REGIONALIZATION OF PARAMETERS OF HYDROLOGICAL MODELS IN SOUTH INDIAN RIVER BASINS

Thesis

Submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

CHANDRASHEKARAYYA G. HIREMATH
165137AM16P01

DEPARTMENT OF WATER RESOURCES & OCEAN ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA
SURATHKAL, MANGALORE – 575 025

OCTOBER 2024

REGIONALIZATION OF PARAMETERS OF HYDROLOGICAL MODELS IN SOUTH INDIAN RIVER BASINS

Thesis

Submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

CHANDRASHEKARAYYA G. HIREMATH

165137AM16P01

Research Guide

Dr. LAKSHMAN NANDAGIRI

Professor

DEPARTMENT OF WATER RESOURCES & OCEAN ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

SURATHKAL, MANGALORE – 575 025

OCTOBER 2024

DECLARATION

By the Ph.D. Research Scholar

I hereby *declare* that the Research Thesis entitled **"Regionalization of Parameters of Hydrological Models in South Indian River Basins"**, which is being submitted to the **National Institute of Technology Karnataka, Surathkal** in partial fulfilment of the requirements for the award of the Degree of **Doctor of Philosophy** in the Department of **Water Resources & Ocean Engineering** (Formerly Department of Applied Mechanics and Hydraulics) is a *bonafide report of the research work carried out by me.* The material contained in this Research Thesis has not been submitted to any University or Institution for the award of any degree.

165137AM16P01, Chandrashekarayya G. Hiremath
(Register Number, Name & Signature of the Research Scholar)
Department of Water Resources and Ocean Engineering
National Institute of Technology Karnataka, Surathkal

Place: NITK-Surathkal

Date: 17.10.2024

CERTIFICATE

This is to *certify* that the Research Thesis entitled **"Regionalization of Parameters of Hydrological Models in South Indian River Basins"** submitted by **Chandrashekarayya G. Hiremath** (Register Number: 165137AM16P01) as the record of the research work carried out by him, is *accepted as the Research Thesis submission* in partial fulfilment of the requirements for the award of degree of **Doctor of Philosophy**.

Lakshman
17/10/2024

Dr. Lakshman Nandagiri
Professor
(Research Guide)

Dept. of Water Resources & Ocean Engineering ★ NITK Surathkal

Varija K
17.10.24

Dr. Varija K
Chairman - DRPC
Department of Water Resources and Ocean Engineering
National Institute of Technology Karnataka, Surathkal

Chairman (DRPC)
Dept. of Water Resources & Ocean Engineering

ACKNOWLEDGEMENT

I wish to extend my profound gratitude and sincere appreciation to my supervisor, Dr. Lakshman Nandagiri, Professor in the Department of Water Resources and Ocean Engineering at NITK, for his invaluable guidance, unwavering encouragement, and gracious support throughout the duration of my research. Dr. Nandagiri's insightful suggestions, enthusiastic cooperation, and uplifting interactions served as a significant driving force behind the successful completion of this research endeavor. His affectionate concern for students and his in-depth appreciation of the problem and timely advice for resolving challenges have rightfully positioned him as an exemplary mentor. I hold Dr. Nandagiri in the highest regard as an exceptional researcher, captivating orator, and a person of utmost integrity and discipline. To me, he embodies the epitome of a distinguished educator and serves as an exemplary role model.

I would also like to convey my sincere thanks to my Research Progress Assessment Committee members: Dr. B. M. Dodamni from the Department of Water Resources and Ocean Engineering, and Dr. P. Srihari from the Department of Electronics and Communication Engineering for their insightful feedback and constructive criticism, which significantly enriched the quality of this research work.

I express my gratitude to all faculty members, with special mention to Dr. G. S. Dwarakish, the late Dr. Paresh Chandra Deka, and Dr. Ramesh H, for their invaluable insights during my coursework. Additionally, I extend my thanks to the HoD, non-teaching and office staff of the Department of Water Resources and Ocean Engineering at NITK for their unwavering support and provision of necessary facilities essential for the completion of my Ph.D. research work.

I am deeply thankful to my friends (*Jaffar, Atmanand, Doley, Subodh, Tushaar, Late Krithi, Nagbushan, Anand. H, Shashibushan, Niranjan, Pramod, Venkanagouda, Arunkumar, Sahaj, Nithyanand, Sachin, Pandu, Sumanth, Mallikarjun & Avinash. R*) and my PG-students (*Nataraj, Vinod, & Nandini*) for their unwavering support, intellectual engagement, and camaraderie throughout the challenges and triumphs of my doctoral journey. Their friendship has been a constant source of joy and strength, enriching my experience profoundly.

I am profoundly grateful to my cherished family for their boundless love, unwavering encouragement, and countless sacrifices. Their steadfast belief in me has served as a constant anchor, and I am deeply thankful to my beloved late father, Shri. Gurusiddhayya, my mother, Mahadevi, my graceful siblings - Jyoti (*including Mama, Shreya & Shrujan*), and Sneha (*including Kumar, Duggu & Chikku*). I also extend my heartfelt thanks to my loving children, *Amogh and Pranav*, for their sweet support.

Last but certainly not least, I want to express my heartfelt appreciation to my wife & close friend- *Vijaya*. Her boundless love, patience, and unwavering belief in me have sustained me through the highs and lows of this academic pursuit. Her sacrifices and unwavering support have been the cornerstone of my success, and I am forever indebted to her.

To all those who have contributed in ways seen and unseen, thank you. This work would not have been possible without your support and encouragement. Above all, it's the most gracious and the most merciful Almighty God for his help to enable me to complete this research work. I bow my head thanking and praying *Mannikeri Mahanteshwar, Godachi Veerabhadreshwara, & Channamala Shivacharya Swamiji Torgal.*

- Chandrashekarayya G. Hiremath

DEDICATED

TO

MY BELOVED FATHER

~ And to his sacrifices, hardships and blessings

(Late Shri. Gurusiddhayya V. Hiremath)

ABSTRACT

Streamflow Prediction in Ungauged Basins (PUB) is imperative in developing countries such as India due to the sparse network of river gauging stations. Notwithstanding this limitation, accurate estimates of streamflow still need to be derived at all locations where either water development projects have to be constructed or optimal water management policies have to be implemented. Such activities require information on the Flow Duration Curve (FDC) and time-series of historical streamflows, which may however be unavailable if the location of interest is ungauged, i.e., historical discharge records are non-existent. While estimates of certain hydroclimatic variables (rainfall, temperature, soil moisture, etc.) at unsampled locations may be derived using spatial interpolation techniques, deriving estimates of streamflow at ungauged locations is not as straightforward and requires implementation of more sophisticated and involved procedures. In particular, a procedure known as Hydrologic Regionalization (also known as Top Down or Darwinian approach) is commonly used for this purpose and involves transfer of historical streamflow information from gauged stations which are hydrologically similar to the ungauged basin of interest.

The present research was taken up with the specific objective of developing a comprehensive Hydrologic Regionalization-based methodology for deriving the FDC and streamflow time-series in ungauged basins located in the peninsular region of South India. The steps involved in the adopted methodology included: 1) compilation of a dataset comprising historical records of streamflow, climate data and catchment attributes of 50 gauged catchments in the study area, 2) delineation of the catchments into hydrologically homogeneous groups using a hierarchical agglomerative cluster analysis and evaluating the accuracy of the analysis, 3) deriving period-of-record FDCs for each catchment from historical streamflow records, 4) regionalization of 9 flow quantiles of the FDC by establishing multiple linear regression (MLR) relationships with relevant catchment

attributes and evaluating the prediction accuracies of the developed MLR models, 5) cluster-wise evaluation of the reliabilities of the developed MLRs when applied to ungauged basins using jackknife cross validation procedures, 6) application and calibration of 3 popular conceptual rainfall-runoff models (ABWM, SIMHYD, Tank) in the 50 gauged catchments and evaluating their relative performances in simulating daily streamflow time series, 7) cluster-wise identification of dominant model parameters for use in ungauged basins in the region.

For the purpose of this study, 50 catchments with largely unregulated flows located in Krishna, Cauvery, Godavari, East and West flowing river basins situated in South India were identified and a dataset of historical daily streamflow records was created for each of them. Depending on the consistency of the available data, the longest discharge record was from 1991 to 2018 (28 years), and the shortest was from 2000 to 2009 (10 years) for the selected catchments. SRTM-derived Digital Elevation Models were created for the purpose of delineating the catchment boundaries upstream of the gauging stations. The DEMs were also used to derive 15 physiographic attributes for each of the catchments. India Meteorological Department (IMD) gridded data products of historical daily rainfall (0.25^0x0.25^0) and daily maximum and minimum air temperatures (1^0x1^0) for the period 1981-2018 were obtained and used for carrying out the trend analysis. In addition, rainfall data and reference evapotranspiration (derived from gridded temperature data) were employed as inputs for the rainfall-runoff modeling. The Theissen polygon method was applied to ascertain the daily average values for all four climatic variables: rainfall, maximum and minimum temperatures, and reference evapotranspiration.

