Data Duration Options of a Practicing Water Manager when Developing a Model for Sustainable Water Resources Development

M.A.D.N. Ariyasena and N.T.S. Wijesekera

Abstract: Water scarcity emphasizes the importance of water resources planning and management. Monthly hydrologic models are commonly used because in practice, monthly time step is the choice for planning activities. Reliability of the model simulations depends upon the nature of the data. Data quality is important and data should be representative of the watershed processes. The information contained in the data and careful extraction is also important.

A detailed review of available guidelines and research findings to identify the best data duration option available for water managers to calibrate and verify a monthly streamflow estimation model for sustainable water resources management was carried out.

The data length selection for monthly water resources models varies between 12 and 780 months and the reasons for selection is mostly individual preference. This study exhibits a qualitative evaluation using five key criteria. Data length use in research and guidelines indicates that 20 to 40 year long data sets are the most favoured for monthly model applications.

A total data length of range 20-30 years or greater should be utilized for monthly water resources model development. It may be prudent to select at least a 20-year data period for model calibration based on the analysis.

Keywords: Data, Data Length, Water Balance Model, Data Quality

1. Introduction

1.1. Modelling Water Resources

Water has become a very scarce resource because of the effects of population growth, industrialization, urbanization and climate change. Hence the gap between the demand and supply of water keeps widening [1]-[3]. The increasing demand for the scarce water resource requires systematic, meaningful and sustainable water resources management and planning. A hydrologic model is the common tool available for water managers to execute this task in a meaningful manner. Hydrologic models vary from simple regression type to detailed physics-based process representations. Irrespective of the type, these models require sufficiently long time series input data such as rainfall and evaporation. Many literature and text book recommendations on the subject of determining model parameters by means of model calibration and verification clearly mention that characteristics and length of data must be evaluated for parameter reliability [4]-[11].

Monthly time scale is the most commonly utilized temporal resolution for practical water resources planning and management [12], [13].

Therefore, most hydrologic mathematical models attempting to fulfil this purpose apply monthly input data to calculate monthly outputs. These monthly models vary between regression and physics-based process type representations. Among the popular types are the monthly water balance models which rationalize their model computations by using principle of mass conservation and other relationships which are closer to known physics [12]-[18]. Monthly water balance models are popular for water resources management, reservoir simulation, drought assessment, long-term drought forecasting, climate change impact evaluation etc. [12]-[14], [16].

Eng. M.A.D.N. Ariyasena, AMIE (Sri Lanka), M.Sc, B.Sc. Eng. (Hons) (Moratuwa), Earth Resources Engineer, Irrigation Department, Sri Lanka, Email: dilanka.ariyasena@gmail.com, <https://orcid.org/0000-0002-9738-4287>

Eng. (Prof.) N.T.S. Wijesekera, C. Eng., FIE (Sri Lanka), MICE (London), B.Sc. Eng. (Hons) (Sri Lanka), P.G.Dip.Hyd Structures (Moratuwa), M.Eng. (Tokyo), D. Eng. (Tokyo), Senior Professor in Civil Engineering, University of Moratuwa, Sri Lanka, Email: sohan@civil.mrt.ac.lk, <http://orcid.org/0000-0003-0964-4331>

This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

1.2. Data Length and Modelling

Recommendations on data length for monthly water balance model development become the first priority of a watershed manager who has completed the model selection. Data length comprises two sections. One is the calibration data length and the other is the verification data length. Therefore, the total data length is the addition of calibration and verification data lengths. There is research that had been conducted on the methods of splitting the dataset for model calibration and verification [19]. Zheng et al.[20] in a related study had considered the splitting of the total dataset as 80% for calibration and 20% for verification. Guidelines on mathematical modelling indicate that a comprehensive data set would include long-term streamflow measurements and rainfall and PET data collected at one or more locations within the catchment, along with land use coverage, vegetation cover and impervious area information, including changes over time [21]. WMO guideline [22] states that daily or even monthly river flow data may be sufficient for assessing regional water security; multi-decadal records of continuous rainfall are necessary for reliable frequency analysis; and a few years of data may be sufficient to calibrate a groundwater model. Recommendations on Hydrologic Practices [23], while indicating the practice of British Meteorological Office for the computation of average annual rainfall that considers a 30 year minimum duration, state that flow records of 20-30 years are required to provide the streamflow pattern representative of the flow regime. For continuous modelling, an observed input containing sufficient events to calibrate all parameters has been recommended [8]. Literature on the selection of data duration for monthly models does not indicate firm recommendations [24]–[26]. Using a daily mathematical model and variety of data lengths, Gupta and Sorooshian [27] indicated that a data length of 500 to 1000 points would achieve acceptable precision of optimized parameters and hence 3-4 years of daily data would be sufficient for model calibration. Interpretation of this result would mean that for monthly modelling a dataset of over 40 years would be required. Subsequently, a study by Yapo et al.[28] recommends a 8-year dataset for calibration and then states that the wettest period of the dataset would identify improved parameters. Loucks and van Beek [29], discussing sensitivity and uncertainty of model parameters, indicate that longer datasets may not lead to better results. This raises the

question whether the number of data points or whether the hydrological characteristics of a watershed are important when selecting a dataset for modelling. On the other hand, there are occasions that cite different data periods are suitable to identify different parameters [30]. Data availability and interpolation without sufficient data also creates a significant impact on parameter estimation [31].

1.3. Problem in Practice

Though there is a desire to use long data lengths, non-availability of data often compels modellers to compromise on the data length in order to fulfil the number of gauging stations required for the model to be realistic. Boughton [17], in his work, cites many other data issues influencing the decision on data length that is used for modelling. It appears that in most modelling work, all available data are used without concerning resource requirements [7]. The present practice of monthly water balance models also does not demonstrate the use of significantly long datasets for model calibration. There is literature indicating the use of 2-5 years of monthly data for model calibration [32]–[36]. A comparative study of 203 watersheds by Chang & Chen [37] had used a 10-year dataset for calibration and verification of monthly watershed models.

In this backdrop, the selection of data length for monthly water resources modelling gives rise to many difficulties. Wide variety of opinions in the available guidance material, constraints due to missing data in datasets, poor quality of data, and constraints due to data accessibility are some of the critical issues faced by practicing watershed managers. Hence, there is a clear necessity to comparatively evaluate available material to identify the best data-length-option when attempting to use an off the shelf model field level operation. Accordingly, the main objective of the present work is to evaluate available guidelines and research findings to identify the best data duration option available for water managers to calibrate and verify a monthly streamflow estimation model for sustainable water resources management.

2. Publications for a Practitioner

2.1. General

A reviewed literature search was conducted to identify the publications available for a practicing water manager to select the appropriate data length for a water resources

modelling application. Guidelines and research publications with direct and indirect guidance on data lengths were summarized for a comprehensive review. Direct publications are those which clearly recommended data length for a particular water resource management purpose. Indirect publications are the reported reviewed publications which convey the use of a data length for a specific water resources application. The investigation included 63 publications which contained 275 and 138 monthly and daily water management applications providing assistance to select a data length for hydrologic modelling.

2.2. Water Resources Models

Length of a dataset for rational water resources modelling is dependent upon the temporal resolution of data used for model computations. Therefore, recommendations on data length depend on the type of model, parameter characteristics and the number of parameters. There are several key factors about models and model parameters that a practitioner needs to be aware when investigating the appropriate model application for a particular application.

Water resources modelling is mostly for planning and management of water as a resource. It is common to carry out these activities using results presented in monthly time scale. This is done in two ways. One is to perform computations and modelling in the daily time scale and then aggregate the results to monthly resolution. The other is to use monthly scale conceptualizations to model and generate monthly and coarser resolution results. This could be noted in the number of literature that was available for review. Daily and monthly models have their own advantages and disadvantages. In this work the length of data used for monthly and daily applications were evaluated to compare the number of data points used for calibration and verification of each model type. Also, daily models provide guidance on parameters to be looked at during applications. The present work concentrated on the assessments to capture the data length selection requirements for monthly water resources modelling applications.

The adequacy of monthly streamflow estimations for planning and management of water resources is common knowledge [16], [23], [38], [39]. Monthly continuous streamflow models play a major role in watershed

management due to many reasons such as: 1) water resources planning and management practice is on monthly or longer time scales; 2) climate change effects on vegetation can be easily specified on a monthly or seasonal scale, daily or hourly; 3) large space scale seeking only averages of watershed characteristics can be selected; and 4) monthly hydro-climatological data are most readily available [40]. Apart from the above, monthly water balance models have specific advantages, such as simple mathematical representations which are closer to known physics, sufficiency of either precipitation only or precipitation, temperature and evaporation as inputs and having a lesser number of parameters compared to finer time resolution models [41]–[44].

Monthly water balance models are used to forecast the monthly streamflow or seasonal yield of a catchment. A monthly hydrograph smoothens out the streamflow fluctuations visible at hourly or daily temporal resolutions. The variations in finer resolution measurements are due to variations of watershed resistance to runoff and storage characteristics. Monthly models provide cumulative effects of actual phenomena aggregated over calendar months. Due to the aggregation of finer variations the monthly models enable simplifications when handling flow and storage. Therefore, monthly models are comparatively simpler and have a smaller number of parameters when compared with daily and hourly models [12].

There are many monthly water balance models in use. They are ‘abcd Model’ [14], ‘PE Model’ [45], ‘Pitman Model and Stanford Model’ [46], TOPMODEL and Xianjiang Model’ [47], ‘HYDROLOG and HRCYCL Model’ [48], ‘The Catchment Land Surface Model (CLSM)’ [49], ‘Tank Model’ [6], ‘Jazim Model’ [50], ‘ReNuMa Model’ [35], ‘WatBal Model’ [33] etc.

