flow equals the total discharge $Q_t$ [$LT^{-1}$], which can be multiplied by the catchment area $A$ [$L^2$] in order to get the volumetric discharge (see remainder of the article). The more detailed mathematical description of the PDM can be found in Appendix A. The model version in this research makes use of 12 parameters. An overview is given in Table 1. Additionally, the estimated lower and upper boundaries of these parameters in Flemish catchments are provided (Cabus, 2008).

## 4 Site description and data availability

Figure 2 shows a preselection of 32 catchments in the Scheldt and Yser basins in Flanders. The drainage areas range from 2 to 265 km$^2$. The size of the catchments considered in this study is thus rather small. The catchments are delineated based on a digital elevation model (DEM) with a spatial resolution of 25 m using the algorithms described in Jenson and Domingue (1988). For every catchment hourly discharge records are available at the outlet for five consecutive years (2006–2010). Hourly precipitation and potential evapotranspiration forcing records were obtained from the Flemish Environment Agency (VMM) monitoring network. Precipitation and potential evapotranspiration time series were available for the period 2005–2010 in respectively 14 and 4 weather stations (see Fig. 2). The year 2005 is used to initialise the PDM. Catchment specific forcing data were obtained using inverse square distance weighing (see Eq. 5) (Weber and Englund, 1992). It was decided to only include the three most nearby weather stations in the interpolation ($N = 3$).

$$z(x_i, y_i, t) = \sum_{k=1}^{N} \frac{\left[(x_i - x_k)^2 + (y_i - y_k)^2\right]^{-1}}{\sum_{j=1}^{N} \left[(x_i - x_j)^2 + (y_i - y_j)^2\right]^{-1}} z(x_k, y_k, t) \quad (5)$$

Herein $z(x_i, y_i, t)$ is the interpolated forcing $z$ at the point of gravity $(x_i, y_i)$ of catchment $i$ at timestep $t$. $z(x_k, y_k, t)$ is the forcing record measured at the $k$-th weather station out of $N$ at location $(x_k, y_k)$ and at timestep $t$. Furthermore, raster data with a spatial resolution of 25 m regarding land cover and soil type were obtained from the Flemish Geographical Information Agency (FGIA).

A subgroup of six catchments (see Table 2) is further selected for the calibration experiments. In the remainder of this paper, these catchments are considered to be ungauged or scarcely gauged and will be referred to with the term "autochthonous catchments". The subgroup is chosen in order to
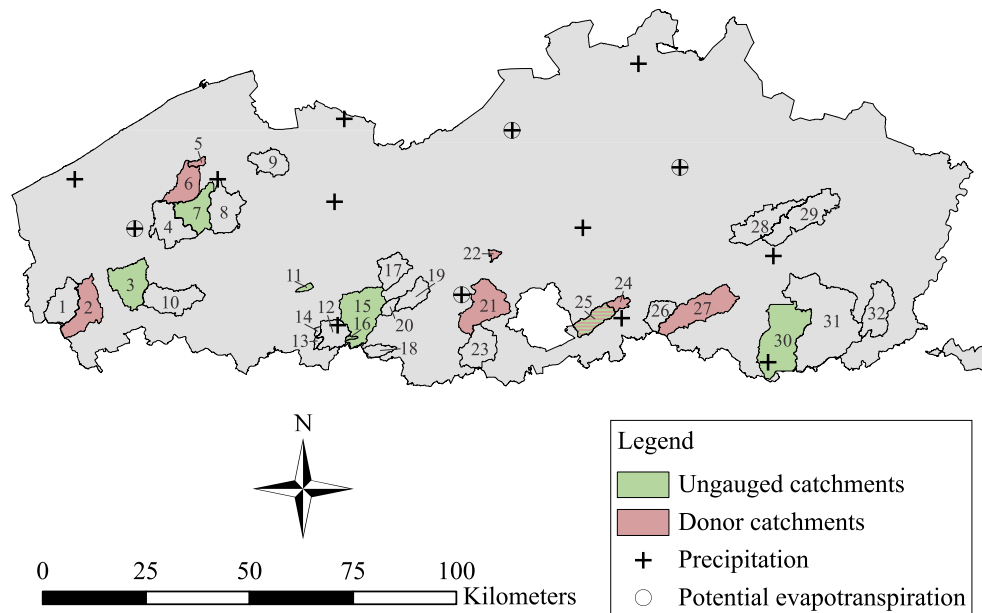
**Fig. 2.** The spatial spreading of the 32 catchments included in this study. The autochthonous catchments considered to be ungauged and the donor catchments are coloured green and red, respectively. Stations with precipitation and potential evapotranspiration measurement are indicated with crosses and circles, respectively.

obtain a certain diversity with respect to geographical location, drainage area, land cover, soil type, geomorphology and morphometry. In this way a certain bias in the conclusions of the calibration experiments should be minimised. In Fig. 2 the autochthonous catchments are coloured green.

## 5   Estimation of the spectral densities

First it should be made clear that the density spectra calculated following Eq. (2) are already approximations of the real spectra because of the finiteness of the used time series and the observation error. However, this is not indicated in the mathematical notation of Eq. (2) ($S$ instead of $\hat{S}$). The notation $\hat{S}$ is used in this study to indicate that the spectrum is indirectly estimated because the catchment is ungauged or scarcely gauged (see below).

### 5.1   Case of spatial gauging divergence

In order to estimate the density spectrum of the autochthonous discharge time series in the case of spatial gauging divergence, a donor catchment approach is introduced. This implies for every autochthonous catchment the identification of the catchment in the population of 31 remaining catchments with the most similar discharge density spectrum. In practice, this identification has to be performed indirectly because the autochthonous density spectrum is unknown. In this research a selection based on five catchment properties is proposed to identify the best donor catchment. The difference in drainage area and the mutual distance between the

points of gravity are considered to be the most determining properties. The drainage area is an important indicator of the discharge magnitude and thus the spectral density magnitude. Also the time constant of the hydrological response is influenced by this property (Post and Jakeman, 1996), which in turn influences the density spectrum. The drainage area dissimilarity between catchments $i$ and $j$ is expressed by a normalised dissimilarity index $\text{NDI}_A\,[-]$:

$$\text{NDI}_A(i, j) = \frac{|A_i - A_j|}{A_{\max} - A_{\min}}. \tag{6}$$

$A_i\,[\text{L}^2]$ is the drainage area of catchment $i$. The subscripts max and min indicate respectively the higest and lowest drainage area value in the population of 32 catchments. The mutual distance between the points of gravity of two catchments can serve as a measure for the difference in the observed meteorologic pattern. Significant differences in the latter can be reflected in the spectral properties of the discharge time series. The normalised dissimilarity with respect to the mutual distance is calculated by the $\text{NDI}_D\,[-]$:

$$\text{NDI}_D(i, j) = \frac{D_{i,j}}{D_{\max}}. \tag{7}$$

$D_{i,j}\,[\text{L}]$ is the distance between catchment $i$ and $j$. $D_{\max}\,[\text{L}]$ is the maximum distance between two catchments in the population. Another catchment property having an influence on the spectral density of a discharge time series is the land topography. For instance, steeper catchments are generally characterised by a higher surface runoff. Therefore, the high

**Table 2.** Overview of the selected autochthonous catchments and corresponding properties. $A$ is the drainage area of the catchment. $S_{me}$ is the local slope (mean slope of a grid cell) averaged over all grid cells within the catchment. sc%$_{max}$ and lc%$_{max}$ are respectively the soil class and land cover class with the highest relative area within the catchment.

| No. | autochthonous catchment | $A$ [km$^2$] | $S_{me}$ [%] | sc%$_{max}$ | lc%$_{max}$ |
|---|---|---|---|---|---|
| 3 | Merkem–Martjevaart | 78 | 2.74 | Sandy loam (81 %) | Cultivated fields (75 %) |
| 7 | Oostkamp–Rivierbeek | 63 | 2.02 | Sand (73 %) | Cultivated fields (49 %) |
| 11 | Huise–Plankbeek | 5 | 3.85 | Sandy loam (98 %) | Cultivated fields (90 %) |
| 15 | Nederzwalm–Zwalmbeek | 106 | 5.48 | Loam (70 %) | Cultivated fields (69 %) |
| 25 | Bertem–Voer | 36 | 6.43 | Loam (75 %) | Cultivated fields (62 %) |
| 30 | Rummen–Melsterbeek | 150 | 2.47 | Loam (75 %) | Cultivated fields (62 %) |

frequency parts of the discharge density spectrum (large $k$) will be higher than will be the case in horizontal catchments. In order to let this property interfere in the selection of the donor catchments, the following normalised dissimilarity index is introduced:

$$\text{NDI}_R(i, j) = \frac{|S_{me,i} - S_{me,j}|}{S_{me,max} - S_{me,min}}. \tag{8}$$

