Observational uncertainty in hydrological modelling using data depth

Singh S.K.[1,*] and Dutta S.[2]

[1]National Institute of Water and Atmospheric Research, Christchurch, New Zealand,

[2]Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur, India.

*to whom all correspondence should be addressed: e-mail: Shailesh.Singh@niwa.co.nz

Abstract

For any river basin management, one needs tools to predict runoff at different time and spatial resolutions. Hydrological models are tools which account for the storage, flow of water and water balance in a watershed, which include exchanges of water and energy within the earth, atmosphere and oceans and utilise metrological data to generate flow. There are several sources of error in meteorological data, namely, through measurement at point level, interpolation, etc. When an erroneous input is passed to a model, one cannot expect an error free output from the prediction. Every prediction is associated with uncertainty. Quantification of these uncertainties is of prime importance in real world forecasting. In this study, an attempt has been made to study uncertainty associated with hydrological modelling, using the idea of data depth. To see the effect of uncertainty in rainfall on flow generation through a model, the input to a model was altered by adding an error and a different realisation was made. A Monte Carlo simulation generated a large number of hydrological model parameter sets drawn from the uniform distribution. The model was run using these parameters for each realisation of the rainfall. The parameters which are good for different realisations are more likely to be good parameters sets. For each parameter set, data depth was calculated and a likelihood was assigned to each parameter set based on the depth values. Based on this, the frequency distribution of the likelihood was analysed as well. The results show that uncertainty in hydrological modelling are multiplicative. The proposed methodology to assign prediction uncertainty is demonstrated using the 'TopNet' model for the Waipara river catchment located in the central east of the South Island, New Zealand. The results of this study will be helpful in calibration of hydrological model and in quantifying uncertainty in the prediction.

Keywords: Half-space depth, Parameter estimation, Predictive uncertainty, TopNet.

1. Introduction

Hydrological models provide simplifications and theoretical approximations of complex natural phenomena. Hence, they cannot perfectly predict what happens in natural systems. Every stage of hydrological modelling acquires some uncertainty. This uncertainty can be broadly grouped into input forecast uncertainty and hydrologic model uncertainty. The input forecast uncertainty implies uncertainty due to forcing (input) data such as precipitation, temperature and other metrological inputs to the model. The hydrologic model uncertainty comes from initial conditions, imperfect model structure and insufficient information for parameter identification, complexity of the hydrological model (Patil, 2008; Regonda *et al.*, 2013; Singh *et al.*, 2013).

In recent years quantification of uncertainty at different stages of hydrological modelling has received attention by several authors (Beven and Binley, 1992; Beven and Freer, 2001; Götzinger and Bárdossy, 2008; Wagener *et al.*, 2003; Xia *et al.*, 2004).

Despite several uncertainties, hydrological models are valuable tools for applications ranging from real time prediction of flood events to the design of policies and structures for mitigating the effects of extreme hydrologic events such as floods and droughts (Gupta *et al.*, 1999).

Most hydrological models require parameters to be estimated by calibration, so that the runoff calculated from the model provides an accurate simulation of recorded runoff from some historical period (Beven, 2004; Jakeman and Hornberger, 1993). The calibration procedure must be conducted very carefully to maximize the reliability of the model for the given application (Gupta *et al.*, 1999). It is well known that the physics governing the path of a drop of water through a catchment to the stream involves complex relationships, but evidence indicates that the information content in a rainfall-runoff record is sufficient to support models of only very limited complexity (Jakeman and Hornberger, 1993). The quality and quantity of observed rainfall and flow data therefore has a great influence on the identification of hydrological model parameters (Andréassian *et al.*, 2001; Singh and Bárdossy, 2012). Ideally, calibration data should be fully representative of the range of flow conditions experienced by the watersheds (Sorooshian *et al.*, 1983).

Uncertainty in model prediction is a cumulative effect of input and hydrological model uncertainty. Each uncertainty has cascading effects. Out of this, observation uncertainty is recognized as a serious issue undermining the reliability of hydrological studies (Montanari and Di Baldassarre, 2013).

In this study, an attempt has been made to study the effect of observational uncertainty on parameter estimation of hydrological models and provide a new way of defining the uncertainty using data depth. We have tried to look into the kind of uncertainty invoked by observational errors in hydrological variables. The input error is very basic and the most important error, as one cannot expect a good result if the input itself is incorrect. These input errors can be mostly due to measurement error, or from methods used in recording the data. The observational error is the first error in modelling and it can multiply further in the modelling. So, observational errors can introduce significant amounts of uncertainty in hydrological modelling. There have been numerous studies (Ibbitt, 1972; Oudin *et al.*, 2006; Singh and Dutta, 2016) to quantify such uncertainties, but there is still need to develop some tools to quantify and reduce such uncertainties.

The objective of this study is to use data depth (Liu, Parelius and Singh, 1999) to yield a likelihood which will be used in quantifying uncertainty in hydrological forecasting arising from input errors. This study was carried out in the Waipara catchment located in the South Island of New Zealand using semi-distributed physically based on the hydrological model TopNet (Bandaragoda *et al.*, 2004)

2.   Methodology

In this section, we first describe the study area and the hydrological model used in this study. Data depth function and its application to quantify uncertainty in hydrological prediction due to input error is described later in this section.