A necessary first step in hydrologic regionalization is to group the gauged catchments into hydrologically homogeneous groups/clusters so as to ensure accurate information transfer to ungauged basins located within them. Therefore, a hierarchical agglomerative cluster analysis utilizing Ward's linkage method was implemented using carefully selected catchment attributes as clustering variables. This resulted in the grouping of the 50 gauged catchments into three homogeneous clusters with Cluster 1 comprising 17 catchments, followed by 11 catchments in Cluster 2, and 22 catchments in Cluster 3. Through the CV test and L-Discordancy measure using the L-Moment ratio, it was confirmed that all three clusters exhibited homogeneity without any discordant stations.

A Flow Duration Curve (FDC) depicts the relationship between the percentage of time (or duration) for which a particular magnitude of discharge is equalled or exceeded at a particular gauging site. It is a valuable hydrological tool in the planning and design of water resources projects and therefore the present study focused on its estimation at ungauged locations. Initially, considering the historical records of daily discharge, frequency analysis was used to derive period-of-record FDCs for each of the 50 gauged catchments. Subsequently, nine flow quantiles representing the discharge magnitudes at durations of 10%, 20%....90% were extracted from the FDCSs of each catchment. The regionalization approach was then adopted, whereby using step-wise regression, each flow quantile was separately related to the catchment attributes through multiple linear regression (MLR) equations. Performances of the developed MLR models were evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and percentage bias (PBIAS) statistics. Cluster-wise performance analysis of the developed MLR models indicated excellent performance with average R^2 values of 0.85, 0.97, and 0.8 for Cluster – 1, 2, and 3 respectively in comparison to poor performance when all 50 stations were considered to be in a single region. In an effort to evaluate the reliabilities of the developed MLR models when applied in ungauged catchments, a cluster-wise leave-one-out jackknife cross validation procedure was implemented. Results indicated mixed performances with regard to the reliability of developed models with performance being good for high flow quantiles and poor for low flow quantiles.

Since historical records of streamflow time-series are also essential at ungauged sites, the present study also focused on the regionalization of conceptual rainfall-runoff (RR) models. For this purpose, the Australian Rainfall-Runoff Library (RRL) toolkit was identified from within which three popular conceptual RR models namely, AWBM, SIMHYD and Tank were selected for use in this study. Using inputs of catchment average rainfall and reference evapotranspiration the 3 lumped RR models were applied separately to each of the 50 gauged catchments for the periods for which historical discharge records were available using a daily time step. The in-built Shuffled Complex Evolution (SCE-UA), Pattern Search Multi-Start (PS-Multi), and Rosenbrock optimization routine in RRL was used to calibrate the models using both split-sampling and full-period approaches and optimal model parameter values were obtained for each of the 50 catchments. Performance evaluation of the models was carried out using coefficient of determination (R^2), Nash-Sutcliffe efficiency (NSE), and Kling-Gupta efficiency (KGE) metrics. Results indicated that the AWBM model exhibited satisfactory performance (average NSE 0.52) in 88% of the basins, surpassing the SIMHYD and Tank models, which showed satisfactory performance (average NSE 0.51) in 80% of the basins. Variabilities of the parameters of the best performing AWBM model across the catchments in each of the homogeneous clusters were examined and cluster-wise median values were extracted. It was assumed that these could be considered the optimal parameter values for application of the AWBM model in ungauged catchments located in a given cluster and thereby circumvent the need for model calibration using gauged flow records. The efficacy of this assumption was tested by applying the AWBM with the median parameter values in 3 gauged catchments located one each in the clusters. The performance of the model (without calibration) was evaluated and found to perform reasonably well in all 3 test catchments with NSE values of 0.55, 0.82, and 0.65 being obtained. This implies that the AWBM model with optimal model parameters derived through regionalization can yield reasonably accurate daily streamflow time-series in ungauged catchments in the study area.

Overall results of this study indicate that hydrologic regionalization using historical flow records from a reasonably large number of unregulated catchments and possessing diverse attributes can lead to the development of reasonably accurate models for predicting flow duration curves and streamflow time series in ungauged catchments.

Keywords: Ungauged basins, Flow Duration Curve, Hierarchical cluster analysis, Regionalization, Multiple Linear Regression, Jack-knife Cross Validation, Conceptual rainfall-runoff models, Hydrological modeling

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PUB	Predictions in Ungauged Basins
FDC	Flow Duration Curve
MLR	Multiple Linear Regression
GIS	Geographical Information System
RR	Rainfall Runoff
RRL	Rainfall-Runoff Library
CRCCH	Cooperative Research Center for Catchment Hydrology
AWBM	Australian Water Balance Model
HRS	Himalayan River System
PRS	Peninsular River System
SRTM	Shuttle Radar Topography Mission
DEM	Digital Elevation Model
USGS	US Geological Survey
EROS	Earth Resources Observation and Science
WRIS	Water Resources Information System
CWC	Central Water Commission
IMD	India Meteorological Department
SPSS	Statistical Package for the Social Sciences
CV	Coefficient of Variation
MAXe	Maximum Elevation
MINe	Minimum Elevation
ΔH	Basin Relief
$\Delta H/P$	Relief Ratio
S	Basin Slope
A	Basin Area
P	Basin Perimeter
L	Basin Length
W	Basin Width

L_p	Longest Flow Path
FF	Form Factor
SF	Shape Factor
R_c	Circulatory Ratio
R_L	Elongation Ratio
D_D	Drainage Density
RMSE	Root Mean Square Error
PBIAS	Percentage Bias
KGE	Kling-Gupta efficiency metric
ET	Evapotranspiration
IET	Impervious ET
PF	Pervious Fraction
PT	Pervious Threshold
ImI	Impervious Incident
IET	Interception ET
PI	Pervious Incident
RISC	Rainfall Interception Store Capacity
IC	Infiltration Capacity
ICo	Infiltration Coefficient
IS	Infiltration Shape
SMF	Soil Moisture Fraction
IR	Interflow Runoff
IntCo	Interflow Coefficient
IAF	Infiltration After Interflow
RC	Recharge Coefficient
SCE-UA	Shuffled Complex Evolution-Developed by the University of Arizona
PS-Multi	Pattern Search Multi-Start
NSE	Nash–Sutcliffe Efficiency

CHAPTER 1

INTRODUCTION

1.1 GENERAL

Water is a vital resource and its management is crucial for the economic development of all nations. However, ecological and hydrological systems in most of the river basins of the world have been disturbed by large-scale human interventions. The risks associated with floods and droughts are increasing, billions of people worldwide are facing the problems of insufficient and insecure water supplies, and the ongoing destruction of riparian ecosystems has resulted in a steady decrease in biodiversity. These impacts seem to be true in many developing countries that face the issues of insufficient hydrometric data, together with problems of land use changes and climatic change impacts, resulting in decreased water availability and degradation of the ecosystem (Sivapalan et al. 2003a). The ultimate solution to these problems can be advanced through the quantitative understanding of catchment-scale hydrological process (Blank and Delleur 1968).

New research has begun to deal with the precise reconstruction of historical river flows, required for sustainable environmental management and also to supply crucial data to projects designed to control and manage natural disasters such as droughts and floods (Montanari et al. 2006). Hydrological studies have mainly focussed on the enhanced physical interpretation of hydrological processes and the development of mathematical techniques required for the management of water resources (Franchini and Pacciani 1991). Time-series information related to river flow is essential for the success of water resources management studies such as assessment of water availability for domestic supply and irrigation, forecasting of floods and droughts, assessing the ecosystem health, and analysis and design of water resources projects (Vogel et al. 1999; Masih et al. 2010; Karki et al. 2023).

In comparison with rainfall data, river flow data in most of the countries are often limited and are hardly accessible for the specific rivers under investigation. Hence, the need to estimate the river flow discharges for managing the water resources has motivated a great deal of research (Xu and Singh 1998). Streamflow information is neither available nor sufficient in terms of quality or quantity resulting in many catchments being classified as ungauged. Despite substantial progress in hydrological research, many countries continue to struggle with problems of insufficient hydrological data. Such issues are difficult to overcome when estimating flows from an ungauged or inadequately gauged basin (Sivapalan et al. 2003a), resulting in improper planning and management of water resources not only at the ungauged site but also at the river basin level (Masih et al. 2010). Thus, the prediction of streamflow in unmonitored basins holds practical importance in planning and managing water resources and has been recognized as a critical research topic by the international hydrologic community (Sivapalan et al. 2003a; Qamar et al. 2016; Guo et al. 2020; Karki et al. 2023).

Therefore, streamflow Prediction in Ungauged Basins (PUB) is imperative in developing countries such as India due to the sparse network of river gauging stations. Notwithstanding this limitation, accurate estimates of streamflow still need to be derived at all locations where either water development projects have to be constructed or optimal water management policies have to be implemented.