$S_{me,i}$ [%] is the mean local slope of catchment $i$. The local slope is calculated at the grid cell scale. The subscripts max and min indicate respectively the highest and lowest mean local slope in the population of 32 catchments. Soil composition and land cover are also incorporated in the selection framework. Both properties have an important influence on the infiltration rate and thus the runoff in a catchment. Therefore, soil composition and land cover can possibly have a proper influence on the pattern of the discharge density spectrum. The NDI$_S$ [−] and NDI$_L$ [−] are proposed to respectively account for dissimilarities in soil composition and land cover.

$$\text{NDI}_S(i, j) = \sum_{k=1}^{N_B} \phi_k \frac{|\text{sc\%}_{k,i} - \text{sc\%}_{k,j}|}{\text{sc\%}_{k,max} - \text{sc\%}_{k,min}} \tag{9}$$

$$\text{NDI}_L(i, j) = \sum_{k=1}^{N_L} \chi_k \frac{|\text{lc\%}_{k,i} - \text{lc\%}_{k,j}|}{\text{lc\%}_{k,max} - \text{lc\%}_{k,min}} \tag{10}$$

$N_B$ and $N_L$ are the number of soil and land cover classes. The relative areas of a certain soil or land cover class $k$ are presented by sc%$_k$ [%] and lc%$_k$ [%]. Again, the highest and lowest values for the considered variables are indicated with the subscripts max and min, respectively. Important to notice are the weights $\phi_k$ [−] and $\chi_k$ [−]. Those are equal to the mean relative area of a particular soil or land cover class in the two catchments to be compared. In this way rare soil or land cover classes cannot have a large influence on the donor catchment selection.

To assess the total dissimilarity between two catchments, a weighted sum of the aforementioned indices is calculated. In this study the following weights are used: 0.3 for the NDI$_A$ and NDI$_D$, 0.2 for the NDI$_R$ and 0.1 for the NDI$_S$ and NDI$_L$.

A high weight is given to the NDI$_D$ because proximity is generally an important indicator for the hydrological resemblance between two catchments (Oudin et al., 2008). The weights given to the other (catchment specific) indices are based on the analysed strength of the relationships between the spectral densities and those properties (drainage area > mean local slope > soil composition, land cover). For every autochthonous catchment the catchment with the lowest general dissimilarity is selected as the donor catchment. Table 3 gives an overview of the selected donor catchments. The same order is preserved as in Table 2; so for example catchment Reninge–Kemmelbeek is the donor catchment for catchment Merkem–Martjevaart. In Fig. 2 the six donor catchments are coloured red.

Subsequently, a rescaling of the donor discharge records is performed in order to improve the autochthonous time-series estimate. This rescaling (see Eq. 11) is based on the drainage area of the autochthonous and donor catchment because of the proper linear relationship (Pearson's correlation coefficient $R = 0.87$) between mean discharge (period 2006–2009) and the drainage area in the population of 32 Flemish catchments.

$$\hat{Q}_{aut}(t) = \frac{A_{aut}}{A_{don}} Q_{don}(t) \tag{11}$$

$\hat{Q}_{aut}(t)$ [L$^3$ T$^{-1}$] is the estimated autochthonous discharge time series; $Q_{don}(t)$ [L$^3$ T$^{-1}$] is the donor discharge time series. $A_{aut}$ [L$^2$] and $A_{don}$ [L$^2$] are the drainage areas of the autochthonous and donor catchment, respectively. Based on the aforementioned relationship between a time series and the corresponding spectral density spectrum, the estimated density spectrum of the autochthonous catchment $\hat{S}_{aut}(k)$ [L$^6$ T$^{-2}$] can be calculated as follows:

$$\hat{S}_{aut}(k) = \frac{A_{aut}^2}{A_{don}^2} S_{don}(k). \tag{12}$$

Herein $S_{don}(k)$ [L$^6$ T$^{-2}$] is the density spectrum of the donor discharge records. In Figs. 3 ($k = 0$) and 4 ($k > 0$) the actual and estimated root squared spectral densities of the discharge time series (period 2006–2009) in the six autochthonous catchments are presented. The maximum time

**Table 3.** Overview of the selected donor catchments and corresponding properties. $A$ is the drainage area of the catchment. $S_{me}$ is the local slope (mean slope of a grid cell) averaged over all grid cells within the catchment. sc%$_{max}$ and lc%$_{max}$ are respectively the soil class and land cover class with the highest relative area within the catchment.

| N° | Donor catchment | $A$ [km$^2$] | $S_{me}$ [%] | sc%$_{max}$ | lc%$_{max}$ |
|----|------------------|---------|---------|-------------|-------------|
| 2 | Reninge–Kemmelbeek | 72 | 2.89 | Sandy loam (86 %) | Cultivated fields (76 %) |
| 5 | Sint-Michiels–Kerkebeek | 59 | 1.83 | Sand (86 %) | Cultivated fields (41 %) |
| 22 | Opwijk–Vondelbeek | 5 | 2.58 | Sandy loam (85 %) | Cultivated fields (69 %) |
| 21 | Essene–Bellebeek | 88 | 4.92 | Loam (77 %) | Cultivated fields (66 %) |
| 24 | Heverlee–Voer | 49 | 6.35 | Loam (77 %) | Cultivated fields (64 %) |
| 27 | Ransberg–Velpe | 97 | 3.50 | Sandy loam (55 %) | Cultivated fields (68 %) |
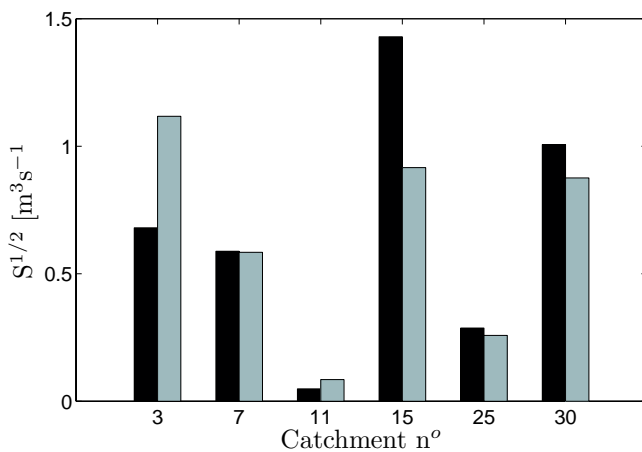


**Fig. 3.** Actual (black) vs. estimated (grey) root squared spectral density ($k = 0$) of the discharge time series for the six autochthonous catchments in the case of spatial gauging divergence (period 2006–2009). The estimates are rescaled root squared densities of the donor discharge time series.

lag $\tau_{max}$ [T] considered is 3 months. For certain catchments (e.g. Oostkamp–Rivierbeek and Bertem–Voer) a good match is obtained. This is however not the case for all catchments (e.g. Merkem–Martjevaart: spectral density for $k = 0$, Rummen–Melsterbeek: spectral densities for $k > 0$).

## 5.2 Case of temporal gauging divergence

The assumption of stationarity has as a consequence an invariable density spectrum. In the case of temporal gauging divergence, it is thus assumed that the density spectrum of the limited non-overlapping discharge time series is a good estimate of the density spectrum of the time series overlapping with the forcing records. In Figs. 5 ($k = 0$) and 6 ($k > 0$) the root squared density spectra of the periods 2006–2007 and 2008–2009 are compared for the six catchments under consideration. A proper match is found, and this to a greater extent for the high frequency parts of the spectra.

## 6 Calibration and validation

### 6.1 Test set-up

In this section, different calibration experiments are carried out in order to optimise the PDM for the autochthonous catchments considered in this study. Evaluation of the overall performance of the RR model is the main objective in this study. Therefore a multiple-year calibration period and a one-year validation period are considered. Furthermore the evaluation is carried out by aggregate assessment indicators. The calibration period runs from 1 January 2006 through 31 December 2009. The year 2005 serves as a initialisation year for the RR model. The first experiment encompasses a comparison between direct temporal calibration and direct spectral calibration. With regard to the latter, also the relationship between the maximum lag $\tau_{max}$ of the correlation function and the resulting model performance is examined. Furthermore the effect of assigning more weight to particular parts of the density spectrum in the objective function is examined. In a second experiment, the case of spatial gauging divergence is further examined. Direct spectral calibration is compared to indirect calibration in the time and frequency domain. For both indirect calibration set-ups, the estimates for the time series and spectral density are based on discharge records at the outlet of the donor catchments. The third experiment focusses on the case of temporal gauging divergence. The autochthonous discharge time series used in the calibration is limited and does not overlap with the forcing records. Additionally, in a fourth experiment a non-overlapping donor discharge time series is used in the calibration to examine the combined effect of spatial and temporal gauging divergence on the calibration of the hydrological model. The code names and properties of all calibration set-ups are listed in Table 4.