2.1. Study Area

The present study was carried out at the Waipara experimental catchment which is located on the South Island of New Zealand, in the headwaters of the Waipara River (see Fig. 1). McMillan and Srinivasan (2015) give a detailed description of the basin and its monitoring network. In this study, the model was set up in the Langs Gully.   The catchment area of the Langs Gully is 0.7km². The elevation ranges from 500 to 723 m a s l. The annual precipitation ranges from 500 to 1100mm/yr. A rather broad range and hard to believe over a long time.   It contains surface slopes of 0.22-34 degrees with a mean slope of 17 degrees. The soil is gravelly sandy loam, with depth ranging from 0.25 to 1.5m and an average depth of 0.5m. The primary vegetation includes grasses and exotic forest.
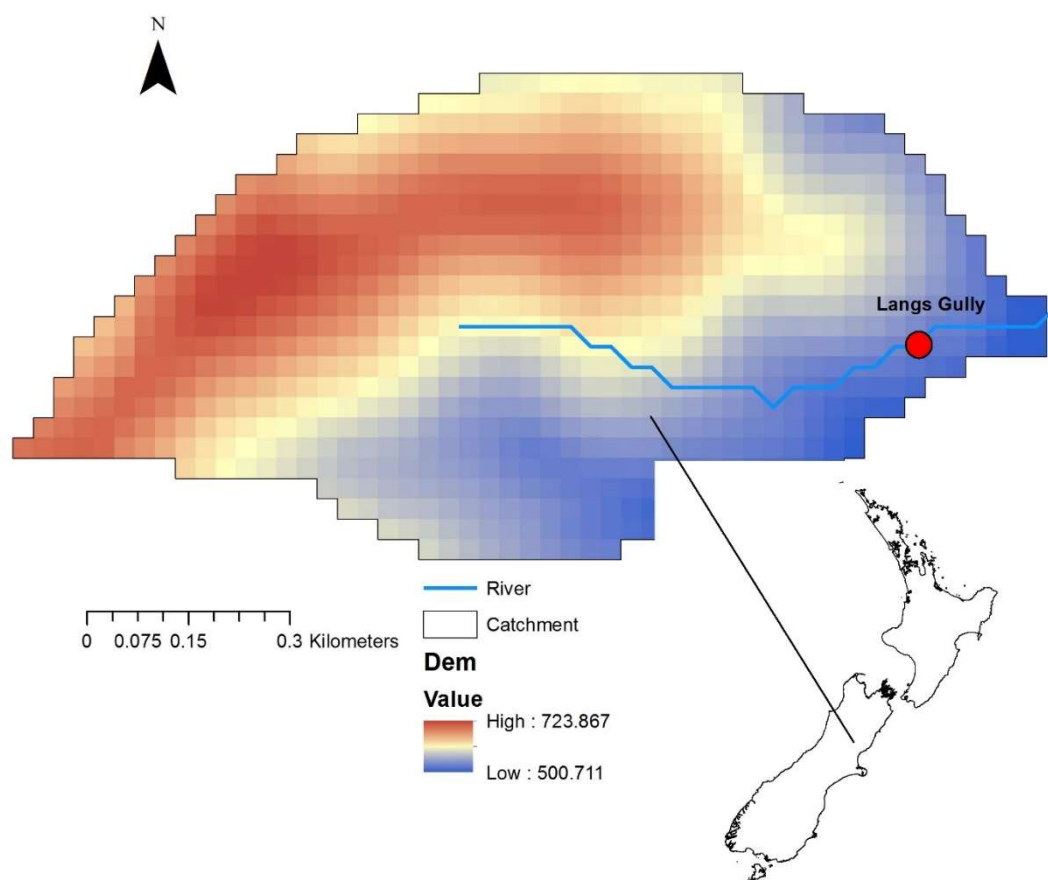

Figure 1. Location map with DEM for the Waipara experimental basin on the South Island of New Zealand, show the area modelled by TopNet

2.2 Model

TopNet is a semi-distributed hydrological model which simulates catchment water balance and river flow. It was developed using the TOPMODEL (Beven and Kirkby, 1979) concepts for parameterization of soil moisture deficit, using a topographic index to model the dynamics of variable source areas contributing to saturation excess runoff (Bandaragoda *et al.*, 2004; Beven and Kirkby, 1979). TopNet models a catchment as a collection of sub-watersheds, linked with a branched river network (Clark *et al.*, 2008). The flow is routed through the river network using kinematic waves using the shock-fitting technique of Goring (1994). Modelled streamflow is generated in three ways:

- rain falls on a location where soil water storage equals its capacity (Saturation excess runoff),
- rain rate exceeds infiltration rate (Hortonian runoff), and
- saturated zone discharge into stream (base flow).