1.2 UNGAUGED BASINS

Ungauged basins represent hydrological catchment areas devoid of direct monitoring or gauging stations to record essential hydrological data, such as streamflow, rainfall, and other critical parameters. The absence of historical data within these basins poses significant challenges in understanding their hydrological behaviour and characteristics (Sivapalan et al., 2003a; Blöschl et al. 2013). Sivapalan et al. (2003) suggest that "*an* *ungauged basin is one with inadequate records (in terms of both data quantity and quality) of hydrological observations to enable computation of hydrological variables of interest (both water quantity or quality) at the appropriate spatial and temporal scales, and to the accuracy acceptable for practical applications*". Such a situation complicates efforts to assess flood risks, water resource availability, and overall basin dynamics. Hence understanding ungauged basins is crucial for water resources management, especially in regions where these basins might constitute a significant portion of the overall water supply. Given the significance of ungauged basins, the International Association of Hydrological Sciences (IAHS) initiated the Decade on Predictions in Ungauged Basins (PUB) 2003-2012 to advance methodologies and techniques for estimating hydrological behaviour in areas lacking direct data. Some of the key issues related to ungauged basins include:

- The absence or limited availability of crucial hydrological data in ungauged basins restricts the development of accurate predictive models and understanding of basin characteristics.
- Models developed using data from one or more gauged basins might not be applied directly to ungauged basins due to differences in hydrological behaviour, land use, and environmental conditions (Wagener et al. 2004).
- Ungauged basins often exhibit significant spatial variability in terms of topography, soil types, and climate patterns, and temporal variability in precipitation, evapotranspiration, and other hydrological processes making it challenging to extrapolate data or models from neighbouring gauged basins (Blöschl et al. 2013).
- Non-availability of data and reliance on indirect estimation methods contribute to uncertainties in flow predictions, affecting water resource management and risk assessment in ungauged basins (Refsgaard and Knudsen 1996).

Water resources management project activities require information on the Flow Duration Curve (FDC) and time-series of historical streamflow, which may however be unavailable if the location of interest is ungauged, i.e., historical discharge records are non-existent. Estimates for certain hydroclimatic variables such as rainfall, temperature, soil moisture, etc., can be obtained for unsampled locations through spatial interpolation techniques. However, estimating streamflow at ungauged locations is not as straightforward and requires implementation of more sophisticated and involved procedures. In particular, a procedure known as Hydrologic Regionalization, also known as Top Down or Darwinian approach (Hrachowitz et al., 2013) is commonly used for this purpose and involves transfer of historical streamflow information from gauged basins that exhibit hydrologically similarity to the ungauged basin of interest.

1.3 FLOW DURATION CURVE AND REGIONALIZATION

The Flow Duration Curve (FDC) of a catchment provides a concise, yet complete description of the runoff regime and therefore its prediction in ungauged basins is considered important (Boscarello et al. 2016; Ma et al. 2023; Yang et al. 2023). The FDC represents the relationship between stream discharge and the percentage of time (duration) (D) that this discharge (Q_D) was equalled or exceeded in the period of record. It has wide applications in the field of water resources assessment and management which include the determination of the abstractable volume of water from rivers for domestic, irrigation, and hydropower projects, evaluation of low-flow statistics to maintain the water-quality standards, flood frequency analysis, wetland inundation mapping, reservoir and lake sedimentation studies and instream flow assessment studies (Fennessey and Vogel 1990; Vogel and Fennessey 1994; Yu et al. 2002; Qamar et al. 2016; Silva 2019; Gaviria and Carvajal 2022; Yang et al. 2023). FDC is one of the most commonly adopted technique for the prediction of flows through regionalization (Boscarello et al. 2016; Qamar et al. 2016). It is for this reason that researchers have devoted significant efforts on the prediction of FDC in ungauged basins using the hydrologic regionalization approach.

Regionalization is one of the commonly used techniques for the analysis of flow characteristics in the ungauged catchments by utilizing the information from one or more gauged stations located within the same hydrological homogenous region (Blöschl and Sivapalan 1995; Sivapalan et al. 2003; Li et al. 2010; Bao et al. 2012; Yang 2017; Guo 2020). In this approach, certain characteristics of the observed FDC derived from historical flow records in gauged basins are transferred to ungauged basins located within hydrologically homogeneous regions (Panthi et al. 2021). Information transfer is achieved by establishing relationships between the flow quantiles extracted from observed FDC and selected catchment characteristics for the gauged basins. The developed relationships are then used to derive flow quantiles/ FDC for the ungauged basins using data pertaining to their catchment characteristics.

While not all studies consider the delineation of hydrologically homogenous regions for regionalization, research (e.g., Yu and Yang 1996; Burgan and Aksoy 2020) has demonstrated that in doing so, increased accuracy in information transfer can be achieved especially if substantial spatial variability in the hydrologic or physiographic features of the catchments exists (Isik and Singh 2008).

1.4 CONCEPTUAL HYDROLOGICAL MODELING

Hydrological models are indispensable tools for analyzing hydrological processes at the watershed scale. Currently, such models are used extensively for various water resource management studies (Zhang et al. 2014). Hydrological models were first developed in the year 1940 for the assessment of regional water resources. Since then, such models have been adopted, modified, and used to solve different hydrological problems (Bergstrom 1991; Xu and Singh 1998).

Hydrological models can be classified as empirical, conceptual, and physically-based (or mechanistic) depending on the adopted structure (Devi et al. 2015; Xu and Singh, 1998). Compared to the empirical models, conceptual and mechanistic models are extensively adopted due to their ability to address a wide spectrum of hydrological problems. Physically based models are distributed models majorly used to assess the effect of watershed transformations on the runoff process. The Present-day physical-based model represents the mechanistic approach demanding the specification of extensive data that are not usually available for large geographical regions (Sivapalan et al. 2003b). In comparison with physically-based models, the conceptual models are less complex with respect to model structure and require comparatively less data. A conceptual model intends to adopt robust approaches allowing the modeler to reduce the issues of over-parameterisation, parameter uncertainty, and equifinality (Perrin et al. 2001; Montanari et al. 2006). A typical conceptual model framework is shown in Figure 1.1.

Literature reveals that the construction of a reliable model is not an easy task due to the complex physical laws and built-in space–time heterogeneity of the hydrological system (James and Burges 1982; Sivapalan et al. 2003b). As per Beven (2006), the testing of different hypotheses in a hydrological system will result in over-parameterisation and equifinality problems. Model complexity often leads to uncertainty (Montanari et al. 2006; Perrin et al. 2001) and inhibits the identification of the dominant physical controls on streamflow variability (Atkinson et al. 2002).

Therefore, scientific research has frequently indicated the importance of simple models for determining the dominant hydrological behaviour of the watershed under investigation (Montanari et al. 2006). The inclusion of additional processes and variables must be considered, to address the causes of variability and to obtain a greater level of prediction. However, such inclusions may lead to over parameterization increasing the model complexity. Hence, the research focus should be on developing the capability to produce quantitative insight into the cause of water flow variability at different scales of interest, rather than formulating a complex mathematical conceptualization of the hydrological system (Farmer et al. 2003; Sivapalan et al. 2003b). In this context, a conceptual hydrological model would be the better choice for the analysis and prediction of flows in ungauged basins.

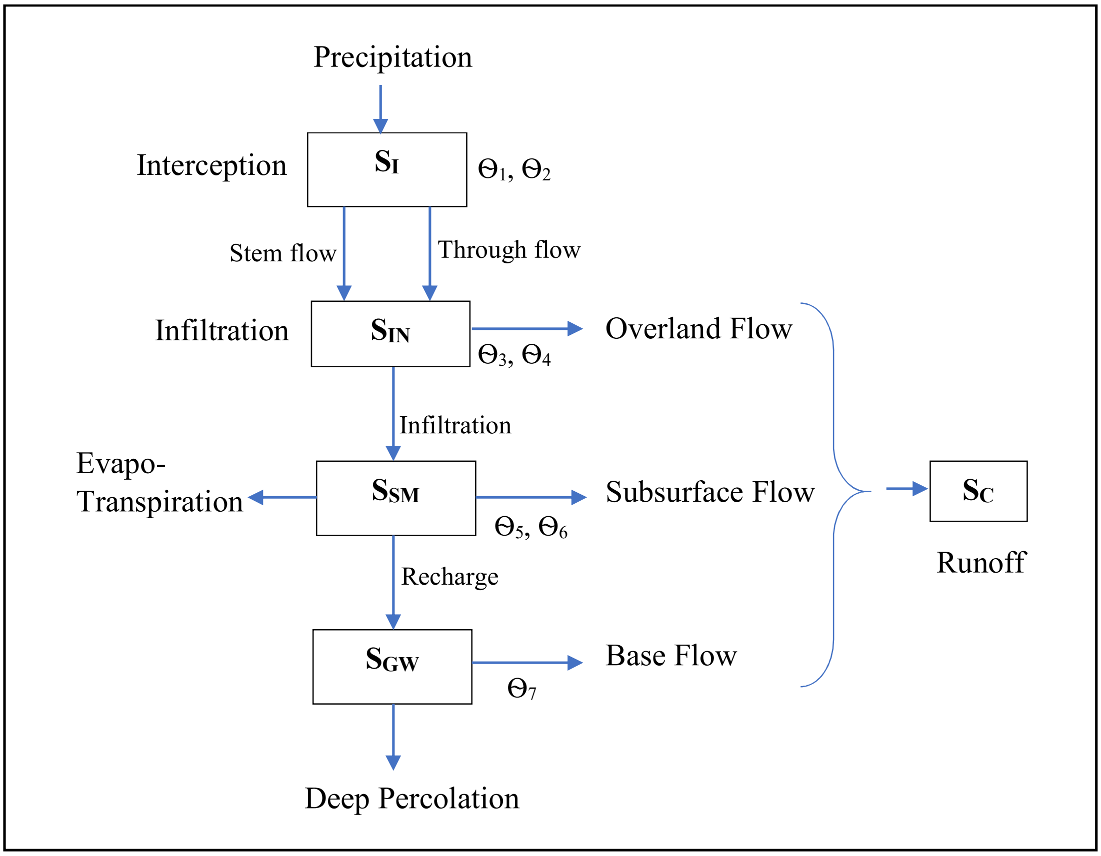


Figure 1.1: Framework of Conceptual model.