Because of the stochastic nature of the calibration algorithm (see below), each calibration set-up is applied three times for every autochthonous catchment. All repeated optimisations are assessed using four aggregate indicators: the Pearson correlation coefficient ($R$), the relative absolute bias (BIASn), the relative root mean square error (RMSEn) and the Nash–Sutcliffe coefficient (NS)
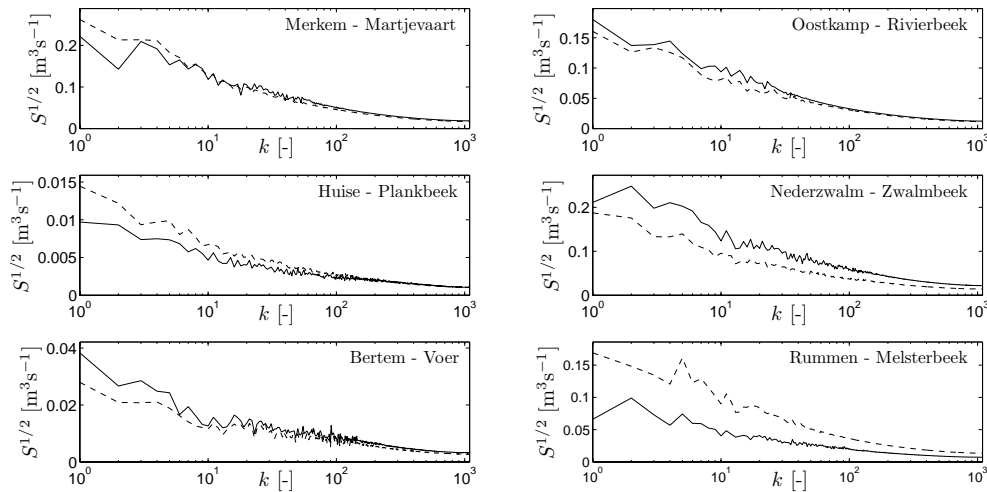
**Fig. 4.** Actual (solid line) vs. estimated (dashed line) root squared density spectrum ($k > 0$) of the discharge time series for the six autochthonous catchments in the case of spatial gauging divergence (period 2006–2009). The estimates are rescaled root squared densities of the donor discharge time series.

**Table 4.** Overview of the applied calibration set-ups.

| Experiment | Code | Domain | Gauging divergence | $\tau_{max}$ [months] | Weight type |
|---|---|---|---|---|---|
| 1, 2, 3, 4 | T-D | Time | – | – | – |
| 1 | F-D-1-0 | Frequency | – | 1 | – |
| 1, 2 | F-D-3-0 | Frequency | – | 2 | – |
| 1 | F-D-12-0 | Frequency | – | 3 | – |
| 1 | F-D-3-1 | Frequency | – | 3 | 1 |
| 1 | F-D-3-2 | Frequency | – | 3 | 2 |
| 1 | F-D-3-3 | Frequency | – | 3 | 3 |
| 2 | T-IS | Time | Spatial | – | – |
| 2 | F-IS-3-0 | Frequency | Spatial | 3 | – |
| 3 | F-IT-3-0 | Frequency | Temporal | 3 | – |
| 4 | F-IST-3-0 | Frequency | Spatio-temporal | 3 | – |

(Nash and Sutcliffe, 1970):

$$R = \frac{\sum\limits_{t=1}^{n}[Q_{\text{obs}}(t) - \overline{Q}_{\text{obs}}][Q_{\text{sim}}(t) - \overline{Q}_{\text{sim}}]}{\sqrt{\sum\limits_{t=1}^{n}[Q_{\text{obs}}(t) - \overline{Q}_{\text{obs}}]^2}\sqrt{\sum\limits_{t=1}^{n}[Q_{\text{sim}}(t) - \overline{Q}_{\text{sim}}]^2}} \quad (13)$$

$$\text{BIASn} = \frac{1}{\overline{Q}_{\text{obs}}}\left|\frac{1}{n}\sum_{t=1}^{n}[Q_{\text{obs}}(t) - Q_{\text{sim}}(t)]\right| \quad (14)$$

$$\text{RMSEn} = \frac{1}{\overline{Q}_{\text{obs}}}\sqrt{\frac{1}{n}\sum_{t=1}^{n}[Q_{\text{obs}}(t) - Q_{\text{sim}}(t)]^2} \quad (15)$$

$$\text{NS} = 1 - \frac{\sum\limits_{t=1}^{n}[Q_{\text{obs}}(t) - Q_{\text{sim}}(t)]^2}{\sum\limits_{t=1}^{n}[Q_{\text{obs}}(t) - \overline{Q}_{\text{obs}}]^2}. \quad (16)$$

$Q_{\text{obs}}[\text{L}^3\text{T}^{-1}]$ and $Q_{\text{sim}}[\text{L}^3\text{T}^{-1}]$ are the observed and simulated discharge values, respectively. $\overline{Q}_{\text{obs}}[\text{L}^3\text{T}^{-1}]$ and $\overline{Q}_{\text{sim}}[\text{L}^3\text{T}^{-1}]$ are the mean observed and simulated discharge, respectively. For every autochthonous catchment and calibration set-up, the assessment indicators of the repetition characterised by the lowest RMSEn are retained. The absolute bias and RMSE are divided by the mean observed discharge in order to obtain four dimensionless indicators. In this form it is possible to average the retained indicators over the six catchments without giving more weight to the indicators of catchments characterised by a high mean discharge.
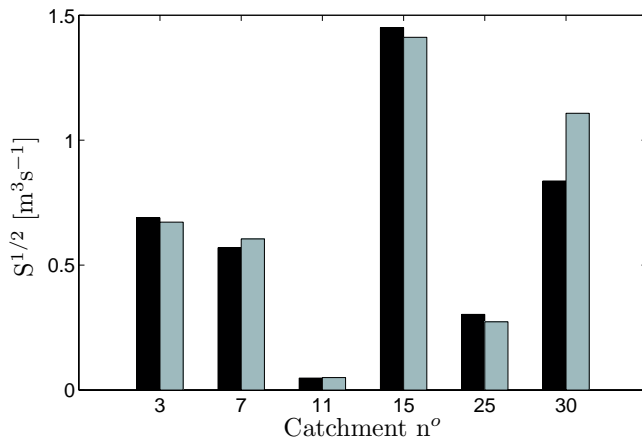
**Fig. 5.** Actual (black) vs. estimated (grey) root squared spectral density ($k = 0$) of the discharge time series for the six autochthonous catchments in the case of temporal gauging divergence (period 2006–2007). The estimates are root squared densities of the autochthonous discharge time series during the period 2008–2009.

As hydrologic models increasingly become more sophisticated, the iterative parameter adjustments in the calibration process are usually performed by specific optimisation algorithms. Commonly used algorithms in hydrologic modelling are, for example, genetic algorithms (Reed et al., 2000) like the shuffled complex evolution algorithm (SCE-UA) (Duan et al., 1992), local and multistart simplex methods (Gan and Biftu, 1996), particle swarm optimisation (PSO) (Kennedy and Eberhart, 1995), simulated annealing (SA) (Thyer et al., 1999), etc. In this research particle swarm optimisation (PSO) (Kennedy and Eberhart, 1995) is applied as optimisation algorithm. The ability of PSO to find optimal solutions for hydrological modelling issues has already been demonstrated in various studies (Gill et al., 2006; Scheerlinck et al., 2009; Tolson et al., 2009; Zhang and Chiew, 2009; Mousavi and Shourian, 2010; Liu and Han, 2010; Pauwels and De Lannoy, 2011). This swarm intelligence algorithm is based on the movement of different particles throughout the $n$-dimensional parameter space. This movement is controlled by the particle's own history of positions (and thus related values of the objective function) and that of neighbouring particles, resulting in a so-called global behaviour. In order to adjust this behaviour so that a convergence to the global optimum is found, a parametrisation of the calibration algorithm is required. The type of PSO applied in this paper is characterised by the parameter vector $\psi = [N_i \, N_k \, c_1 \, c_2 \, w \, \delta]^T$ (for description parameters, see Table 5).