Monthly water balance models have a long history dating back to 1940s. Thornthwaite and Mather [51] and Alley [14] developed a two-parameter monthly water balance model (T Model). Palmer [52] developed a water balance model (P Model) similar to Thornthwaite and Mather [51], dividing soil moisture into two layers. A four parameter ABCD water balance model combined with soil moisture storage was developed by Thomas [53]. Mimikou et al. [38] and Vandewiele et al. [16] describe several other monthly water balance models developed

for Water Resources Planning and Management.

This makes it clear that the available models, their merits and demerits must be carefully evaluated prior to selecting a suitable monthly water resources model and an appropriate data length.

2.3. Parameter Estimation

Parameter estimation is the selection of values for the parameters, so that the model will match the watershed system as closely as possible [42]. Model calibration is the process which determines the model parameters based on the available data and prior knowledge [54]. Among many types of models ranging from black box to physics based, there are model parameters which are either physically interpretable or non-interpretable [43]. A desired characteristic of a particular hydrologic model is the capability to simulate its target hydrologic processes under the hydrologic conditions expected to be experienced by a watershed. Therefore, model parameters estimated should be unique, realistic and independent of the data being used for calibration. Needs and characteristics of model parameters have been discussed by many. It is known that if the models are based on basic principles of physics (mass and energy conservation), then estimation of model parameters is straightforward [9]. However, watershed heterogeneity, such as non-uniformity of vegetation, slopes, existence of macro pores, together with unresolved spatial and temporal variability of meteorological variables, limits the applicability of “known physics”-based models either to laboratories or to well observed small experimental catchments but with limited success.

Therefore, calibration of models with a higher number of parameters would require to select a suitable data length in order to achieve the desired modelling accuracies and representativeness [9]. Practically, calibration is done with a representative dataset and the errors between observed and simulated data are minimized. This is then followed by model verification with an independent dataset in order to ensure the representativeness of the model.

As a result of practical constraints with data, process conceptualization and objective functions, model parameters are inevitably data dependent [5]. This hints that a search for data

lengths required for model applications would require to keep the parameter influence at a low level by selecting a model with a few parameters. In case of data limited situations, a model with a few parameters is usually preferred by hydrologists. This is because higher number of parameters increase the complexity of models and this in turn tends to decrease the model performance [8]. Hence the determination of data length must be linked with the number of parameters associated with the model.

2.4. Selection of Data Length

A practitioner searching for suitable data lengths has two options. They are either to obtain assistance from available guidelines or to evaluate reviewed research publications for acceptable recommendations.

2.4.1. Guideline Suggestions

Guidelines for Rainfall Runoff Modelling by Vaze et al. [55] without a specific reference to a data length indicate the importance of long-term streamflow observations for the development of a model that can produce streamflow estimations with an acceptable level of accuracy.

Hydrological Design Data Estimation Techniques issued by Czech Hydrometeorological Institute (WMO 1993) [56] states the use of 50-year monthly datasets as entry datasets for example studies. In case of rainfall analysis for flood forecasting, water security etc., it is important to consider continuous and multi decadal rainfall [57].

“Good Practice Guidelines-Water Data Management Policy”, issued by Bureau of Meteorology of Australia 2017 (WMO) [22] under water data temporal frequency states requirement of continuous high frequency rainfall and streamflow data for flood forecasting, and the sufficiency of monthly river flow data to assess regional water security. Also, it mentions that under water data longevity of measurements, multi-decadal records of continuous rainfall are necessary for reliable rainfall Intensity Duration Frequency (IDF) analysis.

Australian Rainfall and Runoff (ARR) “A Guide to Flood Estimation”, guideline updated in 2016 [58], states that calibration data length should be covering a range of different flood conditions in order to confidently use calibration results in data scarce situations. This

guideline further states that the parameter uncertainty decreases as the length of the data increases, and increases when the length of data decreases. More than a 10-year period is indicated as suitable for flood modelling. In cases where less than 10 years of data is available then regional approaches are recommended as more appropriate for computing base-flow contribution to design flood estimates. ARR goes on to mention that in an ideal situation, more than 10 years of continuous streamflow data is required to perform detailed site-specific analysis. However, ARR guideline before the update required at least 15 years of flood data for flood analysis.

Guideline issued by Department of Water Resources in Rajasthan, India [59] states that minimum of 25 years of monthly rainfall data and runoff data of maximum available period is required for yield studies. Guide to Hydrological Practices (WMO-No. 168) (1994) [60] states that at least 30 years of data is required to obtain a representative relationship between rainfall and runoff. WMO/TD - No.554 (WMO 1993) Hydrological Design Data Estimation Techniques issued by Czech Hydrometeorological Institute states about entry data sets of example studies done and mention the use of 50 years for monthly data.

Daggupati et al.[61] mentioning expensive resource requirements indicate the need to use shorter datasets instead of lengthier datasets. Discussing the needs for model development, Razavi and Tolson [62] recommend the use of one third of a long dataset for calibration while using a longer dataset for model verification. Zheng [20] splits the data set as 80% data for model calibration and 20% data for model verification based on common practice, e.g. May et al.[63]; Wu et al.[19]. A modellers' guide for rainfall runoff modelling by Kherde [64] recommends a one year dataset for the warm up period, a 10 year data set for calibration and a minimum of two years data for validation.

The variety of guideline statements indicated above, also necessitates a careful review of ongoing research to determine an appropriate data length for water resources model applications.

2.4.2. Research Recommendations

A comparison of reviewed research publications displays the recommendation of data lengths between one year and 30 years for

model calibration. A monthly dataset of 60 years enabling 30 years each for model calibration and verification has been recommended as the best modelling option for Nile River [65]. A comparative monthly water balance modelling effort where 5 non-overlapping 10-year data periods had been evaluated on 5 models and on 10 sites, had reported verification with 40 years of data [14]. Xu and Vandewiele [45], working on 91 catchments in Belgium and China having extents between 16 km² to 3626 km², and using 2, 5, 10, 15, 20 years long monthly data from a 158 year data series, concluded that a 10 year data period is necessary for an adequate model calibration. Selecting six events for calibration and 24 events for validation of simple and complex runoff models in semi-arid watersheds, Michaud and Sorooshian [66] concluded that a minimum data length of 15 years is required for calibration. GÖrgens [46] in a research on two hydrologic models for a semi-arid 73.1 km² catchment with monthly datasets of 3, 6, 10, 15 and 20 years, also concluded that 15 years of monthly data is required for a reliable model optimization. 30 years of data had been identified as the minimum requirement for realistic streamflow estimation to monitor droughts at two Indian catchments [67]. After a modelling study of 7 catchments in USA, Haan [32] stated that at least one and preferably two or three years of observed monthly flows are required for acceptable parameter estimation.

Burn and Elnur [68] recommended a minimum of 20 years of data for a reference hydrometric basin network. Kahya and Kalayci [69] in their work indicated that a 30 year period is long enough to compute a valid mean statistic. Nyunt et al.[70] carried out a climate change study and utilized 20 years of data for modelling. Similarly, for water management purposes, data lengths that had been used in research vary from 1 year to 65 years [7], [10], [73], [12], [14], [25], [32], [42], [45], [71], [72].

Sorooshian and Gupta [71] suggest that for any calibration procedure to be successful, the data should be representative of the various natural phenomena experienced by the watershed during a complete seasonal hydrological cycle. They also suggest choosing a wet year for calibration to activate all the model parameters. James [8] recommends collecting all rainfall and dry events in a dataset to calibrate the parameters, thereby ascertaining that the model is sensitive to all the events in the watershed.

After carrying out a research in Great Usuthu Catchment Gan, Dlamini and Biftu [5] suggested that wet years are preferred over dry years as calibration data, because dry years would not contain sufficiently high flows to excite the models.

There are many recommendations in the research carried out for daily water management applications. Perrin et al.[10], working with 12 catchments varying between 1021 km² and 4421 km² and using 39 years of daily data, stated that 350 days chosen randomly are sufficient to obtain robust estimates for the model. A sample size greater than 500 days has been reported as a threshold to achieve acceptable precision of model parameters [27]. A comparison of data lengths on a 407 km² watershed has highlighted the importance of daily data sets greater than 2-5 years for long term runoff estimation [17]. Anctil et al.[7], in their work for the Serein River Basin in France (1120 km²), concluded that best model performance was with 3 and 5 year calibration data sets. Sufficiency of at least 8 years of daily data for model calibration has been reported in an application using 55 watersheds having extents between 51-1891 km² by Li et al.[74] and by Gupta et al.[25] in their work for the Leaf river basin (1944 km²). It is noteworthy that Ye et al.[26] had discussed the inadequacy of a 5 year dataset for the modelling of low yielding watersheds in Australia. Need of at least one hydrologically appropriate year for the activation of model parameters has been highlighted by several researchers [5], [24], [75]. Choosing a suitable data length for modelling is balancing the length of dataset and the quality of data [24]. Data length options between shorter periods with high quality data and longer datasets with lesser quality must be appropriately chosen to improve the precision of parameters [27].