1.5 JUSTIFICATION FOR THE STUDY

Water resources management-related studies are very much essential to tackle the issues of water scarcity, inaccessibility, and management of natural disasters. Such studies majorly focus on the quantitative understanding of hydrological processes and their variability at the watershed level particularly in the ungauged basins. Runoff estimation is an essential component of water resources management and its estimation has the following significance:

- Ungauged basins suffer from issues related to data scarcity, model transferability, spatial and temporal variability, and uncertainties in predictions (Refsgaard and Knudsen 1996; Wagener et al.2004; Blöschl et al. 2013). Hence, as recognized by the scientific community advanced methodologies and techniques are essential for predicting flows in ungauged basins (Sivapalan et al. 2003a).

- Reliable prediction of river flow data is necessary for the analysis of catchment yield, for expanding the river flow series, to infill the missing discharge data, and for investigation of different hydrological processes (Chiew and McMahon 1994).
- Understanding the transformation of rainfall into runoff at the catchment level has forever been a main objective of research in surface hydrology and will continue so in the probable future mainly because- hydrological events like floods and droughts are caused due to uncertainties connected with climatic and hydrological data (Edijatno et al. 1999; Ahmad and Simonovic 2005; Post and Jakeman 1999)
- The growing population and increasing economic activities along the flood plains and rivers have increased the significance of accurate runoff prediction, especially in the regions that suffer from insufficient hydrometric data. Apart from sustainable water resources management application, the sensible runoff estimation will also assist in reducing the damages to life and property caused by hydrological events. (Ahmad and Simonovic 2005; Mathevet et al. 2006).

This research study is motivated by the need to assess the flow variability at the watershed-scale from the perspective of control exercised by various landscape and meteorological parameters. The biggest problem in successfully applying the conceptual model in ungauged basins is the difficulty in estimating the hydrological parameters (Xu 1999). This study attempts to develop tools for regionalization of model parameters by exploring relationships to physiographic characteristics. Also, it explores the effectiveness of a conceptual hydrological model for simulating river flow from a wide variety of catchments with unregulated flow conditions and with different climatic and physiographic areas. The study focuses on the analysis of unregulated basins, as these basins are essential for determining the natural relationship among different hydrological processes that are unaffected by human interventions. The study is believed to be a contribution towards a greater understanding of the hydrology of catchments located in major South Indian River basins. This would be achieved by using the regionalization technique and the conceptual hydrological models that have the capability to regionalize watershed models and are adequate to determine the dominant processes accountable for runoff generation in South Indian River basin catchments and their interactions appropriate to this region.

1.6 SCOPE OF THE STUDY

The main aim of the research work is to regionalize the conceptual model parameters in unregulated river basins of South India using conceptual models. The overall scope of this research study involves:

- Identification of unregulated gauged basins by cross-verifying the information provided by the competent authority regarding the details of the major and minor water resources projects upstream of the basins.

- Downloading the appropriate digital elevation maps of the study and analyzing these maps for delineating the identified unregulated basins. The characteristics of each of the delineated basins are then calculated and used for the regionalization process.

- Grouping of the identified basins into homogenous regions and subsequent regionalization utilizing the flow and basin characteristics.

- The historical data about climate and river flows for the delineated basins will be collected and analyzed, and further used as input in the development of FDCs and conceptual hydrological modeling.

1.7 OBJECTIVES OF THE STUDY

This research study has the following objectives:

- To identify and delineate river basins in South Indian region having unregulated flow conditions. Characterize stream flow variability using frequency analysis and develop Flow Duration Curves (FDC).
- To develop tools for regionalization of FDCs by exploring relationships to physiographic characteristics.
- To investigate the performances of identified rainfall-runoff models on a large sample of catchments located in different hydro-climatic regions of South India.
- To determine the model components and associated hydrological processes that are dominant at the catchment level.

1.8 OVERVIEW OF THE RESEARCH METHODOLOGY

The present research was taken up with the specific objective of developing a comprehensive Hydrologic Regionalization-based methodology for deriving the FDC and streamflow time-series in ungauged basins located in the peninsular region of South India. The steps involved in the adopted methodology included: 1) compilation of a dataset comprising historical records of streamflow, climate data and catchment attributes of 50 gauged catchments in the study area 2) delineation of the catchments into hydrologically homogeneous groups using a hierarchical agglomerative cluster analysis and evaluating the accuracy of the analysis 3) deriving period-of-record FDCs for each catchment from historical streamflow records 4) regionalization of 9 flow quantiles of the FDC by establishing multiple linear regression (MLR) relationships with relevant catchment attributes and evaluating the prediction accuracies of the developed MLR models 5) cluster-wise evaluation of the reliabilities of the developed MLRs when applied to ungauged basins using jackknife cross validation procedures 6) application and calibration of 3 popular conceptual rainfall-runoff models (ABWM, SIMHYD, Tank) in the 50 gauged catchments and evaluating their relative performances in simulating daily streamflow time series 7) cluster-wise identification of dominant model parameters for use in ungauged basins in the region. The overall methodology adopted for the present study is shown in Figure 1.2.

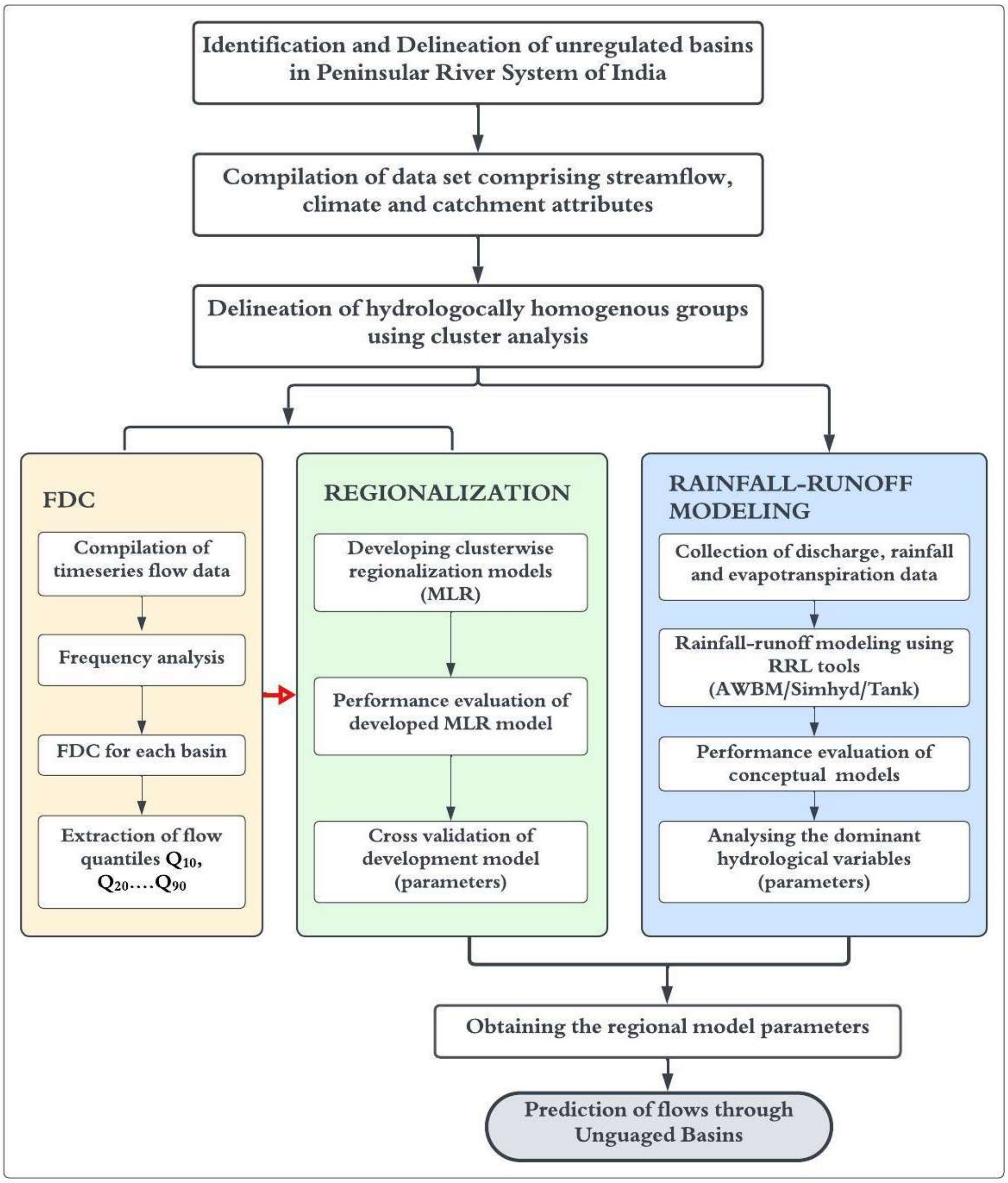


Figure 1.2: Flowchart of the overall research methodology adopted for the study

1.9 ORGANIZATION OF THE THESIS

Chapter 1: This chapter provides the preamble to the research work explaining the significance of time series information for managing the water resources related project, about problems related to the analysis of ungauged basins, prediction of flows in ungauged basins through regionalization concept and using conceptual hydrological model process. This chapter tries to justify the study followed by the scope of work, the objectives of the study, and an overview of the methodology followed.