$N_i$ and $N_k$ are assigned a value of respectively 30 and 36 (Pauwels and De Lannoy, 2011). The main reason for giving a fixed value to $N_i$ and $N_k$ in advance is to reduce the dimension of the parameter vector to be optimised. The values of the remaining parameters are selected out of the following

discrete intervals, also applied in Pauwels and De Lannoy (2011):

1. $c_1 \in \{0.8, 1.0, \ldots, 1.8\}$
2. $c_2 \in \{1.0, 1.2, \ldots, 2.2\}$
3. $w \in \{0.2, 0.4, 0.6\}$
4. $\delta \in \{0.2, 0.4, 0.6\}$.

For every parameter combination a direct temporal calibration of the PDM is performed for the catchment Merkem–Martjevaart. The selection of the optimal parameter vector is based on the RMSE between the $n$ observed discharge records $Q_{\text{obs}}$ and model simulations $Q_{\text{sim}}$. The lowest RMSE ($0.637 \, \text{m}^3 \, \text{s}^{-1}$) is found with $\psi = [30 \, 36 \, 1.8 \, 2.2 \, 0.2 \, 0.4]^T$. There is a problem of equifinality (Beven and Binley, 1992) as a high amount of parameter combinations give near optimal RMSEs. Due to practical considerations the above-mentioned most optimal set is applied in all of the following calibration set-ups. $N_i$ and $N_k$ have an influence on the convergence of the algorithm. Because the case presented in Pauwels and De Lannoy (2011) is not entirely comparable to this study, the convergence had to be checked for the following calibration experiments. Convergence is indeed obtained in the experiments, which indicates that the values for $N_i$ and $N_k$ are well chosen.

The model calibration is followed by a validation. The validation period runs from 1 January 2010 till 31 December 2010. For this the year 2009 is used to initialise the PDM. As in the calibration the same four dimensionless indicators are used (averaged over the six autochthonous catchments) to make an assessment of the calibration set-ups.

### 6.2 Results and discussion

#### 6.2.1 Experiment 1: direct temporal vs. spectral calibration

In this first experiment direct temporal calibration is compared to direct spectral calibration. With regard to the former, the RMSE between the observed and simulated discharges is used as objective function, which implicitly assumes a normal distribution of the residuals. Spectral calibration on the other hand makes use of a spectral objective function. For example this can be the RMSE between the spectral densities of the observed and simulated time series. However, better results are achieved by an RMSE between the root squared spectral densities (Quets et al., 2010; Pauwels and De Lannoy, 2011). This type of calibration is fundamentally different from the above-mentioned time domain calibration. Because the spectral properties are calculated by a Fourier transformation of the correlation function, a matching of the those properties will result more in a matching of the correlation structure of the observed and modelled time series (Montanari and Toth, 2007). So a calibration in the spectral
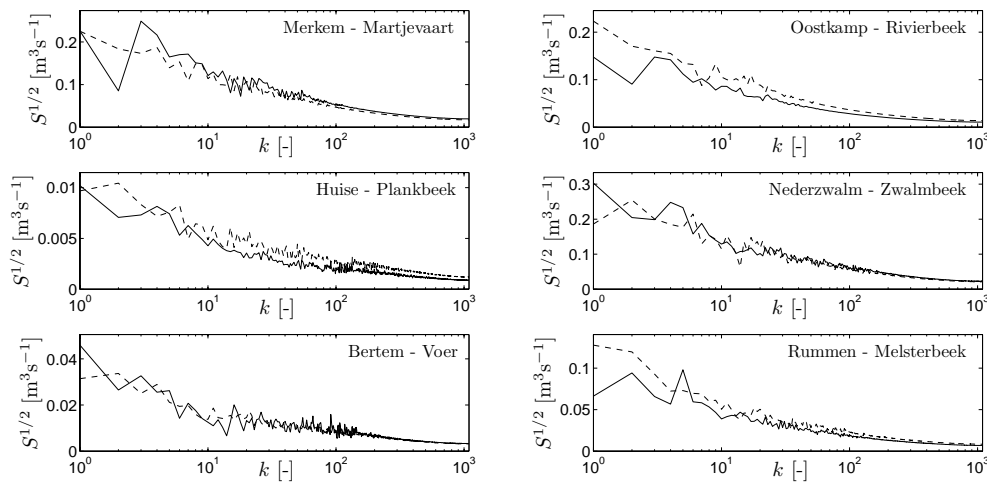
**Fig. 6.** Actual (solid line) vs. estimated (dashed line) root squared density spectrum ($k > 0$) of the discharge time series for the six autochthonous catchments in the case of temporal gauging divergence (period 2006–2007). The estimates are root squared densities of the autochthonous discharge time series during the period 2008–2009.

**Table 5.** Overview of PSO algorithm parameters.

| Parameter | Description |
|---|---|
| $N_i$ | Particle population size |
| $N_k$ | Iterations |
| $c_1$ | Cognitive parameter |
| $c_2$ | Social parameter |
| $w$ | Inertion weight |
| $\delta$ | Velocity limiter |

domain is more about matching the shape characteristics of the observed and modelled time series. As the spectral density for $k = 0$ is an estimate of the time-series mean, giving a considerable weight to this density in the objective function should also result in a bias minimisation. Furthermore, it should be emphasised that the error between observed and modelled root squared spectral densities cannot be assumed to be normal in the case of normal time domain residuals. However, for sake of simplicity the RMSE between the root squared spectral densities is applied as objective function in this research. As mentioned in the introduction, Montanari and Toth (2007) made use of another objective function, based on the Whittle likelihood estimator. First of all this objective function does not include the first harmonic. Furthermore, statistically seen this estimator is more appropriate as a spectral objective function. However, the use of this estimator requires a full characterisation of the model error, which is in fact not entirely known. Therefore it is chosen not to use this estimator.

First, the influence of the maximum lag of the correlation function $\tau_{max}$ on the post-calibration model performance is examined. Three values for $\tau_{max}$ are proposed: 1, 3 and 12 months. In Fig. 7 (upper-left subplot) the averaged assess-

ment indicators for the calibration period (2006–2009) are compared for set-ups T-D, F-D-1-0, F-D-3-0 and F-D-12-0. It is clear that discharge is mostly better simulated after a calibration in the time domain (higher $R$, higher NS coefficient and lower RMSEn). A clearly poorer model performance can be noted when applying a low $\tau_{max}$. This is probably due to the information loss, which is a consequence of the limitation of the correlation function. As can be expected in the validation period (see Fig. 7, upper-right subplot), the BIASn and RMSEn worsened for all four set-ups compared to the calibration period. This is however not entirely the case for the $R$ and NS coefficient. Furthermore, the set-ups relate differently with regard to model performance. For three out of four assessment indicators, temporal calibration still leads to better results; however, the differences with spectral calibration using a $\tau_{max}$ of 3 months are very small. Spectral calibration with the longest correlation function ($\tau_{max} = 12$ months) leads to a poorer BIASn, RMSEn and NS coefficient compared to spectral calibration based on the correlation function with $\tau_{max} = 3$ months. Correlation function values at higher lags represent more noise effects than physical system characteristics. If longer correlation functions are considered, more weight is given to noise effects in the calibration process. This could explain why spectral calibration with a longer correlation function does not necessarily lead to the best model performance in the validation period. Because of this observation a $\tau_{max}$ value of 3 months is proposed for all following spectral calibration set-ups. A second influence on the calibration exercise examined in this experiment is giving weights to certain parts of the density spectrum in the objective function. This is performed by the principle of exponential clustering (see Fig. 8). Three types of weights are proposed:
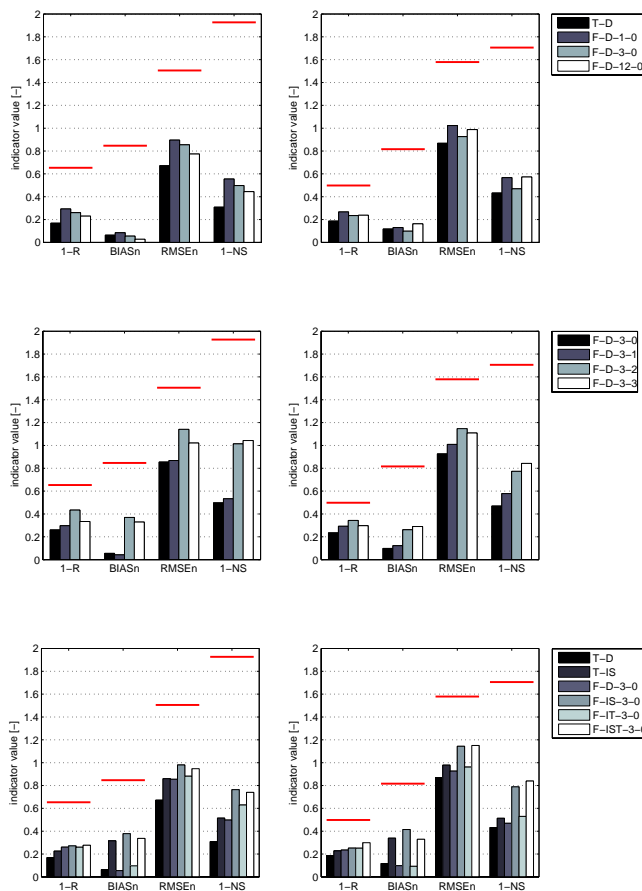
**Fig. 8.** Principle of exponential clustering and averaging (exponent 2).