3. Practical Constraints

3.1. Data Length and Quality

Success of model calibration process depends upon the nature of the data used. Erroneous input data causes a significant adverse impact on model calibration [76]. The dataset used should be representative of the watershed processes. Even though most researchers had tried to achieve the representativeness of data by using longest available data lengths, the information contained in data and the way of extracting these information is considered as more important than the length of data series

[24]. Increased data lengths increase the computational burden to obtain the best parameter set and the cost of data acquisition. Hence the requisite parameter precision, computational constraints and data costs must be optimized to manage the desired quality of data [27]. Poor quality of data leading to unsatisfactory hydrologic modelling is well reflected in the study of 75 Belgian catchments by Vandewiele and Elias [72]. Rainfall data must not contain significant errors, and even random errors in a rainfall series significantly affect model performance and parameter values [77]. Xu and Vandewiele [45], having performed a sensitivity analysis for input data errors in 91 catchments of Belgium and China, concluded that random errors in rainfall data negatively influence model performance and that systematic errors are less important for the estimation of streamflow. However, the systematic errors do have a significant influence on model parameter values and consequently on the estimation of other components of water balance. On the other hand, Das and Bárdossy [11] quoting Chaplot et al.[78] state that models have the capacity to achieve performance criteria by compensating input errors that are within a reasonable range by adjusting their parameters values.

3.2. Missing Data Issues

Many research and work in projects use data without missing data periods and work of Anctil et al.[7] which focused only on periods without missing data is an example. They go on to mention that in most research the length of available observation series is rarely addressed. Missing data is often a major issue when attempting to acquire long datasets. This is especially because of the need to have a common dataset for all gauging stations representing the spatial variability of rainfall within a project area. To overcome the issue of non-continuous data in poorly gauged catchments, Perrin et al.[10] suggested investigations on model behavior with non-continuous data. Adhering to an appropriate data filling technique has been suggested as a measure to overcome the poor performance of models due to lack of long datasets adequately representing the variations in ground elevation and geography of a project area [11]. Though data length requirements are not specifically mentioned, many text books mention methods to fill the missing data in order to ensure common datasets for hydrological computations. There are commonly used data filling methods mentioned in text books. They

are: 1) Arithmetic Mean Method; 2) Normal Ratio Method; 3) Inverse Distance Weighing Method; 4) Linear Regression Method; 5) Weighted Linear Regression Method; 6) Multiple Linear Regression Method; 7) Probabilistic Method; 8) Nearest Station Method; and 9) Areal Precipitation Ratio Method [79]–[85].

Filling with the data from nearest available station has been used in an 8-year long daily dataset for a SWAT model application at six watersheds in North Ohio [86]. In a case study of 144 sub watersheds in 10 separate watersheds, Liew et al.[87] had used the inverse distance weighing method to fill the missing daily data of 31 gauges with the help of four closest gauging stations by constructing a 23 year and 8-year long common dataset. Khandu [88] has used a linear interpolation technique to fill missing rainfall data for a 48 year data set for a two parameter monthly water balance model development. Dissanayake [89], Artan et al.[90] and Gutierrez-Magness and McCuen [91] had filled the missing data to carry-out hydrologic modelling, water balance modelling, and rainfall runoff modelling by using 8 and 15, 13, and 15 year long datasets, respectively.

3.3. Common Data Periods

The data usage for hydrologic modelling varies with the purpose, type of model and model resolution. Occasions where similar studies on the same watershed had utilized different sets of data are shown in Table 1. In this summary, the study by Bastiaanssen and Chandrapala [92] is an exception because the reported work is based on a remote sensing satellite based evaluation. The choice of gauging stations for the same watershed had varied even with the use of same data resolution and the most likely reason for such differences could be the difficulties to access data. As indicated previously, these reasons could be the location preferences, financial limitations, data release constraints, time availability access data, period of interest etc. Figure 1 shows that, in a particular study area, the gauging stations and data duration have differed, thus hinting that the selection of either one of these parameters could influence the other. It also can be noted that the monthly studies have used greater data durations. This could be due to lesser cost of monthly data, and relatively easy accessibility pertaining to monthly data.

4. Reality of Application

4.1. Type of Publications

A summary of reviewed publications was prepared to evaluate Year of Publication, Watershed Area, Data Resolution, Geography of Usage, Type of model, Number of Parameters, Length of Calibration Data, Length of Verification Data, Length of Total Data Period and Purpose of the Study.

The distribution of Catchment areas, Year of Research Publication and Location of study watersheds reviewed in this work are shown in Figure 2, Figure 3 and Figure 4, respectively. The summary table with key parameters and corresponding references is in Table 2. Mixture of applications and studies listed in Table 2 indicate that a wide variety was subjected to the evaluation.

4.2. Monthly Water Resources Modelling

The present review identified 48 monthly water resources models of varying complexities. Of these models, two models had been the center of attraction (Figure 5). They are the PE Model [45] and the 2P Model [12] which indicated application percentages of 25.7 and 22.2 respectively. A summary of mostly used 10 models for monthly water resources modelling is in Table 2 and in Figure 5.

4.3. Models and Number of Parameters

Parameter use in daily and monthly water resource management models is shown in Figure 6. Daily models due to model complexities require a greater number of parameters than monthly models and this is clearly shown by the summary of applications. In general, the monthly model parameter number varies between 2 and 5. The same for daily models is approximately between 3 and 10. In most models commonly used for monthly water resources management applications, number of parameters varies between 2 and 4. A summary of parameters in the most used 10 models are shown in Figure 7 and in Table 3.

4.5. Data Length Water Resources Models

4.5.1. General

Data length used for model development is important to excite a model during parameter estimation and then to verify the model for its representativeness over hydrologic variability. Though this is amply discussed in literature, many reviewed literature neglects to explicitly

indicate either the rationale for selection of data length or the data period used for a particular application. Figure 8 and Figure 9 show the use of data length against the time of exceedance of daily and monthly scale watershed applications. The verification data length curve reaches zero data length at an approximate time of exceedance value of 58% and 41% for calibration and verification, respectively. This shows a lesser prominence given either to perform model verification or when mentioning the verification period. The summary could also be interpreted that the importance is given to calibration and hence verification had not been carried out. This supports the notion that calibration period is of greater importance when models are developed [20].

4.5.2. Calibration Data Length

Individual exceedance curves for calibration and verification cases are in Figures 10 to 14. Shape of the exceedance curve for calibration data length enables the identification of less frequent data usage. The less frequently used long data lengths, shown by highly varying slope of small time of exceedance, are approximately between 0% and 10% time of exceedance for calibration data periods. This segment with highly variable slope over a short % time of exceedance, is followed by a data length range with a shorter bandwidth having milder slopes and spanning over a longer % time of exceedance. This represents the most frequently used data lengths and provides an indication of the percentage usage. Subsequent to this section with a stable slope is another rapidly varying slope which corresponds to less frequent short data lengths. The data length ranges that had been used in applications are summarized in Table 4.

Daily models showed that Calibration data duration has respective values of 21, 19 and 19 years as the maximum, minimum and median of the watersheds subjected to the review between 0 to 10% times of exceedance. Similarly, the maximum, minimum and median of the reviewed watersheds are 19, 0.96 and 5 years respectively for % time of exceedance between 10 - 85%. Maximum, minimum and median of the reviewed watersheds greater than 85% exceedance is 0.96, 0.96 and 0.96 years, respectively.

Monthly models showed that calibration data duration has respective values of 65, 20 and 24.5 years as the maximum, minimum and

median of the watersheds subjected to the review between 0 to 10% times of exceedance. Similarly, the maximum, minimum and median of the watersheds reviewed are 20, 10 and 10 years respectively for % time of exceedance between 10 - 85%. Maximum, minimum and median of the reviewed watersheds greater than 85% exceedance is 10, 1 and 6 years, respectively.

4.5.3. Verification Data Length

The exceedance curve for verification data lengths shows the long and short less frequent data length usage at either end. The threshold values interpreted from the curve are shown in Table 5.

4.5.4. Total Data Length

The exceedance curve for total data lengths shows the long and short, less frequent data length usage at either end. Threshold values interpreted from the curve are shown in Table 6.

4.6. Catchment Size

All case applications reviewed in the present work were plotted to identify the behaviour between the data duration and the size of catchment. Semi logarithmic plots of data length against catchment size indicated a clustering of catchments in Figure 15, Figure 16, and Table 7. Number of applications against the data length is also indicated under each category.

Applications with watershed area between 0 - 10 square kilometers has a median of 3 years [32], [45], [93], [94]. In the range of 10 - 100 km², the median value is 10 years [14], [33], [45], [46]. Watershed with sizes from 100 - 1000 km² has the highest variation between maximum and minimum catchment areas. This is because of a single application with a 65-year data length. In this range, the maximum without this out of ordinary application is 10 years. The median value of data length for 100-1000 km² watershed is 11 years [12], [14], [33], [35], [38], [45], [48]-[50], [66]. Watershed applications between 1000 and 5000 km² indicated the median data length as 15 years [12], [36], [100], [38], [45], [49], [95]-[99]. Greater than 5000 km² study area catchments have a median of 19 years [1], [12], [34], [36], [95], [101]-[105].

A majority of applications in most of the considered watershed range indicated that the median value is between 10 and 19 years. It is felt reasonable to conclude that maximum data

length that had been considered by many for monthly applications is in the order of 30 years and that the most common data length could be approximated as 15 years.

5. Evaluation

The present review revealed that there is a clear deficiency of supporting material for the selection of an appropriate data length to carry out a monthly water resources model development. Available documentation pointed to the criteria that lead to making a rational decision. The following discussion presents the reasoning associated with the selection of criteria while facilitating the determination of available data length alternatives.

5.1. Model Excitation

Hydrological models require the calibration datasets to representatively include both wet and dry periods to achieve the precision of model parameters. In case of daily models, even a single representative year has been found as adequate. Models require data not only for model excitation but also to carry out the fitting of mathematical equations representing the watershed hydrology. Reviewed literature did not point to a minimum data length for the excitation of monthly models. One year of daily data contains 365 data points. Therefore, the minimum number of data points to calibrate a model can be considered similar to a daily model. On the other hand, a single water year in daily scale may contain sufficient hydrological variations. In case of monthly data, 360 points cover 30 years and it is felt that this would be more than sufficient to represent the hydrologic variation in monthly scale. Therefore, it is felt that even a 20-30 year period containing over 240 data points could be safely considered as a highly favourable data length when considering the model excitation requirements. Using a similar rationalization, a data length period between 10-20 years taken as moderately favourable while data lengths less than 10 years were considered as least favourable.