Chapter 2: This chapter gives an overview of the previous research works carried out in the areas of regionalization, multiple regression analysis, multivariate clustering analysis, rainfall-runoff modeling, model complexity, comparison of physical and conceptual models, and comparative analysis of the different hydrological models.

Chapter 3: Details of the study area, its location, climatic conditions, topographical features, and discharge characteristics have been discussed in this chapter. The chapter also explains the description of the various data used in the research work along with the details about delineated catchment characteristics.

Chapter 4: This section explains the delineation of the identified unregulated basins in south peninsular India and how the delineated basins are clustered into hydrological homogenous groups. The details of delineated basins along with cluster analysis outcomes are explained in the results and discussion section of this chapter.

Chapter 5: The research study focuses on the prediction of flows in ungauged basins using regionalization models using FDC. The concepts and methodology adopted for developing FDC for each basin and cluster-wise regionalization have been elaborated in this chapter. The outcomes of the chapter in the form of regionalized flow duration curves along with their cluster-wise performances have been explained in the results and discussions section.

Chapter 6: The research also undertakes conceptual rainfall-runoff modeling using the RRL toolkit for analysis of flows in ungauged basins. This chapter provides the description of the conceptual models utilized, calibration and validation technique along with performance metrics used for evaluating the models. The outcomes in the form of model performance and the dominant hydrological parameters obtained from the analysis are explained in the results and discussion section.

Chapter 7: This chapter summarizes the research findings of the study, highlighting the outcomes obtained from regionalization techniques and conceptual hydrological modeling. Further, the chapter also provides the scope for future enhancement of this research.

CHAPTER 2

LITERATURE REVIEW

2.1 GENERAL

Though the hydrologic cycle is a system that is fairly easy to grasp and understand, it is more challenging to compute the processes in the system. Hence literature review of hydrological studies is a fundamental aspect of the research studies that not only provides a comprehensive understanding of existing knowledge but also offers insights, guidance, and knowledge gaps, avoids redundancy, and provides directions for advancing research in the field of prediction of flows in ungauged basins.

Previous research works in the areas of hydrologic regionalization, flow duration curve, multivariate regression analysis, clustering techniques, rainfall-runoff modeling, model complexity, comparison of physical and conceptual models, and comparative analysis of different conceptual hydrological models have been reviewed in this study. Following are the findings gathered from the literature review.

2.2 FLOW DURATION CURVE AND REGIONALIZATION CONCEPT

2.2.1 Flow Duration Curve

Flow Duration Curve (FDC) is a graphical representation to depict the distribution of streamflow over a specific period. This curve illustrates the percentage of time the specific flow rate has equaled or exceeded. FDC provides valuable insight into the variability and frequency of flow rates within a river system and helps hydrologists in various ways to manage the available water resources (Vogel and Fennessey 1995; Yu et al. 2002; Qamar et al. 2018; Silva 2019; Gaviria and Carvajal 2022; Yang et al. 2023). FDC is one of the most commonly adopted technique for the prediction of flows through regionalization (Boscarello et al. 2016; Qamar et al. 2018). It is for this reason that researchers have devoted significant efforts on the prediction of FDC in ungauged basins using the hydrologic regionalization approach (Panthi et al. 2021). For example, Yu and Yang (1996) assessed regional FDC for Southern Taiwan using multivariate statistical analysis of flow data from 34 sites. Castellarin et al. (2004) developed regional FDC for 51 unregulated river basins in Italy using catchment and morphologic characteristics. Mohamoud (2008) developed a regression model for different exceedance probabilities of flows in more than 40 climatic and landscape regions of the Northeastern US. A regionalization study by Shu and Ouarda (2011) indicated that FDC based method outperformed the area-ratio method in 109 stations of Quebec Provinces in Canada. A recent study by Qamar et al. (2018) presented a nonparametric regionalization procedure for assessing FDC for 124 catchments of Northwestern Italy. Releasing the wide hydrological application of FDC, Chouaib et al. (2019) predicted the daily FDC through regionalization in ungauged basins using the hydroclimatic data of 73 catchments in the eastern USA. Similarly, the regionalization of FDC was demonstrated by Panthi et al. (2021) for predicting streamflow values for the data scares region of the central Himalayas. Considering FDC as a crucial indicator of river basins, Ma et al. 2023 attempted to identify the best-fit function using the regression analysis concept for developing FDC in a semi-arid region of North China. Karki et al. (2023) evaluated the uses and limitations of the regionalization method for developing FDC in 23 medium to small-sized watersheds across Nepal.

2.2.2 Regionalization Concept

The reliable estimation of hydrological quantities in a poorly gauged or ungauged basin is considered to be one of the major concerns in hydrological research studies due to insufficient data needed for calibration and validation (Sivapalan et al. 2003; Jin et al. 2009; Masih et al. 2010; Wagener and Wheater 2006). The formation of an initiative named-Prediction in Ungauged Basins (PUB) by the International Association of Hydrological

Sciences (IAHS) indicates the need to focus research in this area (Sivapalan et al. 2003; Jin et al. 2009). A comprehensive review of the PUB initiative conducted by Hrachowitz et al. (2013) offers a thorough understanding of the significance of hydrological forecasts in basins lacking gauge data and addresses the challenges related to upscaling and downscaling of hydrological processes across spatial and temporal scales. The study explores the methods of regionalization and transferability of hydrological models to address the issues related to the complexity involved in extrapolating predictions from gauged to ungauged basins.

As per Wagener and Wheater (2006), the prevalent method for modeling the ungauged basins is to find a functional relationship between parameters of the conceptual model and basin characteristics in a statistical manner assuming that the distinctiveness of each catchment can be analyzed in a unique combination of catchment characteristics. Such statistical relationships are known as regional models and a typical model structure provided by Wagener and Wheater (2006), is shown in equation 2.1:

$$\hat{\theta}_L = H_R(\theta_R|\phi) + \upsilon_R \qquad (2.1)$$

where $\hat{\theta}$ is the estimated model parameter at the ungauged basin,

H_R(.) is a functional relation for $\hat{\theta}_L$ using a set of physiographic and meteorological catchment characteristics ϕ,

θ_R is a set of regional model parameters and,

υ_R is an error term

This approach of developing relationships is referred to as regionalization or conventional schemes or spatial generalization (Seibert, 1999; Wagener and Wheater, 2006; Bastola et al. 2007). Such regionalization approaches involve partly or completely transferring of parameters from neighbouring gauged basins to the basin of interest (Merz and Bloschl, 2004; Jin et al. 2009).

Regionalization is one of the commonly used techniques for the analysis of flow characteristics in the ungauged catchments by utilizing the information from one or more gauged stations located within the same hydrological homogenous region (Blöschl and Sivapalan 1995; Sivapalan et al. 2003a; Li et al. 2010; Bao et al. 2012; Hrachowitz et al., 2013; Yang 2017; Guo 2020). Xu & Singh (1998), Post & Jakeman (1999), and Xu (1999) have presented reviews on studies that have used regionalization methods to relate rainfall-runoff model parameters to watershed characteristics. The capability to regionalize watershed models can be improved, when hydrologists begin to conceptualize and model regional physical relationships between the parameters of watershed models and the unique characteristics of each watershed, otherwise the regionalization studies will continue to give mixed results. (Vogel et al. 2005; Post and Jakeman 1999)

Regionalization can be carried out by different techniques, viz. regression analysis, area-index, nearest neighbour method, and hydrological similarity method. Among these, the multiple linear regression (MLR) analysis is one of the earliest and most widely used techniques globally for regionalization (Li et al. 2010; Bao et al. 2012; Swain and Patra 2017). This technique aims to develop a relationship between identified catchment characteristics and stream flow information corresponding to gauged catchment through a multiple linear regression equation and has the advantage of evaluating each model parameter independently (Parajka et al. 2005; Yang et al. 2017).

Various researchers have adopted such multivariate regression analysis for analyzing the ungauged catchments. For example, Bao et al. (2012) compared the regionalization approaches based on regression and similarity methods in 55 catchments of China. Vogel et al. (1999) developed the regional regression model relating the hydrologic, geomorphic, and climatic characteristics of a large number of catchments across the United States. Cluster-wise regionalization analysis was investigated by Li et al. (2018) for 15 catchments located in the Yangtze and Yellow River basins of China. An attempt was made by Huang et al. (2015) to integrate the regression concept of regionalization with clustering analysis, and it turned out to be effective in the Yalong River Basin, China. Compared to other methods, the regression approach has more advantages that include- integration of catchment and stream flow characteristics in GIS processing, analysis of climate change impacts on water yields, and most importantly, this concept can be used to quantify the mean and variance of the stream flow for any catchment in the region (Vogel 1999).

As indicated by Seibert (1999), developing the relationship between optimized parameter values and basin characteristics is a challenging task compared to a simple extension of runoff series in time. The regionalization process might be affected if model parameters are not well defined and are uncertain. This supports the conclusions made by Wagener and Wheater (2006), and Bastola et al. (2007), that parameter uncertainty may hamper the identification of model parameters representing certain processes and therefore hinder the regionalization process (Merz and Bloschl 2004).