**Fig. 7.** Comparative bar charts of the assessment indicator values for the different calibration set-ups. Left panels: calibration period, right panels: validation period. The red lines show the average indicator values in case of random sampling the parameters in the confined parameter space, considered in this study.

- Type 1: the spectral densities for $k < 9$ retain their value. Exponential clustering and averaging over those clusters is applied from $k = 9$ to $k = N$. Higher weights are thus assigned to the low frequency part of the density spectrum.

- Type 2: the spectral density for $k = 0$ is not considered in the objective function. Exponential clustering and averaging over those clusters is applied from $k = 1$ to $k = N$. Higher weights are thus assigned to the low frequency part of the density spectrum with a zero weight for the first spectral density.

- Type 3: exponential clustering and averaging over those clusters is applied from $k = N$ to $k = 0$. Higher weights are thus assigned to the high frequency part of the density spectrum.

It can be concluded from Fig. 7 (middle subplots) that giving weights to certain parts of the density spectrum generally
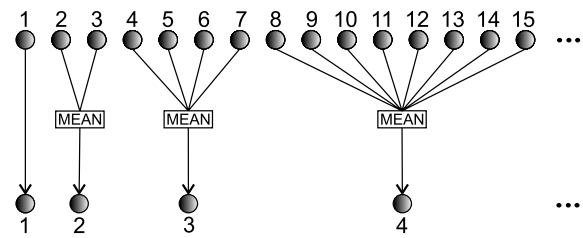
does not lead to a better model performance in the calibration and validation period. This applies to a greater extent for weight types 2 and 3. It is clear that barely or not taking into account the density for $k = 0$ generally leads to a worse model performance. Bias minimisation apparently results in a better model performance. Because of the aforementioned conclusions, no types of weight are applied in the indirect spectral calibrations of experiments 2, 3 and 4.

### 6.2.2 Experiment 2: indirect spectral calibration in the case of spatial gauging divergence (set-up F-IS-3-0)

This second calibration experiment focusses on the discharge prediction in an autochthonous catchment without available discharge records at the outlet. However, discharge records are available at the outlet of a donor catchment in the same time window as the forcing records monitored in the autochthonous catchment (spatial gauging divergence). Two calibration strategies are undertaken: a temporal calibration on the rescaled donor discharges (see Eq. 11) (set-up T-IS) and a spectral calibration on the root squared spectral densities of the rescaled donor discharges (see Eq. 12) (set-up F-IS-3-0). A first observation in the calibration period (see Fig. 7, lower-left subplot) is the reduced model performance after an indirect calibration compared to a direct calibration. Amongst others this can be explained by the rescaling of the donor discharges. Donor catchment selection and rescaling based on the drainage area namely is not a perfect estimating framework for the mean and shape of the discharge signature. Furthermore indirect temporal calibration clearly leads to a better model performance than indirect spectral calibration. $R$, BIASn and NS do not change strongly in the validation period (see Fig. 7, lower-right subplot). RMSEn on the other hand again increases more clearly. In the case of spatial gauging divergence, it seems to be recommended to apply indirect temporal calibration. This should however be nuanced. The spatial dimensions considered in this research are rather small (magnitude kilometres). Due to this proximity the time lapse for a certain meteorological event to happen in the autochthonous and donor catchment will be rather small (magnitude minutes–hours). If however the distance between the autochthonous and donor catchment is larger, the time
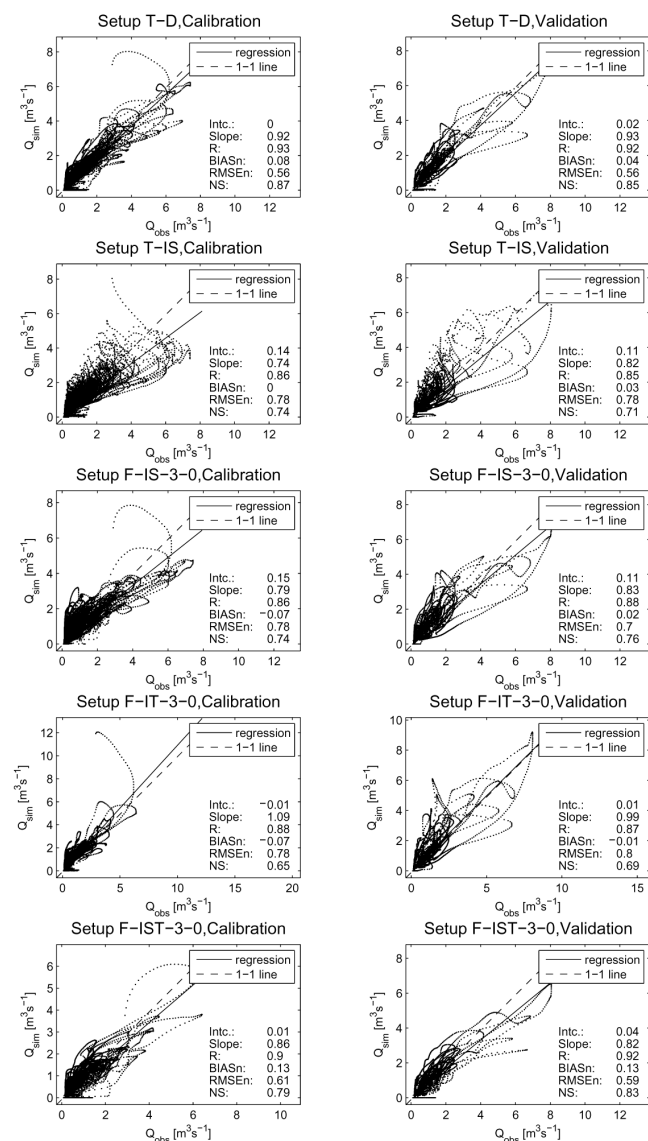
**Fig. 9.** Scatter plots of observed and simulated discharge in catchment Oostkamp–Rivierbeek for calibration set-ups T-D (1), T-IS (2), F-IS-3-0 (3), F-IT-3-0 (4) and F-IST-3-0 (5) in the calibration (left) and validation (right) period.

lapse increases (magnitude hours–days). It is even possible that the same meteorological event will not pass over both catchments. The applicability of indirect temporal calibration should therefore be evaluated in advance based on the proximity between the autochthonous and donor catchment.

### 6.2.3 Experiment 3: indirect spectral calibration in the case of temporal gauging divergence (set-up F-IT-3-0)

In this experiment limited discharge records are available for the autochthonous catchments; however, they are not concomitant with the forcing records. Specifically for this ex-

periment the time window of the observed discharge time series runs from 1 January 2008 through 31 December 2009. Forcing records are available from 1 January 2006 through 31 December 2007. In this way there is no overlap between the forcing and discharge records. The density spectrum of the observed discharge time series serves as an estimate for the density spectrum of the discharge time series concurrent with the available forcing records. The indicator values of set-up F-IT-3-0 in Fig. 7 (lower-left subplot) are based on the discharge simulations over the shortened calibration period (2006–2007). With exception of the NS coefficient, both direct spectral calibration and indirect spectral calibration in the case of temporal gauging divergence exhibit, almost similar indicator values during the calibration period. Compared to indirect spectral calibration in the case of spatial gauging divergence almost similar ($R$) or better indicator values (BIASn, RMSEn and NS) are obtained in the case of temporal gauging divergence. This will be due to the fact that the calibration data are observed at the outlet of the autochthonous catchment itself and not in a donor catchment. For the considered data set interannual discharge variation within the same catchment is thus smaller than discharge variation between the most similar catchments within the same year (in the calibration period). It should be emphasised that in this experiment the non-concomitant discharge and forcing time series are situated very close together in time. It is not unimaginable that the assessment indicators could turn worse in case of a larger time lapse. For example many former colonies dispose of historical hydrological data because post-colonial civil warfare hindered hydrological monitoring (Winsemius et al., 2006, 2009). These historical time series are often incomplete and are characterised by a larger measurement uncertainty. It is also possible that over time the meteorological conditions or the hydrological response of the catchment changes, for example, by an anthropogenic influence (Immerzeel and Droogers, 2008; Coe et al., 2011). All of this can limit the success of indirect spectral calibration in the case of temporal gauging divergence. In the validation period (see Fig. 7, lower-right subplot) the following is observed: all assessment indicators of the indirect spectral calibration (set-up F-IT-3-0) are almost comparable to those obtained with direct spectral calibration (set-up F-D-3-0). Furthermore only the RMSE and NS coefficient are notably worse compared to the value obtained in the calibration period. It can be concluded that indirect spectral calibration based on limited autochthonous discharge time series (set-up F-IT-3-0) results in a better model performance than indirect spectral calibration based on rescaled donor discharges (set-up F-IS-3-0).