5.2. Data Quality

Literature recognizes the need for good quality data for successful model development. In this context, quality represents both the quality of measurements and the quality of hydrological representativeness. The latter is important for the model excitation. Quality of measurements is important for the representation of the

reality. Missing data is a huge challenge when attempting to preserve these characteristics. Modellers, who target the acquisition of a common data period for several rainfall, streamflow, evaporation and other climate data, are often faced with the missing data obstacle. It is a very much easier task to capture shorter common data periods without missing data. Therefore, shorter the desired data length the higher is the chance of successfully developing a watershed model. Hence in the present study, identification of data length selection options was based on the allocation of a lower preference rank to longer datasets. This, in other words, considers that the increase of data length is inversely proportionate to the quality of input data.

5.3. Precision of Parameters

Precision of parameters depends on the model calibration and verification. The review identified that many research work around the globe recommends a lengthy dataset for calibration. In certain occasions there had been recommendations to split the entire data set to allocate 80% for calibration and 20% for verification. Therefore, in case of calibration, not only longer data length classes lead to better model development but also exercise a higher weight on better model development. Even though the priority is low relative to calibration, longer model verification data leads to better model development. The evaluation of the present practice indicated that for shorter data lengths, verification had a much lower weightage compared to calibration. Selection of mid-range data lengths showed that calibration had received greater priority over verification. In the longer datasets it was noted that equal status had been granted for both calibration and verification. Therefore, in the present evaluation of data length options for model development, the calibration data length options were assigned a one-step higher weightage than the verification data length options.

5.4. Model Complexity

Complexity of a hydrologic model increases with the advances in improving mathematical descriptions of hydrological processes. Improved mathematical representations to represent the numerous physical processes observed in the real world inevitably increases the number of model parameters. Optimization of an increased number of model parameters require lengthier datasets which possess better chances of reflecting the variety of process

characteristics embedded in the model. The number of parameters and data length relationship in monthly water resources model applications reveal an increasing trend in the data length when the number of parameters increase.

5.5. Practice amidst Constraints

The practice of selecting a data length in monthly water resources applications was captured during the review. In most of the applications, the selection of data length has not been justified. The behaviour of data length selection in peer reviewed publications is considered as a reflection of executing model development and application amidst practical constraints. Peer reviewed publications, by nature, confirm the trustworthiness of the contents together with scrutiny and acceptance by experts in the same area of work. Therefore, the current selection trend was qualitatively ranked to identify a suitable data length option.

5.6. Assessment of Options

The present study exhibits the possibility of executing a qualitative evaluation using five key criteria. At the outset it is important to identify the available decision alternatives. Data length use in research and guidelines indicate that 20 to 40 yearlong datasets are the most favoured for monthly model applications. In order to meaningfully determine the best data length, it is necessary to evaluate the merits and demerits of various data lengths by dividing the options into shorter ranges. Therefore, the data length selection options were divided into five smaller Likert scale ranges to represent the responses as, "most preferred", "preferred", "acceptable", "low preference" and "very low preference". The five data range options which in turn form the decision alternatives are <10, 10-20, 20-30, 30-40 and >40 years, respectively. Decision criteria, alternatives, responses and a summary of the rationalization used to assign responses are given in Table 8.

Responses for each criterion were coded numerically by assigning values to measure the preferences, such as, 9="most preferred", 7="preferred", 5="acceptable", 3="low preference" and 1="very low preference". Table 9 presents the numerical values and the normalized indicator for each decision alternative. The data ranges greater than 10-20 years were found as acceptable. The most preferred options were to use a data range greater than 30-40 years. Normalised ranks

show that a water resources modeller making attempts to decide on the data length for a water resources management modelling application should at least select a data length between 20-30 years.

6. Conclusions

- The qualitative analysis carried out using the existing literature which is described in analysis indicates that a total data length of range 20-30 years or greater should be utilized for monthly water resources model development.
- The analysis for the literature review also highlights the need for the use of majority of data for model calibration. Therefore, it may be prudent to select at least a 20-year data period for model calibration.
- Decision alternative selection discussed above was based on the observations made during the review. It is necessary to carry out well designed research to determine design alternatives at finer data range.
- Data length selection for model development decisions should consider model excitation, data quality, precision of parameters, model complexity and current practice as decision criteria.

Acknowledgement

The support extended by the UNESCO Madanjeet Singh Centre for South Asia Water Management, Department of Civil Engineering, University of Moratuwa, Sri Lanka is gratefully acknowledged.

References

1. Wang, G., Xia, J., and Che, J., "Quantification of Effects of Climate Variations and Human Activities on Runoff by a Monthly Water Balance Model: A Case Study of the Chaobai River basin in northern China," *Water Resour. Res.*, vol. 45, no. 7, pp. 1-12, 2009, doi: 10.1029/2007WR006768.
2. Martin-Carrasco, F., Garrote, L., Iglesias, A., and Mediero, L., "Diagnosing Causes of Water Scarcity in Complex Water Resources Systems and Identifying Risk Management Actions," *Water Resour. Manag.*, vol. 27, no. 6, pp. 1693-1705, 2013, doi:10.1007/s11269-012-0081-6.
3. Porkka, M., Gerten, D., Schaphoff, S., Siebert, S., and Kummu, M., "Causes and Trends of Water Scarcity in Food Production," *Environ. Res. Lett.*, vol. 11, no. 1, p. 15001, 2016, doi: 10.1088/1748-9326/11/1/015001.

4. William W-G. Yeh, "Reservoir Management and Operations Models '," *Water Resour. Res.*, vol. 21, no. 12, pp. 1797-1818, 1985, doi: 10.1029/WR021i012p01797.
5. Gan, T. Y., Dlamini, E. M., and Biftu, G. F., "Effects of Model Complexity and Structure, Data Quality, and Objective Functions on Hydrologic Modeling," *J. Hydrol.*, vol. 192, no. 1-4, pp. 81-103, 1997, doi: 10.1016/S0022-1694(96)03114-9.
6. Wijesekera, N. T. S., "Parameter Estimation in Watershed Model: A Case Study Using Gin Ganga Watershed," *Inst. Eng. Sri Lanka*, vol. 1, pp. 26-32, 2000.
7. Antil, F., Perrin, C., and Andréassian, V., "Impact of the Length of Observed Records on the Performance of ANN and of Conceptual Parsimonious Rainfall-Runoff Forecasting Models," *Environ. Model. Softw.*, vol. 19, no. 4, pp. 357-368, 2004, doi: 10.1016/S1364-8152(03)00135-X.
8. William, J., *Rules for Responsible Modeling*, 2005.
9. Bárdossy, A., "Calibration of Hydrological Model Parameters for Ungauged Catchments", *Hydrol. Earth Syst. Sci.*, vol. 11, no. 2, pp. 703-710, 2007.
10. Perrin, C., Oudin, L., Andreassian, V., Rojas-Serna, C., Michel, C., and Mathevet, T., "Impact of Limited Streamflow Data on the Efficiency and the Parameters of Rainfall-Runoff Models," *Hydrol. Sci. J.*, vol. 52, no. 1, pp. 131-151, 2007, doi: 10.1623/hysj.52.1.131.
11. Das, T., and Bárdossy, A., "Influence of Rainfall Observation Network on Model Calibration and Application," *Hydrol. Earth Syst. Sci.*, pp. 77-89, 2008.
12. Xiong, L., and Guo, S., "A Two-Parameter Monthly Water Balance Model and its Application," *J. Hydrol.*, vol. 216, no. 1-2, pp. 111-123, 1999, doi: 10.1016/S0022-1694(98)00297-2.
13. Mouelhi, S., Michel, C., Perrin, C., and Andréassian, V., "Stepwise Development of a Two-Parameter Monthly Water Balance Model," *J. Hydrol.*, vol. 318, no. 1-4, pp. 200-214, 2006, doi: 10.1016/j.jhydrol.2005.06.014.
14. Alley, W. M., "On the Treatment of Evapotranspiration, Soil Moisture Accounting, and Aquifer Recharge in Monthly Water Balance Models," *Water Resour. Res.*, vol. 20, no. 8, pp. 1137-1149, 1984, doi: 10.1029/WR020i008p01137.
15. Klemeš, V., "Operational Testing of Hydrological Simulation Models," *Hydrol. Sci. J.*, vol. 31, no. 1, pp. 13-24, 1986, doi: 10.1080/02626668609491024.
16. Vandewiele, G. L., Xu, C. Y., and Ni-Lar-Win, "Methodology and Comparative Study of Monthly Water Balance Models in Belgium, China and Burma," *J. Hydrol.*, vol. 134, no. 1-4, pp. 315-347, 1992, doi: 10.1016/0022-1694(92)90041-S.
17. Boughton, W. C., "Effect of Data Length on Rainfall-Runoff Modelling," *Environ. Model. Softw.*, vol. 22, no. 3, pp. 406-413, 2007, doi: 10.1016/j.envsoft.2006.01.001.
18. Lü, H., Hou, T., Horton, R., Zhu, Y., Chen, X., Jia, Y., Wang, W., and Fu, X., "The Streamflow Estimation using the Xinanjiang Rainfall Runoff Model and Dual State-Parameter Estimation Method," *J. Hydrol.*, vol. 480, pp. 102-114, 2013, doi: 10.1016/j.jhydrol.2012.12.011.
19. Wu, W., May, R. J., Maier, H. R., and Dandy, G. C., "A Benchmarking Approach for Comparing Data Splitting Methods for Modeling Water Resources Parameters using Artificial Neural Networks," *Water Resour. Res.*, vol. 49, no. 11, pp. 7598-7614, 2013, doi: 10.1002/2012WR012713.
20. Zheng, F., Wu, W., Zhang, T., Maier, H. R., Dandy, G. C., and Gupta, H. V., "On Lack of Robustness in Hydrological Model Development Due to Absence of Guidelines for Selecting Calibration and Evaluation Data: Demonstration for Data-Driven Models," *Water Resour. Res.*, vol. 54, no. 2, pp. 1013-1030, 2018, doi: 10.1002/2017wr021470.
21. Cunderlik, J. M., "Hydrologic model selection for the CFCAS project: Assessment of Water Resources Risk and Vulnerability to Changing Climatic Conditions October 2003 Prepared by Juraj M. Cunderlik University of Western Ontario," no. October, 2003.
22. Bureau of Meteorology, *Good practice Guidelines for Water Data Management Policy*, 2017.
23. Elizabeth M. Shaw, Downloaded From: www.EasyEngineering.net, Third. 1994.
24. Sorooshian, S., Gupta, V. K., and Fulton, J. L., "Evaluation of Maximum Likelihood Parameter Estimation Techniques for Conceptual Rainfall-Runoff Models: Influence of C," *Water Resour. Res.*, vol. 19(1), no. 1, pp. 251-259, 1983.
25. Gupta, H. V., Sorooshian, S., and Yapo, P., "Automatic Calibration of Conceptual Rainfall-Runoff Sensitivity to Calibration Data Models," *J. Hydrol.*, vol. 181, pp. 23-48, 1996.
26. Ye, W., Bates, B. C., Viney, N. R., and Sivapalan, M., "of conceptual RESOURCES Performance of Conceptual Rainfall-Runoff Models in Low-Yielding Ephemeral Catchments Appreciation of the Ability to Predict Streamflow in these very difficult Cases than in Humid Researched on the Relationship between Three Ernphe," 1997.
27. Gupta, V. K., and Sorooshian, S., "The Automatic Calibration of Conceptual Catchment Models Using Derivative-Based Optimization Algorithms," *Water Resour. Res.*, vol. 21, no. 4, pp. 473-485, 1985, doi: 10.1029/WR021i004p00473.