2.3 MULTIVARIATE CLUSTER ANALYSIS

While not all studies consider the delineation of hydrologically homogenous regions for regionalization, research (e.g., Yu and Yang 1996; Burgan and Aksoy 2020) has demonstrated that in doing so, increased accuracy in information transfer can be achieved especially if substantial spatial variability in the hydrologic or physiographic features of the catchments exists (Isik and Singh 2008). Prior to regionalization, a homogeneous group of catchments is created by grouping the catchments into clusters according to the characteristics of the variables within the clusters (Isik and Singh 2008; Yu and Yang 1996; Rao and Srinivas 2006).

In the past few decades, different methods for the delineation of homogenous regions using a variety of similarity measures have been proposed (Tasker 1982; Rao and Srinivas 2006; Nobert et al. 2011; Latt et al. 2014; Boscarello et al. 2016, Li et al. 2018; Javadinejad 2021; Song et al. 2022) among which multivariate cluster analysis has proved to be the most efficient one. For example, Yu and Yang (1996) defined homogenous regions using cluster analysis for developing FDC for 34 stream-gauged stations in Southern Taiwan. Stream flow information from 655 gauging stations in Columbia was studied by Gaviria and Carvajal (2022) and they delineated 15 homogenous regions using geological, topographic, and climatic information as clustering variables and K-mean algorithm for grouping. An agglomerative hierarchical clustering algorithm was used by Burn et al. (1997) to define homogeneous regions for regional flood frequency analysis in the Saskatchewan-Nelson River basin in west-central Canada.

The hierarchical clustering approach was adopted by Boscarello et al. 2016 for classifying 46 catchments in the Upper Po River basin in northwest Italy into three homogenous groups that were further used to estimate FDCs using the regionalization technique. Petrakis et al. 2021 used a hierarchical clustering approach to classify sub-basins of Smith Canyon Watershed, USA based on 12 environmental variables related to structural, biophysical, and hydrologic traits. Owing to the simplicity and widespread use of the hierarchical clustering method, Mulaomerović-Šeta et al. 2023 adopted hierarchical clustering for grouping the basins and subsequently predicted the flood quantiles in ungauged basins of West Balkans using the regionalization concept. Similarly, various other studies on identifying homogenous regions using cluster analysis have been carried out by Mosley (1981), Shaban et al. (2010), Goyal and Gupta (2014), Abdolhay et al. (2012), Latt et al. (2014), Li et al. (2018) and Riswandi et al. (2022).

Multivariate clustering techniques are typically classified into hierarchical and flat clustering. The usefulness of hybrid cluster analysis (combination of hierarchical and flat clustering) in regionalization was demonstrated by Rao and Srinivas (2006) for watersheds in Indiana, USA. Hierarchical clustering is preferred over flat clustering when the number of clusters is unknown.

2.4 RAINFALL-RUNOFF MODELING

The essence of hydrology is modeling. A model is a mathematical statement of the response of a system that takes system inputs and transforms them into system outputs (Dawdy 1983). The hydrological models have turned out to be an essential tool for analyzing hydrological processes at the watershed scale. Such model tools are used extensively for various water resource management studies- including reservoir operation, flood routing and inundation prediction, estimation and analysis of stream flow changes, and impact assessment of land cover changes on water flow (Dhami and Pandey 2013; Moradkhani and Sorooshian 2009).

Studies by Bergstrom (1991), and Xu and Singh (1998) demonstrate the use of a simple water balance model for the assessment of regional water resources by hydrologists and agriculturists. These models make use of available climate data and other related physical parameters to generate stream flow hydrographs. Hydrological models were initially developed in the year 1940 for the assessment of regional water resources. Since then, such models have been adopted, modified, and used to solve different hydrological problems. Crawford and Linsley (1966) developed one of the notable complex rainfall-runoff model known as the Stanford Watershed Model to analyze the dynamic of hydrologic processes prevailing in a catchment. Additional examples of conceptual hydrological models include the Xinanjiang Model and the Sacramento Soil Moisture Accounting Model (SAC-SMA), which is an extensively used operational model in the US National Weather Service for flood forecasting (Moradkhani and Sorooshian 2009).

However, soon it was realized that the physical laws that govern these water flows and storages are so complex and associated parameters are so variable in space and time that construction of a reliable model is no easy matter (James and Burges, 1982). The extrapolation of data from gauged to ungauged catchments using any type of model is a major concern with considerable difficulties and uncertainties due to the presence of the large spatio-temporal heterogeneity of landscape and climatic properties (Sivapalan et al. 2003).

2.4.1 Model Complexity and Equifinality

In surface water hydrology, the hydrologic system happens to be distributed, non-linear, and time-variant in its behaviour. This behaviour of the system is the most complex one and quite difficult not only to formulate mathematically but also to solve. Studies by Perrin et al. (2001) indicate that the parameterization of processes is essential to represent different hypotheses about the hydrological system resulting in increased model complexity. This is in line with arguments by Beven (2006), which elaborates that- the testing of different hypotheses in hydrological systems will result in over-parameterization and equifinality problems. The equifinality thesis as discussed by Beven (2006) is intended to focus consideration on the fact that numerous satisfactory representations cannot be easily discarded and should be taken into account for determining the uncertainty linked with predictions. Examples of distributed rainfall-runoff modeling and groundwater modeling suggest that equifinality is prevalent in the problem (Beven 2006).

Complex models tend to be less robust and less stable than simpler ones (Perrin et al. 2001). Complex models with large calibration and validation data requirements are usually not advisable, when the data availability is either inadequate or incomplete (Sivapalan et al. 2003b; Montanari et al. 2006). Further substantial uncertainty is created during the assessment of parameters belonging to complex models having little or no physical meaning (Perrin et al. 2001; Montanari et al. 2006). This supports the study by Atkinson et al. (2002), in which a simple bucket model was used for developing a hypothetical relationship between timescale, climate characteristics, and model complexity for four watersheds in New Zealand. The developed relationship helps the modeler eliminate the use of unwanted parameterizations by selecting a simple but effective model. Figure 2.1 indicates that- the complexity of the model increases with decreasing timescale, and increasing climate characteristics (dryness index).

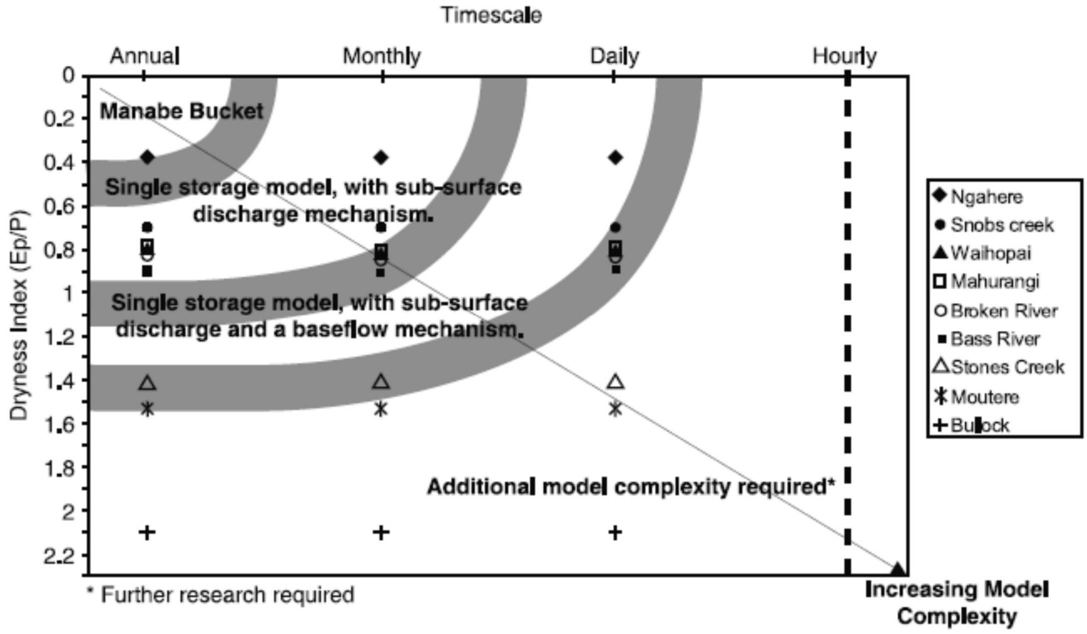


Figure 2.1: Association between timescale, climate characteristics, and model complexity *(Source: Atkinson et al. 2002)*

2.4.2 Conceptual Models

In comparison with physically-based models, simple conceptual models that lump catchment heterogeneity and represent the conversion of precipitation into river flow conceptually are usually easy-to-use tools with low data requirements (Xu and Singh 1998; Perrin et al. 2001). Montanari et al. (2006) describe conceptual models as simple models, that intend to adopt robust procedures allowing the modeler to reduce the risk of over-parameterization and cause fewer issues of parameter uncertainty, which is of the most importance (Perrin et al. 2001).