### 6.2.4 Experiment 4: indirect spectral calibration in the case of spatio-temporal gauging divergence (set-up F-IST-3-0)

Experiment 4 is a combination of experiments 2 and 3. In this case no discharge observations are available at the outlet of the autochthonous catchment. In a very similar donor catchment, limited discharge records are available. However, no overlap exists between the time windows of the donor discharges and the forcing records in the autochthonous catchment. In this particular experimental design, discharge time series are available in the donor catchment from 1 January 2008 through 31 December 2009 and forcing data from 1 January 2006 through 31 December 2007. The density spectrum of the donor discharges will serve as an estimate for the density spectrum of the autochthonous discharge time series concurrent with the available forcing records. The indicator values of set-up F-IST-3-0 in Fig. 7 (lower-left subplot) are also based on the discharge simulations over the shortened calibration period (2006–2007). Because of the double introduced uncertainty in the calibration data, it is not illogical to assume that calibration set-up F-IST-3-0 would lead to a poorer model performance than set-ups F-IT-3-0 and F-IS-3-0. However, the assessment indicators in the calibration and validation period (see Fig. 7, lower subplots) indicate that the model performance is comparable to set-up F-IS-3-0. This again indicates that in this study the introduced error on donor-based calibration data is much larger than the error on non-concomitant autochthonous-based calibration data.

The values presented in Fig. 7 only show values averaged over all six catchments. This is interesting in order to assess the general performance for an average catchment in Flanders. However, a significant variability is observed over the six catchments. As illustration in Figs. 9 and 10 the scatter plots are shown resulting from the direct temporal calibration set-up (T-D) and the four indirect calibration set-ups (T-IS, F-IS-3-0, F-IT-3-0 and F-IST-3-0) for catchment Oostkamp–Rivierbeek and Rummen–Melsterbeek. Figure 11 shows corresponding observed and simulated hydrograph for a certain part of the validation period. In catchment Oostkamp–Rivierbeek very good indirect estimates of the calibration data are obtained. This results in very good indirect calibrations which could, depending on the set-up (e.g. set-up F-IST-3-0), be comparable to the time domain calibration. This shows the real potential of the presented indirect calibration techniques. However like the spectral density estimates and discharge simulations for catchment Rummen–Melsterbeek make clear, so far indirect calibration does not always lead to good model performances. In future research making use of other relevant information as for example correlation properties of precipitation time series (Toth, 2013) could possibly lead to an improved estimation of the calibration data. It should therefore be made clear which catchment properties have the highest predictive power for calibration data in a
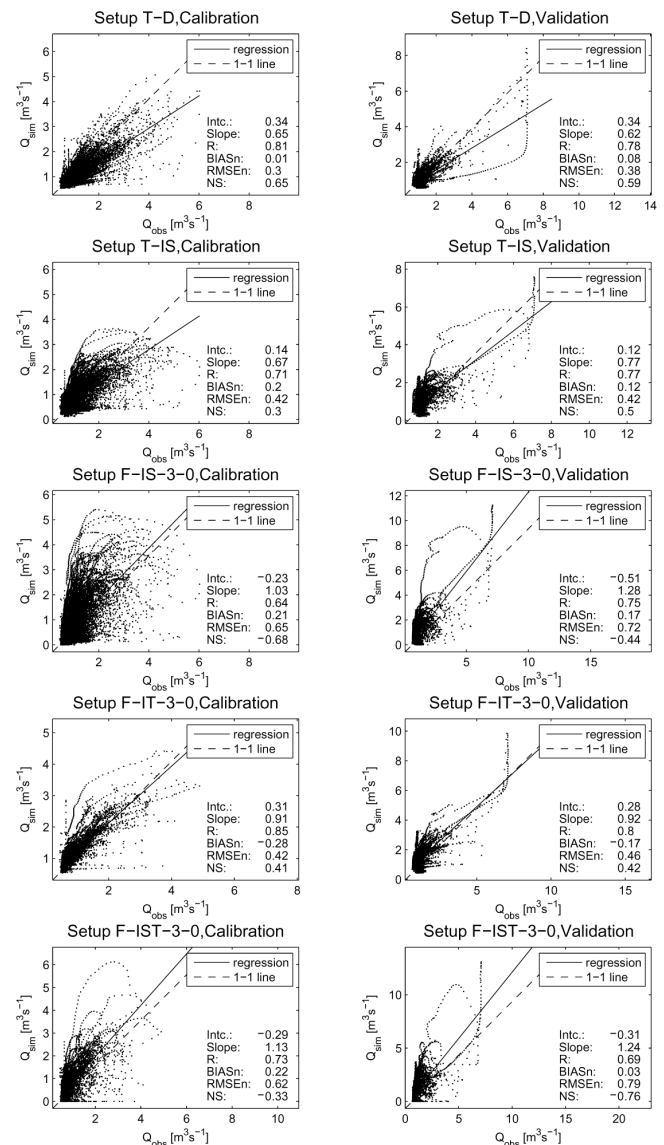


**Fig. 10.** Scatter plots of observed and simulated discharge in catchment Rummen–Melsterbeek for calibration set-ups T-D (1), T-IS (2), F-IS-3-0 (3), F-IT-3-0 (4) and F-IST-3-0 (5) in the calibration (left) and validation (right) period.

particular ungauged catchment. In this way more robust estimation frameworks for calibration data could be developed.

## 7 Conclusions

In this paper, an assessment is made of indirect calibration of the PDM for six (autochthonous) catchments in Flanders, considered to be ungauged or scarcely gauged. As calibration, algorithm PSO is applied. The described results are based on rather small catchments (magnitude 10–100 km$^2$) with a high proximity to each other (magnitude kilometres). The different calibration set-ups are evaluated on the
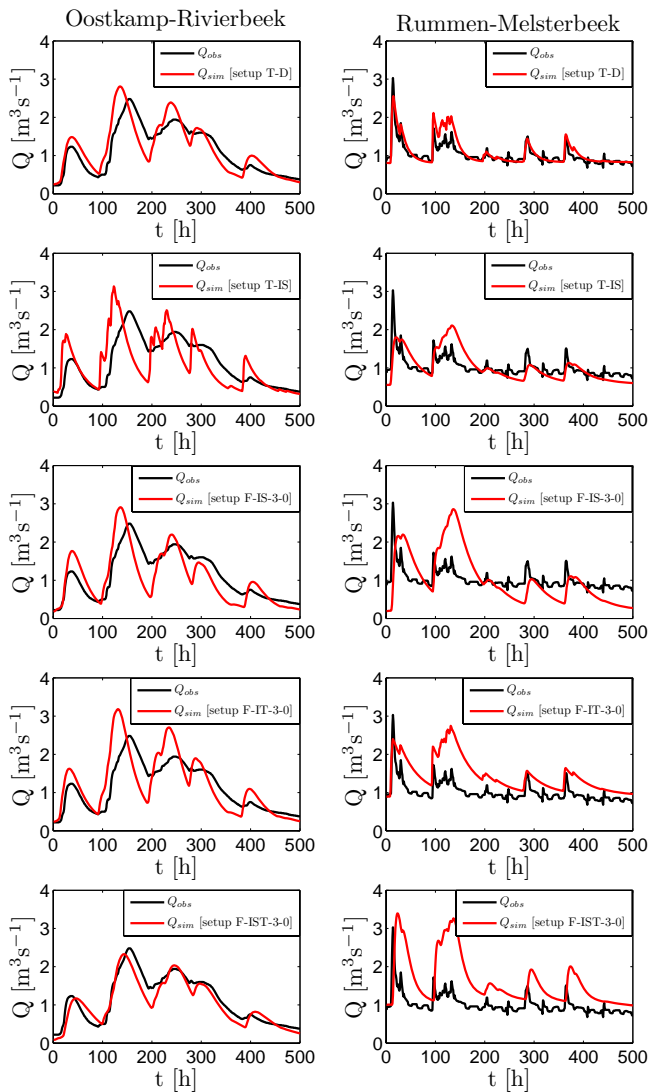
**Fig. 11.** Comparison of the observed and simulated time series for calibration set-ups T-D (1), T-IS (2), F-IS-3-0 (3), F-IT-3-0 (4) and F-IST-3-0 (5) in catchments Oostkamp–Rivierbeek and Rummen–Melsterbeek. A part of the validation period is shown.