28. Yapo, P. O., Gupta, H. V., and Sorooshian, S., "Automatic Calibration of Conceptual Rainfall-Runoff Models: Sensitivity to Calibration Data," *J. Hydrol.*, vol. 181, no. 1-4, pp. 23-48, 1996, doi: 10.1016/0022-1694(95)02918-4.
29. Loucks, D. P., and van Beek, E., "Water Resource Systems Planning and Management: An Introduction to Methods, Models, and Applications," *Water Resour. Syst. Plan. Manag. An Introd. to Methods, Model. Appl.*, pp. 1-624, 2017, doi: 10.1007/978-3-319-44234-1.
30. Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V., and Sorooshian, S., "A Framework for Development and Application of Hydrological Models," *Hydrol. Earth Syst. Sci.*, vol. 5, no. 1, pp. 13-26, 2001, doi: 10.5194/hess-5-13-2001.
31. Blöschl, G., and Sivapalan, M., "Scale Issues in Hydrological Modeling," *Hydrol. Process.*, vol. 9, pp. 251-290, 1995, doi: 10.1029/96eo00131.
32. Haan, C. T., "A Water Yield Model for Small Watersheds," *Water Resour. Res.*, vol. 8, no. 1, pp. 58-69, 1972.
33. Szolgay, J., Hlavčová, K., Kohnová, S., and Daníhlík, R., "Regional Estimation of Parameters of a Monthly Water Balance Model," *J. Hydrol. Hydromechanics*, vol. 51, pp. 256-273, 2003.
34. Hughes, D. A., "Incorporating Groundwater Recharge and Discharge Functions into an Existing Monthly Rainfall-Runoff Model," *Hydrol. Sci. J.*, vol. 49, no. 2, pp. 297-312, 2004, doi: 10.1623/hysj.49.2.297.34834.
35. Sha, J., Liu, M., Wang, D., Swaney, D. P., and Wang, Y., "Application of the ReNuMa Model in the Sha He River Watershed: Tools for Watershed Environmental Management," *J. Environ. Manage.*, vol. 124, pp. 40-50, 2013, doi: 10.1016/j.jenvman.2013.03.030.
36. Qi, Z., Kang, G., Chu, C., Qiu, Y., Xu, Z., and Wang, Y., "Comparison of SWAT and GWLF Model Simulation Performance in Humid South and Semi-Arid North of China," *Water (Switzerland)*, vol. 9, no. 8, 2017, doi: 10.3390/w9080567.
37. Chang, W., and Chen, X., "Monthly Rainfall-Runoff Modeling at Watershed Scale: A Comparative Study of Data-Driven and Theory-Driven Approaches," *Water (Switzerland)*, vol. 10, no. 9, p. 1116, Aug. 2018, doi: 10.3390/w10091116.
38. Mimikou, M., Kouvousopoulos, Y., Cavadias, G., and Vayianos, N., "Regional Hydrological Effects of Climate Change," *J. Hydrol.*, vol. 123, no. 1-2, pp. 119-146, 1991, doi: 10.1016/0022-1694(91)90073-Q.
39. WMO, "Annual Report of World Meteorological Organization," Geneva, Switzerland, 1976.
40. Mohseni, O., and Stefan, H. G., "A Monthly Streamflow Model," *Water Resour. Res.*, vol. 34, no. 5, pp. 1287-1298, 1998.
41. Xu, C., and Singh, V. P., "A Review on Monthly Water Balance Models for Water Resources Investigations," pp. 31-50, 1998.
42. Boyle, D. P., Gupta, H. V., and Sorooshian, S., "Toward Improved Calibration of Hydrologic Models Combining the Strengths of Manual and Automatic Methods," *Water Resources Research Volume 36, Issue 12,* *Water Resour. Res.*, vol. 36, no. 12, pp. 3663-3674, 2000.
43. Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S., "A Shuffled Complex Evolution Metropolis Algorithm for Optimization and Uncertainty Assessment of Hydrologic Model Parameters," *Water Resour. Res.*, vol. 39, no. 8, 2003, doi: 10.1029/2002WR001642.
44. Xu, C. Y., and Vandewiele, G. L., "Sensitivity of Monthly Rainfall-Runoff Models to Input Errors and Data Length," *Hydrol. Sci. J.*, vol. 39, no. 2, pp. 157-176, 1994, doi: 10.1080/02626669409492731.
45. GÖrgens, A. H. M., "Reliability of Calibration of a Monthly Rainfall-Runoff model: The semiarid case/Validité de l'étalonnage d'un modèle pluie-débit à pas de temps mensuel: Cas d'un bassin semi aride," *Hydrol. Sci. J.*, vol. 28, no. 4, pp. 485-498, 1983, doi: 10.1080/02626668309491990.
46. Chen, X., Chen, Y. D., and Xu, C., "A Distributed Monthly Hydrological Model for Integrating Spatial Variations of Basin Topography and Rainfall," vol. 21, no. April 2006, pp. 242-252, 2007, doi: 10.1002/hyp.
47. Porter, J. W., and McMahon, T. A., "A Model for the Simulation of Streamflow Data from Climatic Records," vol. 13, pp. 297-324, 1971.
48. Mahanama, S. P. P., Koster, R. D., Reichle, R. H., and Zubair, L., "The Role of Soil Moisture Initialization in Subseasonal and Seasonal Streamflow Prediction - A Case Study in Sri Lanka," *Adv. Water Resour.*, vol. 31, no. 10, pp. 1333-1343, 2008, doi: 10.1016/j.advwatres.2008.06.004.
49. Tayefeh Neskili, N., Zahraie, B., and Saghafian, B., "Coupling Snow Accumulation and Melt Rate Modules of Monthly Water Balance Models with the Jazim Monthly Water Balance Model," *Hydrol. Sci. J.*, vol. 62, no. 14, pp. 2348-2368, 2017, doi: 10.1080/02626667.2017.1373778.
50. Thornthwaite, C. W., and Mather, J. R., "The Water Balance," *Clim. Lab. Clim. Drexel Inst. Technol.*, vol. 8, no. 1, pp. 1-104, 1955.
51. Palmer, W. C., "Meteorologic Drought. Res. Pap.," *U. S. Weather Bur.*, p. 45-58 pp, 1965.
52. Thomas, H. A., "Improved Methods for National Water Assessment," 1981.
53. Janssen, P. H. M., and Heuberger, P. S. C., "Calibration of Process-Oriented Models," *Ecol. Modell.*, vol. 83, no. 1-2, pp. 55-66, 1995, doi: 10.1016/0304-3800(95)00084-9.
54. Vaze, G., Jordan, J., Beecham, P., Frost, R., Summerell, A., *Guidelines for Rainfall-Runoff Modelling Towards Best Practice Model Application*. 2012.
55. Novicky, D., Kasperek, L., and Kolarova, S., "Hydrological Design Data Estimation Techniques. TD/WMO; N 554." 1993.