TopNet assumes that available soil water storage can vary within a sub-watershed because of topographic effects - valley bottoms and flat places are wetter than the ridges. TopNet provides a prediction of flow in each modelled reach within a catchment (Bandaragoda *et al.*, 2004; Clark *et al.*, 2008; Ibbitt *et al.*, 2000). The model inputs are rainfall and temperature time series (e.g., at hourly time steps with rain from one or more locations), relative humidity, solar radiation, maps of elevation, vegetation type, soil type and rainfall patterns. These map data are used with tables of model parameters for each soil and vegetation type, to produce initial estimates of the model parameters. A schematic representation of the model is given in Figure 2. TopNet has 31 parameters to define the hydrological processes of a catchment. Whenever possible, parameter values are determined from physical catchment properties; however, 15 parameters (see Table 1) typically require calibration. During calibration, TopNet model uses a spatially constant multiplier for each parameter, to adjust the parameters while retaining the relative spatial pattern obtained from the soil and vegetation data (Bandaragoda *et al.*, 2004). This procedure is adopted in order to reduce the dimensionality of the calibration problem.

Table 1. TopNet model parameters which need calibration, description and allowed range for the parameter multiplier

Parameter	Description	Initial Min	Initial Max
topmodf	TOPMODEL f parameter (m^{-1})	0.001	2
hydcon0	Saturated hydraulic conductivity (ms^{-1})	0.01	9999
swater1	Drainable water(m)	0.05	20
swater2	Plant-available water(m)	0.05	20
Dthetat	Soil water content(m)	0.1	10
Overvel	Overland flow velocity (ms^{-1})	0.1	10
Gucatch	Gauge under-catch for snowfall(-)	0.5	1.5
th_accm	Threshold for snow accumulation(K)	272.16	275.16
th_melt	Threshold for snow melt (K)	272.16	275.16
snowddf	Mean degree-day factor for snow melt (mm K-1 day-1 = kg m-2 K-1 day-1)	0.1	7.5
Minddfd	Minimum degree-day-factor day (julian day: 1 to 366)	1	366
maxddfd	Maximum degree-day-factor day (julian day: 1 to 366)	1	366
snowamp	Seasonal amplitude of degree-day factor for snow melt (mm K-1 day-1 = kg m-2 K-1 day-1)	0	7.5
cv_snow	Coefficient of variation in sub-grid SWE(-)	0.5	1.5
r_man_n	Manning's n(-)	0.1	10

Figure 2. Systematic Representation of TopNet model structure

2.3. Data depth function

Data depth is a quantitative measurement of how central (or deep) a point is with respect to a multivariate data cloud, or a probability distribution. It gives us a central outward ordering of multivariate data points and yields simple new ways to quantify complex multivariate features of the underlying multivariate distribution (Liu *et al.*, 1999). The first depth function was introduced by Tukey (1975) to identify the centre (a generalized median) of a multivariate dataset. Generalizations of this concept have been defined by (Barnett, 1976; Liu, 1990; Liu *et al.*, 2006; Zuo and Serfling, 2000). Several types of data depth functions have been developed, e.g., half-space, projection and Mahalanobis depth functions. The half-space depth function (Tukey, 1975) was used in this study because it satisfies all the properties of a depth function, and it leads to robustness.

The idea is perhaps best conveyed by first focusing on the univariate case. Given some point x and a finite collection of points X. Consider any partitioning of all real numbers into two components: those values below x and those above. All values less than or equal to x form a closed halfspace. In a similar manner, all points greater than or equal to x also represent a closed halfspace. Tukey's halfspace depth associated with the value x is intended to reflect how deeply x is nested within the data cloud X. It is estimated to be the smaller of two proportions, namely, the proportion of observed values less than or equal to x, and the proportion greater than or equal to x. Clearly, the maximum possible halfspace depth (which is close to 0.5) in a sample of observations corresponds to the sample median. While the minimum possible halfspace depth (which is 0) corresponds to any observation which lies outside the interval with the lower (respectively, the higher) end point as the minimum (respectively, the maximum) value of the observations in the set X.

Now, we generalize to the bivariate case, d = 2. For any line, the points on or above this line form a closed halfspace, as do the points on or below the line. For d = 3, any plane forms two closed halfspaces: those points on or above the plane, and the points on or below the plane. The notion of a halfspace generalizes in an obvious way for any dimension d.

For the general d-variate case, consider any point p and H be any closed halfspace containing the point p. Let P(H) be the empirical probability associated with H. Then, the halfspace depth of the point p is the smallest value of P(H) among all halfspaces H containing p. More formally, Tukey's halfspace depth is

D(p) = inf{P(H): H is a closed halfspace containing p}.

For p > 1, halfspace depth can be defined instead as the least depth of any one-dimensional projection of the data (Donoho and Gasko, 1992). To elaborate, let us consider any point p and any d-dimensional (column) vector u having unit Euclidean norm, i.e., $||u|| = u_1^2 + \ldots + u_p^2 = 1$. Then, a one-dimensional projection of p is defined as the inner product $\langle u, p \rangle$. For any projection, meaning any choice for u, depth is defined as in the univariate case

above. In the d-variate case, the depth of a point is defined to be its minimum depth among all possible projections u. From a practical point of view, this does not immediately yield a viable algorithm for computing halfspace depth based on a sample of n observations, but it suggests an approximation that has been found to be relatively effective.