basis of four assessment indicators ($R$, BIASn, RMSEn and NS). Those are averaged over the six autochthonous catchments. Consequently, the discussed results are valid for an average catchment in Flanders. The first calibration experiment focusses on direct spectral calibration. For this, the root squared spectral densities of the autochthonous discharge time series are incorporated in the objective function. The experiment revealed that higher values for the maximum lag of the correlation function ($\tau_{max}$) result in a better model performance during the calibration period. This is not necessarily the case in the validation period. Furthermore giving more weight to certain parts of the density spectrum during the spectral calibration generally does not lead to better model performances. In experiments 2, 3 and 4 indi-

rect calibration is examined in the cases of spatial, temporal and spatio-temporal gauging divergence. In those cases the calibration is based on the root squared spectral densities of rescaled donor discharge records, non-overlapping autochthonous discharge records and non-overlapping rescaled donor discharge records. Except for some specific indicator values, the model performance decreased compared to direct calibration but remained at an acceptable level for most catchments. With regard to indirect calibration in the case of spatial gauging divergence, better results were obtained using indirect temporal calibration vs. indirect spectral calibration. Indirect temporal calibration is however impossible to execute in the case of temporal and spatio-temporal gauging divergence. Therefore only indirect spectral calibration is applied in experiments 3 and 4. Generally better model performances were obtained in experiment 3 compared to experiment 2. This is due to the high uncertainty associated with the estimation of a density spectrum of an autochthonous discharge time series based on donor discharges. For certain catchments this can introduce a bias in the model results. It was expected that the model performance in experiment 4 would be worse than in experiment 2 and 3 because a double source of uncertainty is introduced in the calibration data. However, the assessment indicators show that this does not have to be the case.

This paper has shown that a certain potential exists for indirect calibration of a rainfall-runoff model (in particular indirect spectral calibration) in the case of spatial, temporal and spatio-temporal gauging divergence in ungauged catchments. Future research should focus on a refinement of the calibration framework (e.g. more complex objective functions, better estimation of the spectral densities, etc.) and on making the estimation of the calibration data more robust to guarantee a good indirect calibration. It could be challenging but interesting to examine the link between specific catchment properties (e.g. topography, land cover, precipitation signatures, etc.) and the resulting model performance after indirect calibration. For example it may be worthwhile to test whether indirect spectral calibration is more suitable in catchments with certain shape characteristics of the hydrograph (e.g. a short time to peak, a long recession period).

## Appendix A

### Equations of the PDM (Moore, 2007)

In the following section not further specified variables are model parameters and can be found in Table 1.

In the preparatory first step, five constants need to be calculated. The calculation of the maximum store capacity of the soil moisture storage $S_{max}$ [L] is based on the minimum ($c_{min}$ [L]) and maximum absorption capacity ($c_{max}$ [L]):

$$S_{max} = \frac{b(c_{min} + c_{max})}{b + 1}. \tag{A1}$$

The constants $\delta_1$ [−], $\delta_2$ [−], $\omega_0$ [−] and $\omega_1$ [−] are computed using the time constants of the first ($k_1$ [h]) and second surface storage ($k_2$ [h]):

$$\delta_1 = e^{-\frac{\Delta t}{k_1}} + e^{-\frac{\Delta t}{k_2}} \tag{A2}$$

$$\delta_2 = e^{-\frac{\Delta t}{k_1}} e^{-\frac{\Delta t}{k_2}} \tag{A3}$$

$$\omega_0 = \frac{k_1(e^{-\frac{\Delta t}{k_1}} - 1) - k_2(e^{-\frac{\Delta t}{k_2}} - 1)}{k_2 - k_1} \tag{A4}$$

$$\omega_1 = \frac{k_2(e^{-\frac{\Delta t}{k_2}} - 1)e^{-\frac{\Delta t}{k_1}} - k_1(e^{-\frac{\Delta t}{k_1}} - 1)e^{-\frac{\Delta t}{k_2}}}{k_2 - k_1}. \tag{A5}$$

The actual evapotranspiration $aET\,[LT^{-1}]$ is calculated on the basis of the potential evapotranspiration $pET\,[LT^{-1}]$, the store capacity in the soil moisture storage $S_1$ [L] at the previous time step and $S_{max}$ [L].

$$aET(t) = pET(t) \left( 1 - \left[ \frac{S_{max} - S_1(t-1)}{S_{max}} \right]^{b_e} \right) \tag{A6}$$

For the calculation of the drainage $Q_{dr}\,[LT^{-1}]$, $S_1$ [L] of the previous time step and the parameters $k_g$ [T], $S_t$ [L] and $b_g$ [−] are required:

if $S_1(t-1) \le S_t$,
$\quad Q_{dr}(t) = 0$;

if $S_1(t-1) > S_t$,
$\quad Q_{dr}(t) = \frac{1}{k_g}(S_1(t-1) - S_t)^{b_g}$. $\tag{A7}$

Next, the net precipitation $\pi\,[LT^{-1}]$ can be calculated as follows:

$$\pi(t) = P(t) - aET(t) - Q_{dr}(t). \tag{A8}$$

Consequently the direct runoff $Q_{di}\,[LT^{-1}]$ can be computed using the critical store capacity $C^*$ [L] of the previous time step, $\pi\,[LT^{-1}]$ and the parameters $c_{min}$ [L], $c_{max}$ [L] and $b$ [−]:

if $C^*(t-1) + \pi(i)\Delta t < c_{max}$,
$\quad Q_{di}(t)\Delta t = \pi(t)\Delta t - \frac{c_{max}-c_{min}}{b+1}\left[\frac{c_{max}-C^*(t-1)}{c_{max}-c_{min}}\right]^{b+1}$
$\qquad + \frac{c_{max}-c_{min}}{b+1}\left[\frac{c_{max}-C^*(t-1)-\pi(t)\Delta t}{c_{max}-c_{min}}\right]^{b+1}$;

if $C^*(t-1) + \pi(t)\Delta t \ge c_{max}$,
$\quad Q_{di}(t)\Delta t = \pi(t)\Delta t - \frac{c_{max}-c_{min}}{b+1}\left(\left[\frac{c_{max}-C^*(t-1)}{c_{max}-c_{min}}\right]^{b+1}\right)$
$\qquad + C^*(t-1) + \pi(t)\Delta t - c_{max}$. $\tag{A9}$

Once $\pi\,[LT^{-1}]$ and $Q_{di}\,[LT^{-1}]$ are known for the current time step, $S_1$ [L] can be calculated for the current time step:

if $S_1(t) \le 0$,
$\quad S_1(t) = 0$;

if $0 < S_1(t) < S_{max}$,
$\quad S_1(t) = S_1(t-1) + (\pi(t) - Q_{di}(t))\Delta t$;

if $S_1(t) \ge S_{max}$,
$\quad S_1(t) = S_{max}$. $\tag{A10}$

$C^*$ [L] can be calculated on the current time step on the basis of $S_1$ [L] at the current time step, $S_{max}$ [L] and the parameters $c_{min}$ [L], $c_{max}$ [L] and $b$ [−]:

if $C^*(t) \le 0$,
$\quad C^*(t) = 0$;

if $C^*(t) \ge c_{max}$,
$\quad C^*(t) = c_{max} - (c_{max} - c_{min})\left(\left[\frac{S_{max}-S_1(t)}{S_{max}-c_{min}}\right]^{\frac{1}{b+1}}\right)$;

if $C^*(t) \ge c_{max}$,
$\quad C^*(t) = c_{max}$. $\tag{A11}$

The capacity store of the subsurface storage $S_3$ [L] is computed using $S_3$ at the previous time step, $Q_{dr}\,[LT^{-1}]$ at the current time step and the parameter $k_b\,[L^{-2}T^{-1}]$.

$$S_3(t) = S_3(t-1) - \frac{e^{-3k_b S_3^2(t-1)\Delta t} - 1}{3k_b S_3^2(t-1)}$$
$$(Q_{dr}(t) - k_b S_3^3(t-1)) \tag{A12}$$

Making use of $S_3$ [L] and $k_b\,[L^{-2}T^{-1}]$, the base flow $Q_b\,[LT^{-1}]$ can be calculated:

if $Q_b(t) \le 0$,
$\quad Q_b(t) = 0$;

if $Q_b(t) > 0$,
$\quad Q_b(t) = k_b S_3^3(t)$. $\tag{A13}$

The surface runoff $Q_r\,[LT^1]$ is calculated as follows:

$$Q_r(t) = -\delta_1 Q_r(t-1) - \delta_2 Q_r(t-2) + \omega_0 Q_d(t)$$
$$+ \omega_1 Q_d(t-1). \tag{A14}$$

Eventually the total discharge $Q_t\,[LT^1]$ can be calculated as the sum of $Q_r$ and $Q_b$:

$$Q_t(t) = Q_r(t) + Q_b(t). \tag{A15}$$

during the research. Finally the authors also want to express their gratitude to Elena Toth (editor), Bettina Schaefli, Vahid Nourani and an anonymous reviewer for their constructive comments during the reviewing process.