56. Australia Bureau of Meteorology, "Guidelines for water data management policy," p. 60, 2017.
57. "Australian Rainfall Runoff, 'A Guide to Flood Estimation,'" 2016.
58. Department of Water Resources, "Basic Data/ Information Required for Hydrological Design," Rajasthan, India, 2015. [Online]. Available: http://waterresources.rajasthan.gov.in/6guidelines_hydrology.asp.
59. WMO, Guide to hydrological practices, 5th Editio. 1994.
60. P. Daggupati et al., "A Recommended Calibration and Validation Strategy for Hydrologic and Water Quality Models," *Trans. ASABE*, vol. 58, no. 6, pp. 1705–1719, 2016, doi: 10.13031/trans.58.10712.
61. Razavi, S., and Tolson, B. A., "An Efficient Framework for Hydrologic Model Calibration on Long Data Periods," *Water Resour. Res.*, vol. 49, no. 12, pp. 8418–8431, 2013, doi: 10.1002/2012WR013442.
62. May, R. J., Maier, H. R., and Dandy, G. C., "Data Splitting for Artificial Neural Networks using SOM-based Stratified Sampling, Neural Networks," *Neural Networks*, vol. 23, no. 2, p. 283, 2010.
63. Kherde, R. V., "Rainfall Runoff Modelling – Modellers Guide," *Int. J. Appl. Eng. Res.*, vol. 11, no. 3, pp. 1900–1907, 2016.
64. Keshtegar, B., Allawi, M. F., Afan, H. A., and El-Shafie, A., "Optimized River Stream-Flow Forecasting Model Utilizing High-Order Response Surface Method," *Water Resour. Manag.*, vol. 30, no. 11, pp. 3899–3914, 2016, doi: 10.1007/s11269-016-1397-4.
65. Michaud, J. and Sorooshian, S., "Comparison of Simple Versus Complex Distributed Runoff Models on a Midsized Semiarid Watershed," *Water Resour. Res.*, vol. 30, no. 3, pp. 593–605, 1994, doi: 10.1029/93WR03218.
66. Jain, V. K., Jain, M. K., and Pandey, R. P., "Effect of the Length of the Streamflow Record on Truncation Level for Assessment of Streamflow Drought Characteristics," *J. Hydrol. Eng.*, vol. 19, no. 7, pp. 1361–1373, 2014, doi: 10.1061/(ASCE)HE.1943-5584.0000922.
67. Burn, D. H., and Elnur, M. A. H., "Detection of hydrologic trends and variability. *Journal of Hydrology*," *J. Hydrol.*, vol. 255, no. 1, pp. 107–122, 2002.
68. Kahya, E., Kalayci, and S., "(2004). Trend Analysis of Streamflow in Turkey. *Journal of Hydrology*, 289(2): pp 128–144," *J. Hydrol.*, vol. 289, no. 2, p. pp 128–144, 2004.
69. Nyunt, C. T., Koike, T., Yamamoto, H., Yamamoto, A., Nemoto, T., and Kitsuregawa, M., "a Pplication of Bias-Correction and Downscaling Method To," *Annu. J. Hydraul. Eng. JSCE*, vol. 68, no. 4, p. I_115-I_120, 2012.
70. Antil, F., Lauzon, N., Andréassian, V., Oudin, L., and Perrin, C., "Improvement of Rainfall-Runoff Forecasts through Mean Areal Rainfall Optimization," *J. Hydrol.*, vol. 328, no. 3–4, pp. 717–725, 2006, doi: 10.1016/j.jhydrol.2006.01.016.
71. Sorooshian, S. and Gupta, V. K., "Automatic Calibration of Conceptual Rainfall-Runoff Models: The Question of Parameter Observability and Uniqueness," *Water Resour. Res.*, vol. 19, no. 1, pp. 260–268, 1983, doi: 10.1029/WR019i001p00260.
72. Vandewiele, G. L., and Elias, A., "Monthly Water Balance of Ungauged Catchments Obtained by Geographical Regionalization," *J. Hydrol.*, vol. 170, no. 1–4, pp. 277–291, 1995, doi: 10.1016/0022-1694(95)02681-E.
73. Li, C. Z., Wang, H., Liu, J., Yan, D. H., Yu, F. L., and Zhang, L., "Effect of Calibration Data Series Length on Performance and Optimal Parameters of Hydrological Model," *Water Sci. Eng.*, vol. 3, no. 4, pp. 378–393, 2010, doi: 10.3882/j.issn.1674-2370.2010.04.002.
74. Refsgaard, J. C., "Model and Data Requirements for Simulation of Runoff and Land Surface Processes," *L. Surf. Process. Hydrol.*, vol. 146, pp. 423–452, 1997, doi: 10.1007/978-3-642-60567-3.
75. Yen, H., Wang, X., Fontane, D. G., Harmel, R. D., and Arabi, M., "A Framework for Propagation of Uncertainty Contributed by Parameterization, Input Data, Model Structure, and Calibration/Validation Data in watershed modeling," *Environ. Model. Softw.*, vol. 54, pp. 211–221, 2014, doi: 10.1016/j.envsoft.2014.01.004.
76. Oudin, L., Lauzon, N., Andre, V., and Perrin, C., "Improvement of Rainfall-Runoff Forecasts through Mean Areal Rainfall Optimization," vol. 328, pp. 717–725, 2006, doi: 10.1016/j.jhydrol.2006.01.016.
77. Chaplot, V., Saleh, A., and Jaynes, D. B., "Effect of the Accuracy of Spatial Rainfall Information on the Modeling of Water, Sediment, and NO₃-N loads at the watershed level," *J. Hydrol.*, vol. 312, no. 1–4, pp. 223–234, 2005, doi: 10.1016/j.jhydrol.2005.02.019.
78. Te Chow, V., Maidment, D. R., and Mays, L. W., *Applied Hydrology*. McGraw-Hill Book Company, 1988.
79. De Silva, R. P., Dayawansa, N. D. K., and Ratnasiri, M. D., "A Comparison of Methods used in Estimating Missing Rainfall Data," *J. Agric. Sci.*, vol. 3, no. 2, p. 101, 2016, doi: 10.4038/jas.v3i2.8107.
80. Singh, V. P., *Elementary Hydrology*. New Delhi: Prentice Hall of India, 1994.
81. Allison, P. D., *Missing data (Vol. 136)*. Sage publications, 2001.
82. Hasan, M. M., and Croke, B., *Filling Gaps in Daily Rainfall Data: A Statistical Approach*. 2013.
83. Chen, F. W., and Liu, C. W., "Estimation of the Spatial Rainfall Distribution using Inverse Distance Weighting (IDW) in the Middle of Taiwan," *Paddy Water Environ.*, vol. 10, no. 3, pp. 209–222, 2012.