Formally, the half-space depth of a point p with respect to the finite set X in the d-dimensional space is defined as the minimum number of points of the set X lying on one side of a hyperplane through the point p. The minimum is calculated over all possible hyperplanes. Formally, the half-space depth of the point p with respect to set X is defined as:

$$D_x(p) = \inf\left(\min(|\{x \in X \mid \langle u, x-p \rangle > 0\}|, \quad (1)\right.$$
$$\left. |\{x \in X \mid \langle u, x-p \rangle < 0\}|)\right)$$

Here $\langle x, y \rangle$ is the scalar product of the d-dimensional vectors, u is an arbitrary unit vector in the d-dimensional space representing the normal vector of a selected hyperplane and |S| is the cardinality of the set S. If the point p is outside the convex hull of X, then its half-space depth is 0. The convex hull of a set of points S is the smallest convex set (e.g., convex polygon in 2D) which encloses S (Fig. 3). Points on and near the boundary of this convex hull have low depth, while points deep inside have high depth. One advantage of this depth function is that it is invariant to affine transformations of the space. This means that the different ranges of the variables have no influence on the calculated depth.
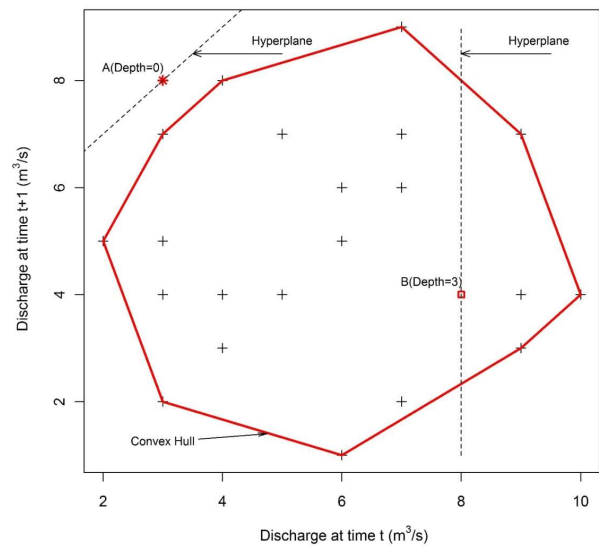

Figure 3. Example of a convex hull using hydrological variables in 2-D

2.4. Quantify the uncertainty using data depth function

Let us assume that the accuracy of the measured precipitation P_M is A% due to observational errors. Thus, the real but unknown precipitation P_E can be written as:

$$P_E(t) = P_M(t) + \xi(t) \quad (2)$$

where ξ(t) is a random error. This random error is due to uncertainties of the aerial distribution of precipitation, recording of precipitation, instrument error etc. Here, we assume that the error follows a normal distribution N(0, A/100). Using Eq. (2), and A=20%, 100 perturbed precipitation series were generated in this study.
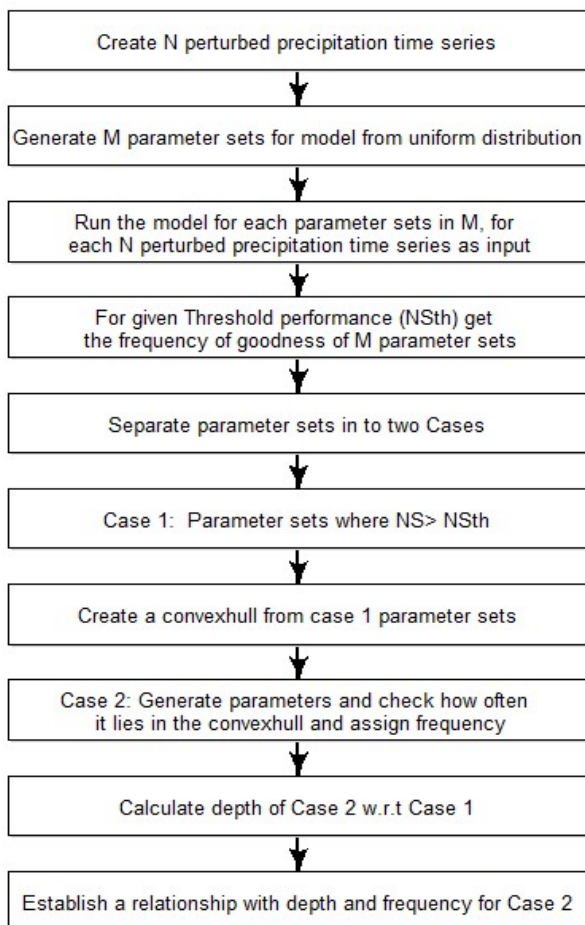
Considering upper and lower range of the model parameters (Table 1), 10000 parameters sets from uniform distribution were generated for the 'TopNet' hydrological model. The model was run using these parameters with the 100 perturbed precipitation time series as input. The likelihood is assigned to each parameter based on a threshold Nash-Sutcliffe coefficient greater than 0.6. A convex hull of those parameters (which have higher values of the likelihood) was constructed. A new parameter set was generated, and it was checked if it lies within, or outside the convex hull. If it is within this hull, a likelihood value was assigned using logistic regression. Depth of the new parameters sets (which are inside the convex hull) was calculated with respect to the space of the convex hull. A flow chart of the methodology adopted in this study is given in Figure 4.