Edited by: E. Toth

# References

Bárdossy, A.: Calibration of hydrological model parameters for ungauged catchments, Hydrol. Earth Syst. Sci., 11, 703–710, doi:10.5194/hess-11-703-2007, 2007.

Beven, K. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, Hydrol. Process., 6, 279–298, doi:10.1002/hyp.3360060305, 1992.

Brown, R. G. and Hwang, P. Y. C.: Introduction to random signals and applied Kalman filtering, John Wiley & Sons, 1992.

Cabus, P.: River flow prediction through rainfall-runoff modelling with a probability-distributed model (PDM) in Flanders, Belgium, Agr. Water Manage., 95, 859–868, doi:10.1016/j.agwat.2008.02.013, 2008.

Castiglioni, S., Lombardi, L., Toth, E., Castellarin, A., and Montanari, A.: Calibration of rainfall-runoff models in ungauged basins: A regional maximum likelihood approach, Adv. Water Resour., 33, 1235–1242, doi:10.1016/j.advwatres.2010.04.009, Workshop on New Frontiers of Hydrology, Rome, Italy, JUL, 2009, 2010.

Coe, M. T., Latrubesse, E. M., Ferreira, M. E., and Amsler, M. L.: The effects of deforestation and climate variability on the streamflow of the Araguaia River, Brazil, Biogeochemistry, 105, 119–131, doi:10.1007/s10533-011-9582-2, 2011.

Duan, Q., Sorooshian, S., and Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resour. Res., 28, 1015–1031, doi:10.1029/91WR02985, 1992.

Dunne, T. and Black, R.: Partial area contributions to storm runoff in a small New-England watershed, Water Resour. Res., 6, 1296–1311, doi:10.1029/WR006i005p01296, 1970.

Gan, T. and Biftu, G.: Automatic calibration of conceptual rainfall-runoff models: Optimization algorithms, catchment conditions, and model structure, Water Resour. Res., 32, 3513–3524, doi:10.1029/95WR02195, 1996.

Gill, M. K., Kaheil, Y. H., Khalil, A., McKee, M., and Bastidas, L.: Multiobjective particle swarm optimization for parameter estimation in hydrology, Water Resour. Res., 42, W07417, doi:10.1029/2005WR004528, 2006.

Immerzeel, W. W. and Droogers, P.: Calibration of a distributed hydrological model based on satellite evapotranspiration, J. Hydrol., 349, 411–424, doi:10.1016/j.jhydrol.2007.11.017, 2008.

Jenson, S. and Domingue, J.: Extracting topographic structure from ditital elevation data for geographic information-system analysis, Photogramm. Eng. Remote Sens., 54, 1593–1600, 1988.

Kennedy, J. and Eberhart, R.: Particle swarm optimization, in: Neural Networks, Proceedings, IEEE International Conference on, 1942–1948 Vol.4, IEEE, doi:10.1109/ICNN.1995.488968, 1995.

Liu, J. and Han, D.: Indices for calibration data selection of the rainfall-runoff model, Water Resour. Res., 46, W04512, doi:10.1029/2009WR008668, 2010.

Merz, R. and Bloschl, G.: Regionalisation of catchment model parameters, J. Hydrol., 287, 95–123,

doi:10.1016/j.jhydrol.2003.09.028, 2004.

Montanari, A. and Toth, E.: Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins?, Water Resour. Res., 43, W05434, doi:10.1029/2006WR005184, 2007.

Moore, R. J.: The PDM rainfall-runoff model, Hydrol. Earth Syst. Sci., 11, 483–499, doi:10.5194/hess-11-483-2007, 2007.

Mousavi, S. J. and Shourian, M.: Adaptive sequentially space-filling metamodeling applied in optimal water quantity allocation at basin scale, Water Resour. Res., 46, W03520, doi:10.1029/2008WR007076, 2010.

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.

Oudin, L., Andreassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments, Water Resour. Res., 44, W03413, doi:10.1029/2007WR006240, 2008.

Papoulis, A.: Probability, Random Variables and Stochastic Processes, McGraw Hill, New York, 1965.

Parajka, J., Merz, R., and Blöschl, G.: A comparison of regionalisation methods for catchment model parameters, Hydrol. Earth Syst. Sci., 9, 157–171, doi:10.5194/hess-9-157-2005, 2005.

Pauwels, V. R. N. and De Lannoy, G. J. M.: Multivariate calibration of a water and energy balance model in the spectral domain, Water Resour. Res., 47, W07523, doi:10.1029/2010WR010292, 2011.

Post, D. and Jakeman, A.: Relationships between catchment attributes and hydrological response characteristics in small Australian mountain ash catchments, Hydrol. Process., 10, 877–892, doi:10.1002/(SICI)1099-1085(199606)10:6<877::AID-HYP377>3.0.CO;2-T, 1996.

Quets, J. J., De Lannoy, G. J. M., and Pauwels, V. R. N.: Comparison of spectral and time domain calibration methods for precipitation-discharge processes, Hydrol. Process., 24, 1048–1062, doi:10.1002/hyp.7546, 2010.

Reed, P., Minsker, B., and Goldberg, D.: Designing a competent simple genetic algorithm for search and optimization, Water Resour. Res., 36, 3757–3761, doi:10.1029/2000WR900231, 2000.

Scheerlinck, K., Pauwels, V. R. N., Vernieuwe, H., and De Baets, B.: Calibration of a water and energy balance model: Recursive parameter estimation versus particle swarm optimization, Water Resour. Res., 45, W10422, doi:10.1029/2009WR008051, 2009.

Seibert, J.: Regionalisation of parameters for a conceptual rainfall-runoff model, Agr. Forest Meteorol., 98-9, 279–293, doi:10.1016/S0168-1923(99)00105-7, 1999.

Shannon, C.: Communication in the presence of noise (reprinted), Proc. IEEE, 72, 1192–1201, doi:10.1109/PROC.1984.12998, 1984.

Sivapalan, M., Takeuchi, K., Franks, S., Gupta, V., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J., Mendiondo, E., O'Connell, P., Oki, T., Pomeroy, J., Schertzer, D., Uhlenbrook, S., and Zehe, E.: IAHS decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, Hydrolog. Sci. J., 48, 857–880, doi:10.1623/hysj.48.6.857.51421, 2003.

Thibos, L.: Fourier Analysis for Beginners, e-book, School of Optometry, Indiana University, 2003.

Thyer, M., Kuczera, G., and Bates, B.: Probabilistic optimization for conceptual rainfall-runoff models: A comparison of the shuffled complex evolution and simulated annealing algorithms, Water Resour. Res., 35, 767–773, doi:10.1029/1998WR900058, 1999.

Tolson, B. A., Asadzadeh, M., Maier, H. R., and Zecchin, A.: Hybrid discrete dynamically dimensioned search (HD-DDS) algorithm for water distribution system design optimization, Water Resour. Res., 45, W12416, doi:10.1029/2008WR007673, 2009.

Toth, E.: Catchment classification based on characterisation of streamflow and precipitation time series, Hydrol. Earth Syst. Sc., 17, 1149–1159, doi:10.5194/hess-17-1149-2013, 2013.

Weber, D. and Englund, E.: Evaluation and comparison of spatial interpolators, Math. Geol., 24, 381–391, doi:10.1007/BF00891270, 1992.

Whittle, P.: Estimation and information in stationary time series, Ark. Mat., 2, 423–434, doi:10.1007/BF02590998, 1953.

Winsemius, H. C., Savenije, H. H. G., Gerrits, A. M. J., Zapreeva, E. A., and Klees, R.: Comparison of two model approaches in the Zambezi river basin with regard to model reliability and identifiability, Hydrol. Earth Syst. Sci., 10, 339–352, doi:10.5194/hess-10-339-2006, 2006.

Winsemius, H. C., Schaefli, B., Montanari, A., and Savenije, H. H. G.: On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information, Water Resour. Res., 45, W12422, doi:10.1029/2009WR007706, 2009.

Zhang, Y. and Chiew, F. H. S.: Relative merits of different methods for runoff predictions in ungauged catchments, Water Resour. Res., 45, W07412, doi:10.1029/2008WR007504, 2009.