The problems of equifinality as observed in physical models are reduced in conceptual models due to their comparatively simple structure and restricted number of parameters and therefore minimize the uncertainty during the analysis of the main hydrological processes occurring within the basin (Braun and Renner 1992; Montanari et al. 2006). Such models have proved efficient as reported by Perrin et al. (2001) and are certainly useful for water managers and engineers. A few of the model applications include- flood forecasting and reservoir operation, time-series extension of runoff records, estimation of design floods, and hydrochemical transport (Bergstrom 1991; Xu 2002). The features of hydrological models, as discerned from the research work conducted by Xu and Singh (1998), as well as Beven (1989) through the application of lumped conceptual models include:

- They explain conceptually the land-based hydrologic processes which are spatially averaged or lumped. The equations used can only offer an approximate depiction of the actual world, inevitably introducing errors stemming out from the model structure.
- Estimation of some of their model parameters is carried out by fitting them with observed hydrological data, such as rainfall and streamflow.
- The problem of over-parameterization may arise if one tries to stimulate every hydrological process deemed suitable and strives to match all those parameters through optimization against observed data. Experiences from the studies (Perrin et al. 2001) suggest that three to five parameters must be adequate to reproduce information in a hydrological record.

Although, the simplification of natural process complexity can be carried out by conceptual models, but it is apparent that such representations are far from reality (Uhlenbrook et al 1999). Further, the calibrating of such models necessitates abundant hydrological and meteorological data, which might not always be accessible. Their calibration involves curve fitting, making it challenging to interpret the physical significance of the obtained parameter values (Beven 1989).

Currently, there is a rising trend to exercise new models in hydrology that offer distributed information on catchment characteristics. Although such a model might help understand the hydrological processes, but has restrictions when applied in an operational context (Xu and Singh 1998; Perrin et al. 2001). Various hydrological conceptual models have been developed across the world for rainfall-runoff modeling including the Tank model (Sugawara et al. 1983); ARNO model (Todini 1996), Australian Water Balance Model

(AWBM) (Boughton 2004), Simhyd (Chiew et al. 2002); Sacramento (Brazil and Hudlow 1981), etc. Rainfall-Runoff Library (RRL) is a catchment runoff model available under the eWater Toolkit development by the Co-operative Research Center for Catchment Hydrology (CRCCH), eWaters Australia. The RRL includes five models, AWBM, SIMHYD, Tank, etc, to simulate the runoff using rainfall and evapotranspiration data of the basin. Apart from calibration and validation activity, the RRL models allow evaluation of different model types, display wettest and driest years, and can carry out parameter sensitivity analysis (Podger 2004).

2.5 DOMINANT PARAMETER ANALYSIS

For any model form, there exist some unknown constants known as parameters, which are used to represent the physical process. These model parameter values must be estimated in such a way they must result in the best agreement between modeled and observed runoff (Duan et al. 1992; Xu 2002). Parameters are a part of the model and have no meaning outside the model. If the modeler builds a physically/conceptual-based model, then the parameters are abstractions that may approximate some physically meaningful quantity (Dawdy et al. 1968). Most of the models in hydrology are based on the conceptualization of physical processes that govern the flow of water. Two types of parameters usually exist in such models: one is physical and the other one is process parameters (Sorooshian and Gupta 1995).

- Physical parameters: represent physically measurable watershed properties. Examples: watershed area, surface area of open water bodies and streams, surface slopes, and so on.
- Process parameters: represent watershed properties that are not directly measurable. Examples: lateral inflow rate, effective depth of surface soil moisture storage, and so on (Xu 2002).

As per Atkinson et al. (2002), the dominant parameters refer to a model parameter to which model predictions are very delicate and change when altered over a reasonable range of values. Assessment of such parameters using water balance models (Xu and Singh 1998) improves the understanding of how climate and landscape heterogeneities control the space-time variability of the main hydrological processes in different hydro-climatic regions (Sivapalan et al. 2003b). It is unfortunate that most of the catchments in the world, are either ungauged or inadequately gauged, and the dominant processes controlling their streamflow response are still poorly understood (Montanari et al. 2006).

Franchini and Pacciani (1991) conducted a comparison study on six conceptual rainfall-runoff models including the STANFORD Watershed Model (IV), SACRAMENTO, APIC, TANK model SSARR, ARNO, and Xinanjiang model applied on the Sieve watershed, a tributary stream of river Arno in Italy. The study reveals that the complications developed during the calibration phase were closely associated with the number of model parameters (degree of complexity). To address such issues Post and Jakeman (1999), recommend - the formulation of a modeling framework that focuses on key hydrologic processes taking place in a basin and should reduce the problems of over-parameterization. Similarly, Nash and Sutcliffe (1970) also presented an approach to avoid the issues of excess parameterization by developing a model from a simple type to a complex model. This approach closely matches the downward approach described by Klemes (1983) and tested by Sivapalan et al. (2003). Hence a conceptual model with a simplex structure and a restricted number of parameters is preferred for determining the influential hydrological model parameters.

As per Jothityangkoon et al. (2001), modeling capability to determine the dominant factor governing runoff variability in a region can be improved significantly if precise information on the input data can be generated based on field studies. Furthermore, it is difficult to conduct the sensitivity analysis of the parameter estimates concerning input data or objective function unless the best set of parameters connected with a known calibration data set can be found (Duan et al. 1992).

2.6 RESEARCH GAPS

- Observed conditions can be extrapolated with some confidence and can come up with reliable estimates if the selected hydrological tool is capable of demonstrating the essence of how a watershed operates hydrologically. Hence the ability to understand the catchment system functions is equivalent to the problem of acquiring the predictive power for future conditions in gauged and ungauged basins (Sivapalan et al. 2003a).

- Even today, conducting hydrological studies concerning ungauged basins for establishing rainfall-runoff relationships remains a challenging task. Uncertainties in model parameters and climatic data result in considerable approximation of the present generation of hydrological models which are used to predict streamflow responses in inadequately gauged or ungauged basins. This was also observed by Montanari et al. (2006) who attempted to carry out a hydrological study using the downward approach in tropical regions. Tropical regions (where the South Indian rivers exist) are some of the most challenging and less understood regions of the world, especially with respect to the sources of uncertainties (Montanari et al. 2006). Tropical river systems are important as they serve several wetland ecosystems, drain into different economically viable regions, and also because of the density of population they cover (Panda et al. 2011).

- As discussed in the literature, the basic hydrological prediction method includes a model, a set of parameters, and appropriate meteorological inputs. But, each of these components especially the model parameters and climate data are either identified inappropriately or not known at all, due to the intrinsic multiscale space-time heterogeneity of the hydrological system, particularly in ungauged catchments (Franchini and Pacciani 1991; Sivapalan et al. 2003).

- Stream flow assessment using the precipitation data has always motivated a great deal of research in hydrology as it is a known fact- that there is generally plenty of precipitation data as compared to stream flow data. Stream flow data are usually inadequate and are rarely accessible for the river under examination (Xu and Singh 1998).

- Even after many advances in hydrological studies, the biggest problem for the successful application of rainfall-runoff models is the estimation of hydrological parameters (Xu 1999; Montanari et al. 2006)

- Studies by Atkinson et al. (2002) point out that the guidance required for selecting a model of suitable complexity for stream flow prediction, given the vegetation, climate, soil, and topographical characteristics is insufficient in present-day hydrological studies. It is therefore required to develop a method, which helps in choosing a model of appropriate complexity which will aid in investigating the dominant physical controls on variation of stream flow within the catchment. Such model development will assist the modeler in adopting the best approaches for determining the essential parameter values, with an aim to reduce uncertainty issues and improve prediction accuracy (Atkinson et al. 2002).

- Previous studies by Chiew and McMahon (1994), Xu and Vandewiele (1995), and Perrin et al. (2001) have evaluated the performance of water-balance models on a large number of catchments outside the Indian Peninsular region. Therefore, such an assessment would be the first of its kind from the Indian context and would create a deeper understanding of hydrological system function in tropical space-time heterogenic conditions.

CHAPTER 3

STUDY AREA AND DATA

3.1 GENERAL

Recent research has been supported by modern computation facilities and capabilities, which help in extensively testing the simple models in a large number of catchments, under a range of different geographical, climatological, and geological conditions (Perrin et al. 2000; Bergstrom 1991). Furthermore, the inadequate data sets usually make the responses extremely dependent on the hydro-climatic conditions. As per Perrin et al (2000), the model is said to be more reliable if it is tested under diverse conditions. Hence south Indian rivers having diverse hydro-metrological conditions are selected for this study.

Also, previous studies seem to have explicitly checked whether or not recorded streamflow in the selected gauged catchments is influenced by upstream diversions or regulations due to the presence of dams/reservoirs. This may prove to be a critical issue since the effect of such anthropogenic modifications on the natural runoff regime of the gauged catchments may be transferred to the ungauged basin in the process of regionalization. Therefore, it is imperative to ensure that the selected gauged catchments possess unregulated flows if the hydrological predictions in the ungauged basins are to represent natural conditions. Even though a large number of new water resources projects are being planned in India in general and in South India in particular, few previous studies seem to have been taken up to develop tools to predict streamflow and FDCs in ungauged catchments in this region.

3.2 STUDY AREA

Indian rivers are broadly classified into those belonging to the Himalayan River System (HRS) in the North and those belonging to the Peninsular River System (PRS) in the South. While the Himalayan Mountains form the headwater catchments for rivers in the HRS, rivers in the PRS mostly originate in the Western Ghat Mountains which run parallel to the West Coast of India and travel eastward across the peninsula before emptying into the Bay of Bengal. The major rivers of the PRS are the Tapi, Narmada, Mahanadi, Godavari, Krishna, and Cauvery. The PRS also comprises several short West Flowing rivers which originate in the Western Ghats and flow westward into the Arabian Sea and several short East Flowing rivers which originate in the Eastern Ghats and flow into the Bay of Bengal. Compared to perennial Himalayan Gangetic rivers, the tropical South Indian River flows are seasonal. This seasonality of flow has resulted in the construction of multi-purpose reservoirs to manage water supply problems and to protect flood-prone deltaic regions (Panda et al. 2011). Water resource is said to be scarce if the availability reduces to less than 1000 m^3 per capita per year creating problems for social and economic activities. South Indian rivers like Krishna, Cauvery, Pennar, and Tapi are some of the basins, which come under this category (Integrated Hydrological Data Book, 2017). South Indian rivers are considered important due to the density of the population that they serve and by considering the economically feasible region that they drain (Panda et al. 2011).