3.   Results and Discussion

3.1 Effect of input error on the output of the model

The TopNet hydrological model was setup for the Langs Gully of Waipara experimental catchment using Strahler order 1 catchments. The model was calibrated using Robust Parameter estimation (ROPE) algorithm (Bárdossy and Singh, 2008). Unlike other global optimisation algorithm, ROPE try to find an optimal parameter space instead of a best parameter set. Figure 5 shows the model and the observed hydrograph. The model was able to reproduce the observed hydrograph dynamics very well. But it reproduces the volume slightly less than observed.This can be due to several factors like error in precipitation, model structure, and error in flow measurement.

Figure 6 illustrates the perturbed precipitation series, and the corresponding hydrograph using an optimal parameter set obtained during the calibration of the hydrological model.
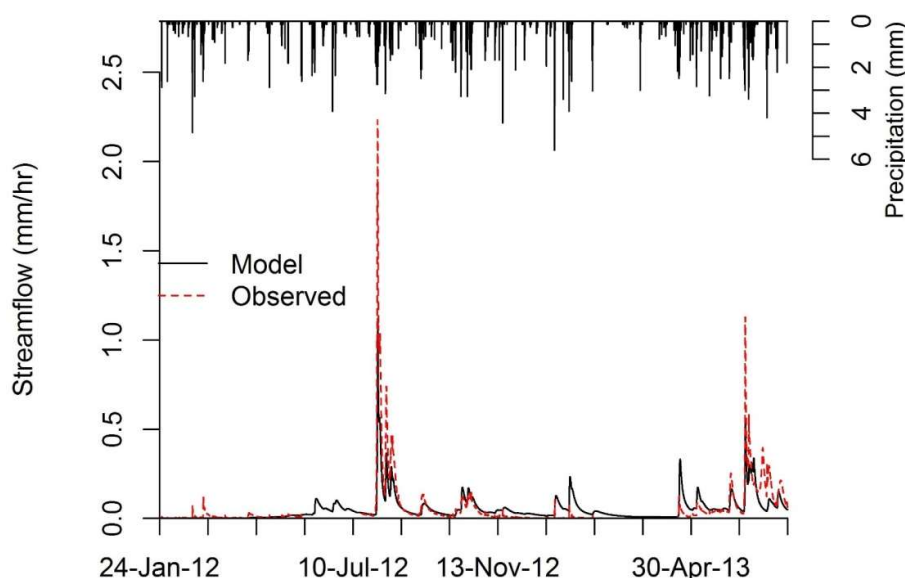

Figure 4. Flow chart of the methodology adopted in this study

Figure 5. Observed and model hydrograph, along with precipitation during calibration time period

The following equations were used to quantify the effect of perturbation error at each time step on the precipitation to the total precipitation over the study time period and the total flow as follows:

$$P_{precipitation} = \frac{\sum P - \sum P_{perturbed}}{\sum P} * 100 \tag{3}$$

$$P_{discharge} = \frac{\sum Q_{obs} - \sum Q_{model}}{\sum Q_{obs}} * 100 \tag{4}$$

At the first step, we calculate the percentage error in model discharge if there is no perturbation. In this case, the model had an error of 3.7%. This error can come from many sources, e.g., model structure error, parameterization error, or error in the precipitation series. We assume this is a basic error for further comparison. Using Equations 3 we calculated the percentage error in the precipitation time series. The error in model discharge with respect to observed discharge due to perturbed precipitation was calculated using Equation 4. An example is given in Figure 7. We can see that +5% error in total precipitation can lead to +13 to -5% error in total discharge. Even if we consider the median discharge error of +3.7% (low bias) which may come from the model itself, the range of percentage errors from input precipitation to discharge is almost double. It was also found that the Gaussian error in the input leads to a Cauchy distribution in the output error. The uncertainty started with known error in precipitation and includes uncertainty associated with the model. This illustrates that uncertainty is not additive, but it is multiplicative. Error in the input leads to an error in fitting the model parameters. This can lead to unrealistic catchment process representation, wrong model structure and ultimately error in the output. Trivial statement.
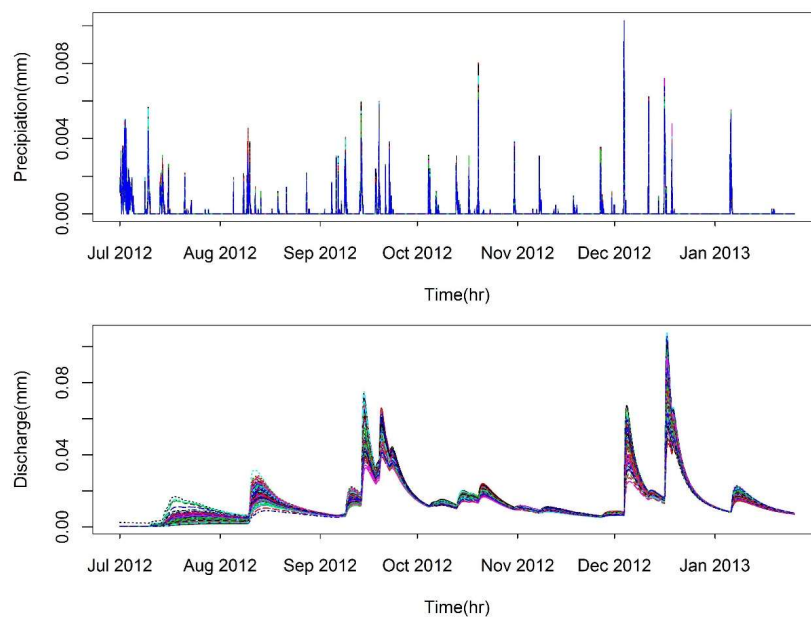