84. Campozano, L., Sánchez, E., Aviles, A., and Samaniego, E., "Evaluation of Infilling Methods for Time Series of Daily Precipitation and Temperature: The Case of the Ecuadorian Andes," 2014.
85. Bosch, N. S., Allan, J. D., Dolan, D. M., Han, H., Richards, and R. P., "Application of the Soil and Water Assessment Tool for six Watersheds of Lake Erie: Model Parameterization and Calibration," *J. Great Lakes Res.*, vol. 37, no. 2, pp. 263-271, 2011, doi: 10.1016/j.jglr.2011.03.004.
86. Van Liew, M. W., Arnold, J. G., and Garbrecht, J. D., "H s a w : c b t m," vol. 46, no. 6, pp. 1539-1551, 2003.
87. Khandu, D., "A Monthly Water Balance Model for Evaluation of Climate Change Impacts on the Streamflow of Ginganga and Kelani Ganga Basins, Sri Lanka," (unpub), 2015.
88. Dissanayake, P. K. M., "Applicability of a Two Parameter Water Balance Model To Simulate Daily Rainfall Runoff - Case Study of Kalu and Gin River Basins in Sri Lanka," (unpub), 2017.
89. Artan, G., Gadain, H., Smith, J. L., Asante, K., Bandaragoda, C. J., and Verdin, J. P., "Adequacy of Satellite Derived Rainfall Data for Stream Flow Modeling," *Nat. Hazards*, vol. 43, no. 2, pp. 167-185, 2007, doi: 10.1007/s11069-007-9121-6.
90. Gutierrez-Magness, A., McCuen, and R., "Accuracy Evaluation of Rainfall Disaggregation Methods," *J. Hydrol. Eng.*, vol. 9, no. 2, pp. 71-78, 2004, doi: 10.1061/(ASCE)1084-0699(2004)9:2(71).
91. Bastiaanssen, W. G. M., and Chandrapala, L., "Water Balance Variability Across Sri Lanka for Assessing Agricultural and Environmental Water Use," *Agric. Water Manag.*, vol. 58, no. 2, pp. 171-192, 2003, doi: 10.1016/S0378-3774(02)00128-2.
92. Allred, B., and Haan, C. T., "SWMHMS-SMALL WATERSHED MONTHLY HYDROLOGIC MODELING SYSTEM," *J. Am. Water Resour. Assoc.*, vol. 32, no. 3, pp. 541-552, 1996.
93. Bobba, A. G., Singh, V. P., Jeffries, D. S., and Bengtsson, L., "Application of a Watershed Runoff Model to North-east Pond River, Newfoundland: To Study Water Balance and Hydrological Characteristics Owing to Atmospheric Change," *Hydrol. Process.*, vol. 11, no. 12, pp. 1573-1593, 1997, doi: 10.1002/(SICI)1099-1085(19971015)11:12<1573::AID-HYP491>3.0.CO;2-V.
94. dos Santos, F. M., de Oliveira, R. P., and Mauad, F. F., "Lumped versus Distributed Hydrological Modeling of the Jacaré-Guaçu Basin, Brazil," *J. Environ. Eng.*, vol. 144, no. 8, p. 04018056, 2018, doi: 10.1061/(asce)ee.1943-7870.0001397.
95. Hughes, D. A., "Monthly Rainfall-Runoff Models Applied to Arid and Semiarid Catchments for Water Resource Estimation Purposes," *Hydrol. Sci. J.*, vol. 40, no. 6, pp. 751-769, 1995, doi: 10.1080/02626669509491463.
96. Fernandez, W., Vogel, R. M., and Sankarsubramanian, A., "Regional Calibration of a Watershed Model," *Hydrol. Sci. J.*, vol. 45, no. 5, pp. 689-707, 2000, doi: 10.1080/02626660009492371.
97. Jothityangkoon, C., Sivapalan, M., and Farmer, D. L., "Process Controls of Water Balance Variability in a Large Semi-Arid Catchment: Downward Approach to hydrological model development," *J. Hydrol.*, vol. 254, no. 1-4, pp. 174-198, 2001, doi: 10.1016/S0022-1694(01)00496-6.
98. Martinez, G. F., and Gupta, H. V., "Toward Improved Identification of Hydrological Models: A Diagnostic Evaluation of the 'abcd' Monthly Water Balance Model for the Conterminous United States," *Water Resour. Res.*, vol. 46, no. 8, pp. 1-21, 2010, doi: 10.1029/2009WR008294.
99. Machado, F., Mine, M., Kaviski, E., and Fill, H., "Monthly Rainfall-Runoff Modelling using Artificial Neural Networks," *Hydrol. Sci. J.*, vol. 56, no. 3, pp. 349-361, 2011, doi: 10.1080/02626667.2011.559949.
100. Chen, X., Chen, Y. D., Xu, and C., "A Distributed Monthly Hydrological Model for Integrating Spatial Variations of Basin Topography and Rainfall," *Hydrol. Process.*, vol. 21, pp. 242-252, 2006, doi: 10.1002/hyp.6187 A.
101. Jain, V. K., Jain, M. K., and Pandey, R. P., "Effect of the Length of the Streamflow Record on Truncation Level for Assessment of Streamflow Drought Characteristics," *J. Hydrol. Eng.*, vol. 19, no. 7, pp. 1361-1373, 2013, doi: 10.1061/(asce)he.1943-5584.0000922.
102. Bai, P., Liu, X., Liang, K., and Liu, C., "Comparison of Performance of Twelve Monthly Water Balance Models in Different Climatic Catchments of China," *J. Hydrol.*, vol. 529, pp. 1030-1040, 2015, doi: 10.1016/j.jhydrol.2015.09.015.
103. Jajarmizadeh, M., Kakaei Lafdani, E., Harun, S., Ahmadi, and A., "Application of SVM and SWAT Models for Monthly Streamflow Prediction, A Case Study in South of Iran," *KSCE J. Civ. Eng.*, vol. 19, no. 1, pp. 345-357, 2015, doi: 10.1007/s12205-014-0060-y.
104. Vilaysane, B., Takara, K., Luo, P., Akkharath, I., and Duan, W., "Hydrological Stream Flow Modelling for Calibration and Uncertainty Analysis Using SWAT Model in the Xedone River Basin, Lao PDR," *Procedia Environ. Sci.*, vol. 28, no. SustaiN 2014, pp. 380-390, 2015, doi: 10.1016/j.proenv.2015.07.047.
105. Perera, K. R. J., and Wijesekera, N. T. S., "Potential on the Use of GIS Watershed Modeling for River Basin Planning - Case Study of Attanagalu Oya Basin, Sri Lanka," *Eng. J. Inst. Eng. Sri Lanka*, vol. 45, no. 4, p. 13, 2012, doi: 10.4038/engineer.v45i4.6922.

106. Wijesekera, N. T. S., and Rajapakse, R. L. H. L., "Mathematical Modelling of Watershed Wetland Crossings for Flood Mitigation and Groundwater Enhancement - Case of the Attanagalu Oya River Basin," *Eng. J. Inst. Eng. Sri Lanka*, vol. XXXXVT, no. 03, pp. 55-67, 2013.
107. Wickramaarachchi, T. N., Ishidaira, H., and Wijayarathna, T. M. N., "Applicability of Global Public Domain Data Versus Local Detailed Data for Distributed Hydrological Modelling: A Study form Gin River Basin Sri Lanka," *Appl. Water Sci.*, vol. 3, no. 2, pp. 545-557, 2013, doi: 10.1007/s13201-013-0102-5.
108. Samarasinghe, S. M. J. S., Nandalal, H. K., Weliwitiya, D. P., Fowze, J. S. M., Hazarika, M. K., and Samarakoon, L., "Application of Remote Sensing and GIS for Flood Risk Analysis: A Case Study at Kalu-Ganga River, Sri Lanka International Archives of the Photogrammetry," *Remote Sens. Spat. Inf. Sci.*, no. 38(Part 8), 2010.
109. Schulz, L., and Kingston, D. G., "GCM-Related Uncertainty in River Flow Projections at the Threshold for 'dangerous' Climate Change: The Kalu Ganga River, Sri Lanka," *Hydrol. Sci. J.*, vol. 62, no. 14, pp. 2369-2380, 2017, doi: 10.1080/02626667.2017.1381965.
110. Wijesekera, N. T. S., and Musiake, K., "Streamflow Modelling of Sri Lankan Catchments (2) - Kalu River Catchment at Putupaula," *Seisan Kenkyu*, vol. 42, no. 11, pp. 645-648, 1990.
111. Kanchanamala, D., Herath, H., and Nandalal, K. D. W., "Impact of Catchment Scale on Rainfall Runoff Modeling: Kalu Ganga River Catchment upto Ratnapura," *Eng. J. Inst. Eng. Sri Lanka*, vol. 49, no. 2, pp. 1-7, 2016, doi: 10.4038/engineer.v49i2.7003.
112. Jayadeera, P., "Development of a Rainfall Runoff Model for Kalu Ganga Basin of Sri Lanka using HEC-HMS model," no. May, p. 187, 2016, doi: 10.1079/BJN20051538.
113. De Silva, M. M. G. T., Weerakoon, S. B., and Herath, S., "Modeling of Event and Continuous Flow Hydrographs with HEC-HMS: Case Study in the Kelani River Basin, Sri Lanka," *J. Hydrol. Eng.*, vol. 19, no. 4, pp. 800-806, 2014, doi: 10.1061/(ASCE)HE.1943-5584.0000846.
114. Wijesekera, N. T. S., and Musiake, K., "Streamflow Modelling of Sri Lankan Catchments (1) - Mahaweli River Catchment at Peradeniya," *Seisan Kenkyu*, vol. 42, no. 11, pp. 645-648, 1990.
115. Wijesinghe, W. M. D., and Wijesekera, N. T. S., "Recognising Temporal Variability of Lumped Watershed Behaviour and Evaluating Average Performance of a Hydrological Model," in *Civil Engineering Research for Industry*, 2012, pp. 85-92.
116. Thapa, G., and Wijesekera, N. T. S., "Computation and Optimization of Snyder 's Synthetic Unit Hydrograph Parameters," pp. 83-88, 2017.
117. Nandalal, H. K., and Ratnayake, U. R., "Figure I. Kalu-Gan ga River with," vol. 14, pp. 486-488, 2009.
118. Dahanayake, A. C., and Wijesekera, N. T. S., "Potential of Water Balance Modelling with Surface Water Pollution Considerations to Manage Ungauged Watersheds with an Emphasis on Multi User Concepts - Demonstrating an Application at a Watershed in Dampe , Sri Lanka," pp. 21-28, 2017.
119. Wijesekera, N. T. S., "Parameter Estimation in Watershed Model: A Case Study Using Gin Ganga Watershed TRANSACTIONS 2000 VOLUME I - PART B," vol. 1, pp. 26-32, 2000.
120. Gunasekara, D. N., " ' abcd ' Monthly Water Balance Model for Kalu Ganga and Gin Ganga Basins and Its Application Potential for Water Resources Investigation," 2018.
121. Ampitiyawatta, A., and Guo, S., "Precipitation Trends in the Kalu Ganga basin in Sri Lanka," *J. Agric. Sci.*, vol. 4, no. 1, p. 10, 2010, doi: 10.4038/jas.v4i1.1641.
122. Xiong, L., and Guo, S., "A Two-Parameter Monthly Water Balance Model and Its Application," *J. Hydrol.*, vol. 216, no. 1-2, pp. 111-123, 1999, doi: 10.1016/S0022-1694(98)00297-2.
123. Mahanama, S. P. P., Koster, R. D., Reichle, R. H., and Zubair, L., "The Role of Soil Moisture Initialization in Subseasonal and Seasonal Streamflow Prediction - A Case Study in Sri Lanka," *Adv. Water Resour.*, vol. 31, no. 10, pp. 1333-1343, 2008, doi: 10.1016/j.advwatres.2008.06.004.
124. Kirchner, J. W., "Catchments as Simple Dynamical Systems: Catchment Characterization, Rainfall-Runoff Modeling, and doing Hydrology Backward," *Water Resour. Res.*, vol. 45, no. 2, pp. 1-34, 2009, doi: 10.1029/2008WR006912.