Figure 6. Illustration of the perturbed precipitation series (above) and corresponding hydrograph (below) using an optimal parameter set
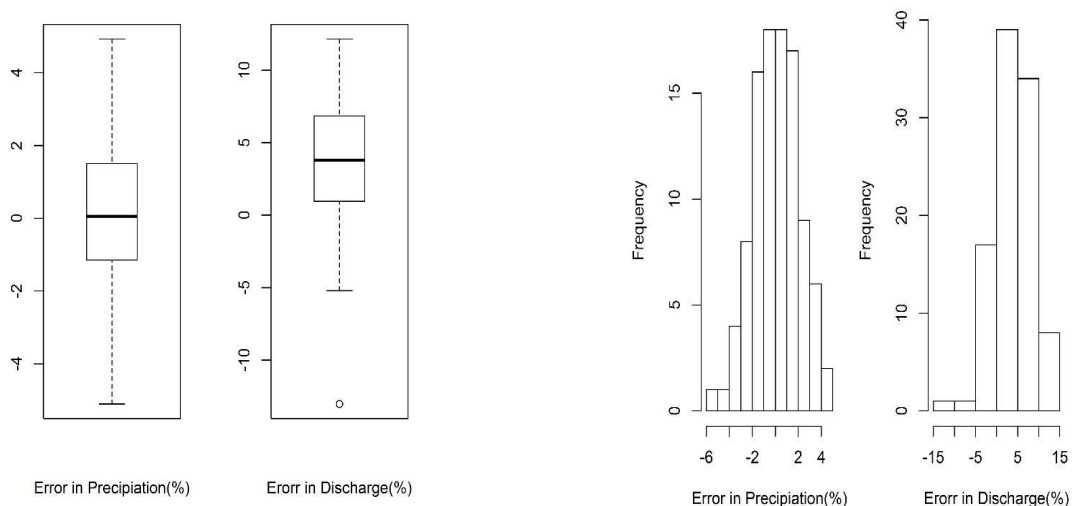

Figure 7. Illustration of translation percentage error from precipitation to discharge