Table 1 - Summary of a few Sri Lankan Modelling Studies

Reference	River Basin	Gauging Station	Data Resolution	No. of Gauging Stations	Calibration Data Duration	Catchment Area (km ²)
Perera & Wijesekera [106]	Attanagalu Oya Basin	Karasnagala	Daily	5	10	53
Wijesekera & Rajapakse [107]	Attanagalu Oya Basin	Karasnagala	Daily	3	1	54
Bastiaanssen and Chandrapala [92]	Attanagalu Oya Basin	Negombo Lagoon	Daily	1	1	736
Wickramaarachchi, Ishidaira and Wijayaratna [108]	Gin River	Agaliya	Daily	6	4	780
Bastiaanssen and Chandrapala [92]	Gin River	Galle	Daily	1	1	932
Dissanayake, [89]	Gin River	Thawalama	Daily	4	8	364
Wickramaarachchi, Ishidaira and Wijayaratna [108]	Gin River	Thawalama	Daily	6	4	470
Dissanayake [89]	Kalu River	Ellagawa	Daily	5	4	1372
Bastiaanssen and Chandrapala [92]	Kalu River	Kalutara	Daily	1	1	2719
Nyunt et al [70]	Kalu River	Putupaula	Daily	26	20	2598
Samarasinghe, Nandalal, Weliwitiya, Fowze, Hazarika, & Samarakoon, [109]	Kalu River	Putupaula	Daily	13	10	2658
Schulz and Kingston [110]	Kalu River	Putupaula	Daily	1	4.75	2765
Wijesekera & Musiaka [111]	Kalu River	Putupaula	Daily	6	3	2598
Kanchanamala, Herath and Nandalal, Dissanayake, Schulz and Kingston [89], [110], [112]	Kalu River	Ratnapura	Daily	7	5	603
Jayadeera [113]	Kalu River	Ellagawa	Daily	5	4	1350
Jayadeera [113]	Kalu River	Ratnapura	Daily	4	4	635.1
Bastiaanssen and Chandrapala [92]	Kelani River	Colombo	Daily	1	1	2292
De Silva, Weerakoon and Herath [114]	Kelani River	Hanwella	Daily	8	2	2230
Wijesekera & Musiaka [115]	Mahaweli River	Peradeniya	Daily	4	4	1167

Table 1 - Summary of a few Sri Lankan Modelling Studies (Continued)

Reference	River Basin	Gauging Station	Data Resolution	No. of Gauging Stations	Calibration Data Duration	Catchment Area (km ²)
Wijesinghe and Wijesekera [116]	Attanagalu Oya Basin	Karasnagala	Event Based/Daily	1	30 events	52.8
Thapa and Wijesekera [117]	Attanagalu Oya Basin	Karasnagala	Event Based/Daily	2	30 events	52.58
Nandalal and Ratnayake [118]	Kalu River	Ratnapura	Event Based/Daily	1	4 Events	604
Dahanayake and Wijesekera [119]	Dampe	Ratmalana	Monthly	1	1	0.62
Wijesekera [120]	Gin River	Thawalama	Monthly	4	3	361
Khandu [88]	Gin River	Thawalama	Monthly	4	20	368.75
Gunasekara [121]	Gin River	Thawalama	Monthly	4	13	360
Ampitiyawatta and Guo [122]	Kalu River	Putupaula	Monthly	8	40	2598
Gunasekara [121]	Kalu River	Ellagawa	Monthly	5	10	1383
Khandu [88]	Kelani River	Deraniyagala	Monthly	4	24	182.54
Perera, & Wijesekera [106]	Kelani River	Glencourse	Monthly	6	10	1537

Table 2 - Summary of Mostly Used Monthly Water Resources Models

No.	Name of Model	Model Reference	Number of Parameters	Number of Monthly Applications	% of Applications
1	PE Model	Xu & Vandewiele [45]	3	81	25.7
2	2P MWBM	Xiong & Guo [123]	2	70	22.2
3	The Catchment Land Surface Model (CLSM)	Mahanama, Koster, Reichle, & Zubair [124]	**	22	7.0
4	WatBal	Szolgay, Hlavčová, Kohnová & Daníhlík [33]	5	14	4.4
5	abcd Model	Alley [14] Bai, Liu, Liang, & Liu [103]	4	13	4.1
6	T Model	Alley [14]	2	10	3.2
7	T _α Model	Alley [14]	3	10	3.2
8	T _γ Model	Alley [14]	3	10	3.2
9	P Model	Alley [14]	3	10	3.2
10	4 Parameter RR Model	Kirchner, Haan [32], [125]	4	8	2.5

** Number of Parameters were unspecified

Table 3 - Summary of Mostly Used Monthly Water Resources Models

No.	Name of Model	Model Reference	Number of Parameters	Number of Monthly Applications	% of Applications
1	PE Model	Xu & Vandewiele[45]	3	81	25.7
2	2P MWBM	Xiong & Guo [123]	2	70	22.2
3	The Catchment Land Surface Model (CLSM)	Mahanama, Koster, Reichle & Zubair [124]	**	22	7.0
4	WatBal	Szolgay, Hlavčová, Kohnová, & Daníhlík [33]	5	14	4.4
5	abcd Model	Alley, Bai, Liu, Liang, & Liu [14], [103]	4	13	4.1
6	T Model	Alley [14]	2	10	3.2
7	T α Model	Alley [14]	3	10	3.2
8	T τ Model	Alley [14]	3	10	3.2
9	P Model	Alley [14]	3	10	3.2
10	4 Parameter RR Model	Kirchner, Haan [32], [125]	4	8	2.5

Table 4 - Data Length and Temporal Scale (Calibration Period)

	Data Length (Monthly Scale)			Data Length (Daily Scale)		
% Time of Exceedance	Maximum	Minimum	Median	Maximum	Minimum	Median
< 10	65	20	24.5	21	19	19
10 - 85	20	10	10	19	0.96	5
> 85	10	1	6	0.96	0.96	0.96

Table 5 - Data length and Temporal Scale (Verification Period)

	Data Length (Monthly Scale)			Data Length (Daily Scale)		
% Time of Exceedance	Maximum	Minimum	Median	Maximum	Minimum	Median
< 28	101	40	40	20	11	20
28 - 84	40	5	6	11	4	6
> 84	4	2	3	4	2	3

Table 6 - Data Length and Temporal Scale (Total Data Period)

	Data Length (Monthly Scale)			Data Length (Daily Scale)		
% Time of Exceedance	Maximum	Minimum	Median	Maximum	Minimum	Median
< 19	101	50	50	31	19	31
19% - 94	40	10	15	19	0.96	19
> 94	9	3	8	0.96	0.96	0.96

Table 7 - Calibration Data Duration Variation According to Watershed Area

Range	0-10	10-100	100-1000	1000-5000	>5000
Max	10	16	65	30.4	33
Min	1	10	5	7	6
Average	3	11	15	15	18
Median	3	10	11	15	19

Table 8 - Decision Criteria, Alternatives and Associated Preferences

	Decision Alternatives Data Length Selection Options (number of Years)					Remarks / Rationalization
Criteria *	<10	10-20	20-30	30-40	>40	
Model Excitation	Very Low Preference	Low Preference	Acceptable	Preferred	Most Preferred	30 years with 360 data points would be sufficient to fulfil statistical requirements if any. 30 years would adequately represent the desired hydrological characteristics.
Data Quality	Most Preferred	Preferred	Acceptable	Low Preference	Very Low Preference	Shorter data sets have a higher likelihood to possess high quality
Precision of Parameters Calibration	Very Low Preference	Low Preference	Acceptable	Preferred	Most Preferred	Quality of parameter estimation enhances with the length of data used for calibration
Precision of Parameters Verification	Low Preference	Acceptable	Preferred	Most Preferred	Most Preferred	Verification data length has a lower priority according to data splitting recommendations and also as noted in practice

Model Complexity	Very Low Preference	Low Preference	Acceptable	Preferred	Most Preferred	Models representing reality prefer better representations with lengthier datasets.
Practice amidst Constraints	Very Low Preference	Most Preferred	Preferred	Preferred	Very Low Preference	Use of clusters noted by the change of slope in percentage exceedance curves
* The Likert scale: Most Preferred, Preferred, Acceptable, Low Preference and Very Low Preference are assigned 9,7,5,3, and 1 to reflect the difference in the normalized rank between each data length option						

Table 9 - Normalized Rank and Decision Alternatives for Data Length Selection

Criteria*	Decision Alternatives				
	Data Length Selection Options (number of Years)				
	<10	10-20	20-30	30-40	>40
Model Excitation	1	3	5	7	9
Data Quality	9	7	5	3	1
Precision of Parameters Calibration	1	3	5	7	9
Precision of Parameters Verification	3	5	7	9	9
Model Complexity	1	3	5	7	9
Practice amidst Constraints	1	9	7	7	1
Normalized Rank for Decision	2.67	5	5.67	6.67	6.33

* The Likert scale: Most Preferred, Preferred, Acceptable, Low Preference and Very Low Preference are assigned 9,7,5,3, and 1 to reflect the difference in the normalised rank between each data length option

Figure 1 - Data Length, Station Numbers and Resolution of a few Sri Lankan Modelling Studies

Figure 2 - Geographical Distribution of Study Watersheds

Figure 3 - Distribution of Catchment Area of Study Watersheds

Figure 4 - Distribution of Applications and Year of Publication

Figure 5 - Mostly Used Monthly Water Resources Models in Literature

Figure 6 - Occurrence of Parameter Use in Water Resources Applications

Figure 7 - Parameters in Mostly Reported Monthly Water Resources Models

Figure 8 - Percentage Time of Exceedance of Data Length in All Applications (Daily Models)

Figure 9 - Percentage Time of Exceedance of Data Length in All Applications (Monthly Models)

Figure 10 - Percentage Time of Exceedance of Calibration and Verification Data Period in Associated Applications (Daily Models)

Figure 11 - Percentage Time of Exceedance of Calibration and Verification Data Period in Associated Applications (Monthly Models)

Figure 12 - Calibration Data Length of Daily and Monthly Temporal Scale

Figure 13 - Verification Data Length of Daily and Monthly Temporal Scale

Figure 14 - Total Data Length of Daily and Monthly Temporal Scale

Figure 15 - Calibration Data Duration Variation According to Watershed Area

Figure 16 - Distribution of Calibration Data Duration According to Size of Watershed