3.2. Data depth function to quantify the uncertainty

Following the methodology described in Section 2, the calibrated TopNet hydrological model was run for 10000 parameters set with 100 perturbed rainfall time series, and hence we obtain 10000*100 simulations. Nash-Sutcliffe (NS) coefficient (Nash and Sutcliffe, 1970) of 0.6 was used as threshold to get the frequency of goodness for each parameter set. Using the NS threshold, the parameter set were selected and called as Case 1. The Case 1 was further utilised to create a convex hull. From the given range of model parameters, randomly parameters were generated from uniform distribution and the frequency of the parameter set that lies within the convex hull of Case 1 was evaluated. The generated parameter sets were termed as Case 2. Data depth of Case 2 with respect to Case 1 was calculated, and correlated with the frequencies of goodness of fit in 100 ensembles. Figure 8 shows the relationship between the frequency and the depth values. There is no clear relationship between the two but we can see that higher depth corresponds to high values of the frequency. This is because deeper parameters perform better even if there is a slight error in the input. Hence, depth was classified in a different bin and a contour of convex hull was constructed. The inner convex hull parameters which have higher depth correspond to lower uncertainty. However, if the parameters for a higher convex hull have been used, then uncertainty will be higher. An example is given in Figure 9. We can see that 95% and 5% discharge bound is narrow for higher depth convex hull as compared to the lower depth outer convex hull.
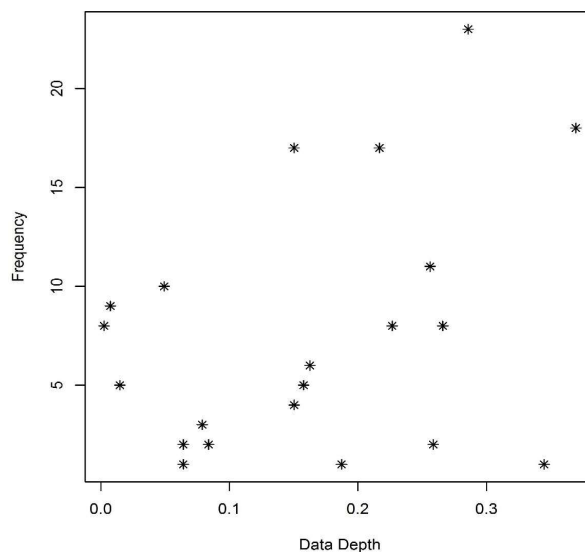
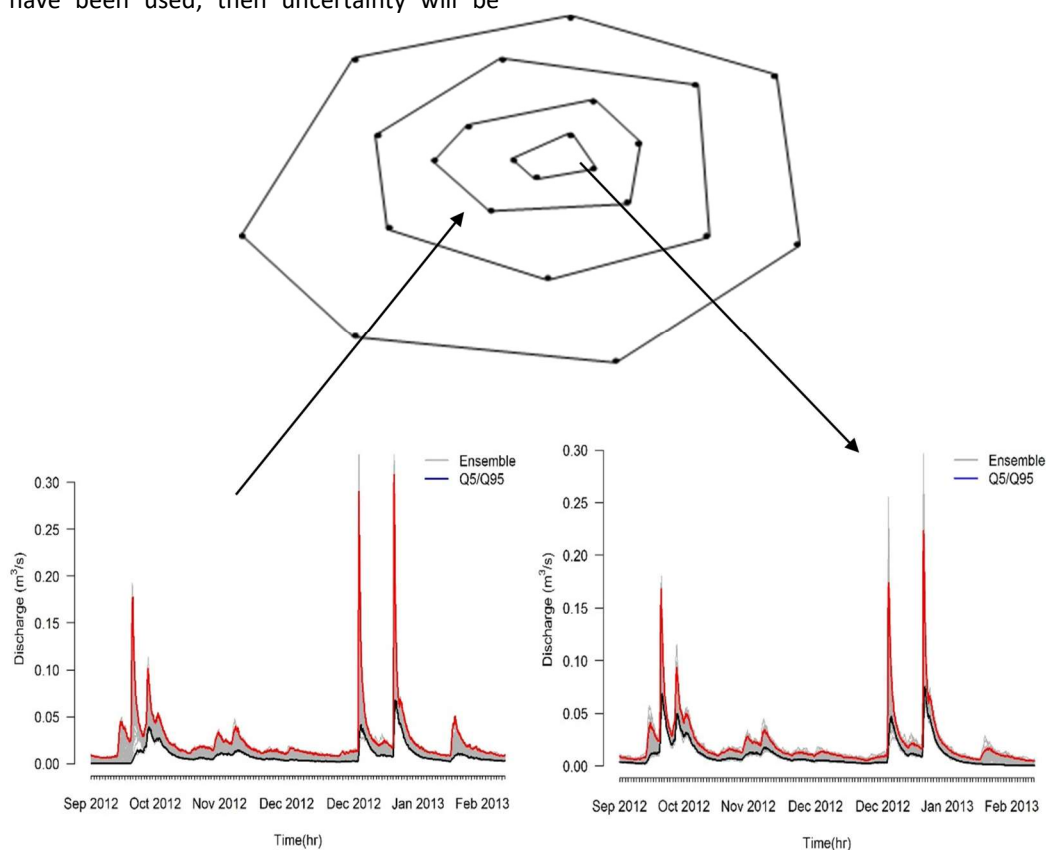

Figure 8. Data Depth vs Frequency

Figure 9. Example of uncertainty sample from different location in depth contour

In a previous study, Bárdossy and Singh (2008), showed that parameter sets having higher depths contain robust parameters which are good in transferability (transferring parameters for other time periods), and are less sensitive (small change in parameters set does not significantly change the performance of the model). This phenomena is attributed to the fact that the parameter set is trying to compensate all types of uncertainty. Hence, if we have a very compact and structured parameter set it will have higher depth and it will have higher model performance, and less predictive error which leads to less uncertainty. The results from this study demonstrate that depth of parameter sets can be used as a surrogate to define the uncertainty in our prediction.

4. Conclusions

The present study analyses the effect of the input error on the output of the hydrological model. It has been found that the observed error can give very different parameter sets. The results indicate that the input to output uncertainty is multiplicative. The other results are based on an application of data depth function and demonstrate that the depth of parameters set can be used as a surrogate to define uncertainty in our prediction. Hence, we can conclude that geometrical properties (namely, depth) of parameter sets can be used to define uncertainty as well as location of parameter sets in parameter space can tell us about the uncertainty associated with that prediction. The results of this study will be helpful in robust calibration of hydrological models, and in quantifying the uncertainty in predictions. A further study with other model and area may require to generalise the finding of this study.

References

Andréassian V., Perrin C., Michel C., Usart-Sanchez I. and Lavabre J. (2001), Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models, *Journal of Hydrology*, **250**(1–4), 206-223.

Bandaragoda C., Tarboton D.G. and Woods R. (2004), Application of TOPNET in the distributed model intercomparison project, *Journal of Hydrology*, **298**(1-4), 178-201.

Bárdossy A. and Singh S.K. (2008), Robust estimation of hydrological model parameters, *Hydrology and Earth System Sciences*, **12,** 1273-1283.

Barnett V. (1976), The ordering of multivariate data (with discussion), *Journal of Royal Statistical Society A*, **139**, 318-354.

Beven K. and Binley A. (1992), The future of distributed models: model calibration and uncertainty prediction, *Hydrological Processes*, **6**(3), 279-298.

Beven K. and Freer J. (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *Journal of hydrology*, **249**(1), 11-29.

Beven K.J., 2004. Rainfall-runoff modelling: the primer. Wiley, Chichester.

Clark M.P., Rupp D.E., Woods R.A., Zheng X., Ibbitt R.P., Slater A.G., Schmidt J. and Uddstrom M.J. (2008), Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model, *Advances in Water Resources*, **31**(10), 1309-1324.

Götzinger J. and Bárdossy A. (2008), Generic error model for calibration and uncertainty estimation of hydrological models, *Water Resources Research*, **44**(12), W00B07.

Gupta H.V., Sorooshian S. and Yapo P.O. (1999), Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration, *Journal of Hydrologic Engineering*, **4**(2), 135-143.

Ibbitt R.P. (1972), Effects of random data errors on the parameter values for a conceptual model, *Water Resources Research*, **8**(1), 70-78.

Ibbitt R.P., Henderson R.D., Copeland J. and Wratt D.S. (2000), Simulating mountain runoff with meso-scale weather model rainfall estimates: a New Zealand experience, *Journal of Hydrology*, **239**(1–4), 19-32.

Jakeman A.J., Hornberger G.M. (1993), How much complexity is warranted in a rainfall-runoff model?, *Water Resources Research*, **29**(8), 2637-2650.

Liu R. (1990), On a notion of data depth based on random simplices, *The Annals of Statistics*, **18,** 405-414.

Liu R.Y., Parelius J.M. and Singh K. (1999), Multivariate analysis by data depth: descriptive statistics, graphics and inference (with discussion and a rejoinder by Liu and Singh), *The Annals of Statistics*, **27**(3), 783-858.

Liu R.Y., Serfling R.J. and Souvaine D.L. (2006), Data depth: robust multivariate analysis, computational geometry, and applications, 72. American Mathematical Society.

McMillan H. and Srinivasan M. (2015), Characteristics and controls of variability in soil moisture and groundwater in a headwater catchment, *Hydrology and Earth System Sciences*, **19**(4), 1767-1786.

Montanari A. and Di Baldassarre G. (2013), Data errors and hydrological modelling: The role of model structure to propagate observation uncertainty, *Advances in Water Resources*, **51**, 498-504.

Nash J.E. and Sutcliffe J. (1970), River flow forecasting through conceptual models part I—A discussion of principles, *Journal of Hydrology*, **10**(3), 282-290.

Oudin L., Perrin C., Mathevet T., Andréassian V. and Michel C. (2006), Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models, *Journal of Hydrology*, **320**(1), 62-83.

Patil S.R., 2008. Regionalization of an event based Nash cascade model for flood predictions in ungauged basins, PhD Thesis Nr. 175, University of Stuttgart, Germany, Stuttgart.

Regonda S.K., Seo D.-J., Lawrence B., Brown J.D. and Demargne J. (2013), Short-term ensemble streamflow forecasting using operationally-produced single-valued streamflow forecasts – A Hydrologic Model Output Statistics (HMOS) approach, *Journal of Hydrology*, **497,** 80-96.

Singh S.K. and Bárdossy A. (2012), Calibration of hydrological models on hydrologically unusual events, *Advance in Water Resources*, **38**(0), 81-91.

Singh S.K., McMillan H. and Bárdossy A. (2013), Use of the data depth function to differentiate between case of interpolation

and extrapolation in hydrological model prediction, *Journal of Hydrology*, **477,** 213-228.

Singh S.K. and Dutta S. (2016), Effect of observational uncertainty on hydrological modelling, In Proc. 56th New Zealand Hydrological Society's The Water and Infrastructure and Environmental conference, 28 Nov – 2 Dec 2016, Queenstown, New Zealand.

Sorooshian S., Gupta V.K. and Fulton J.L. (1983), Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: Influence of calibration data variability and length on model credibility, *Water Resources Research*, **19**(1), 251-259.

Tukey, J., 1975. Mathematics and picturing data. In Proceedings of the 17th International Congress of Mathematics, 2: 523-531.

Wagener T., McIntyre N., Lees M.J., Wheater H.S. and Gupta H.V. (2003), Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis, *Hydrological Processes*, **17**(2), 455-476.

Xia Y., Yang Z.L., Jackson C., Stoffa P.L. and Sen M.K. (2004), Impacts of data length on optimal parameter and uncertainty estimation of a land surface model, *J. Geophys. Res*, **109**: D07101.

Zuo Y. and Serfling R. (2000), General notions of statistical depth function, *The Annals of Statistics*, **28**(2), 461-482.