The main criterion for selecting the catchments for use in this study was the absence of any major water resources project upstream of the gauging station so that the recorded flows could be considered as being virgin or unregulated flows. The preliminary information about the basins was extracted from the river basin report by the Central Water Commission (CWC), Government of India. Based on the basin report (Integrated Hydrological Data Book, 2019, Government of India) and verification from the India – Water Resources Information System (WRIS) portal (http://india-wris.nrsc.gov.in/HydroObservationStationApp.html), the presence of major flow regulation through dams on upstream side of the river gauge stations was verified and unregulated catchments were identified in major river basins of Peninsular River systems of India.

Accordingly, 50 catchments with largely unregulated flows located in Krishna, Cauvery, Godavari, East and West flowing river basins situated in South India were identified and a dataset of historical daily streamflow records was created for each of them. Table 3.1 shows details of the selected catchments identified by the river basin in which they are located and the names of the gauging stations. Figure 3.1 shows the location map of the selected gauging stations.

Table 3.1: List of selected stream gauging stations

River Basin Name	No. of Stations	Station Name
West Flowing Rivers	18	Santeguli, Avershe, Yennehole, Addoor, Bantwal, Erinjipuzha, Kidangoor, Kalloopara, Thumpaman, Ayilam, Kuniyil, Karathodu, Kalampur, Mahuwa, Haladi, Nanipalasan, Ozerkheda, Pulamanthole
Krishna	08	Kellodu, Talikot, Navalgund, Balehonnur, Khanapur, Marol, Halia, Naguleru @Dachepalli
Cauvery	11	Sakleshpura, K M Vadi, E_Mangalam, Bendrahalli, Hogenakkal, Kudlur, Thevur, Thoppur, Nellithurai, Thengumarahada, T. Bekuppe
East Flowing Rivers	05	Kashipatnam, Seedhi, Ambasamudram, Salur, Gunupur
Godavari	08	Pedagedadda, Ramakona, Wairagarh, Amabal, Tumnar, Cherribeda, Gandlapet, Sonarpal

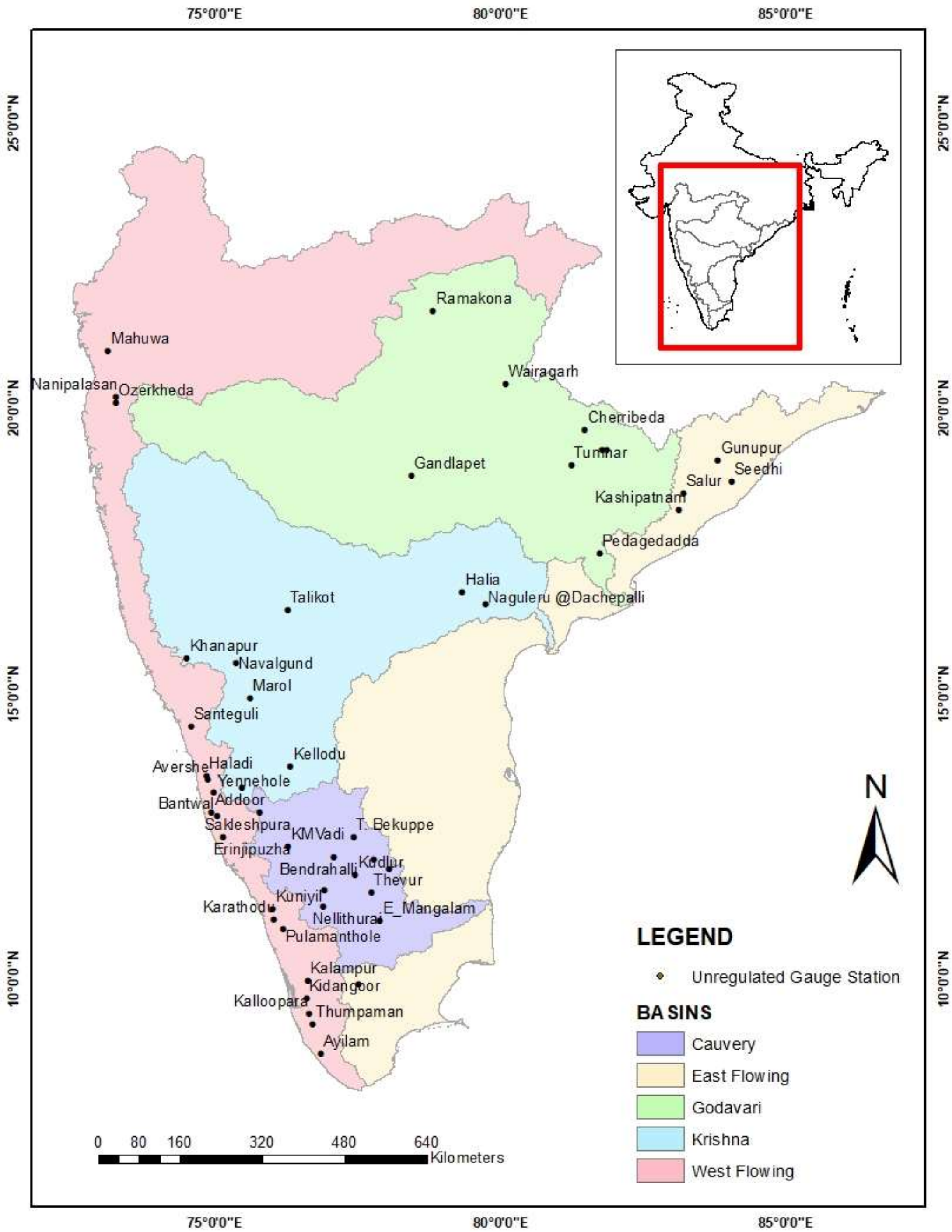


Figure 3.1: Map showing the location of selected stream gauge stations in the Peninsular River System of India

The areas of the delineated catchments upstream of the gauging stations varied from a minimum of 171 km^2 to a maximum of 6930 km^2, with six catchments having catchment areas in excess of 3000 km^2. The topography of the study area consists of hilly terrain in the West and a flat plateau towards the East with the average elevation varying from 251 m to 1404 m. The mean annual rainfall of the selected catchments ranged from a minimum of 564 mm to a maximum of 4029 mm. The major part of river flow occurs during the wet monsoon months from June to September, and the flow is negligible in most of the rivers during the other months. The average daily temperature in the identified catchments ranges between 21 ^{0}C to 31^{0}C.

3.2.1 Basin-wise details of the study area

The basin-wise details of the identified catchments are explained below:

West Flowing River Basin:

A total of eighteen unregulated gauge stations were identified in West flowing river basin with an upstream basin area varying from a minimum of 276 km^2 to a maximum of 3205 km^2. The topography of the study area consists of hilly terrain rolling towards the west side. The climate of the region is humid and gets a good amount of rainfall with a mean annual rainfall of the selected catchments ranging from 1272 mm to a maximum of 4029 mm and the average daily temperature in the identified catchments ranges between 21 ^{0}C to 31^{0}C. West flowing river basin can be divided into two regions- Tapi to Tadri basin and Tadri- to Kanyakumari basin. In this study, four identified stations fall in the Tapi to Tadri basin and the remaining fourteen stations are in the Tadri- to Kanyakumari basin. The details of the location (especially the Indian- State) and the name of the river on which the stations are situated are mentioned in Table 3.2.

Table 3.2: Details of the selected stream gauge stations in West Flowing River basin

Sl. No	Station Name	Area (km^2)	Sub-basin	River Name	Indian-State
1	Santeguli	988	Tapi to Tadri (Vasishti and Others)	Aghnashini/ Tadri	Karnataka
2	Avershe	276	Tadri to Kanyakumari (Netravati and Others)	Sita	Karnataka
3	Yennehole	334		Swarna	Karnataka
4	Addoor	690		Gurpur	Karnataka
5	Bantwal	3205		Netravati	Karnataka
6	Erinjipuzha	852		Payaswami	Kerala
7	Haladi	524		Haladi	Karnataka
8	Kidangoor	593	Tadri to Kanyakumari (Periyar and Others)	Minachil Ar	Kerala
9	Kalloopara	700		Manimala Ar	Kerala
10	Thumpaman	796		Achankovil Ar	Kerala
11	Ayilam	533		Vamanapuram	Kerala
12	Kalampur	358		Kaliyar Puzha	Kerala
13	Kuniyil	1998	Tadri to Kanyakumari (Varrar and Others)	Chaliyar	Kerala
14	Karathodu	770		Kadalundi	Kerala
15	Pulamanthole	904		Tuta Puzha/ Pulantod	Kerala
16	Mahuwa	1701	Tapi to Tadri (Bhatsol and Others)	Purna	Gujarat
17	Nanipalasan	719		Damanganga	Gujarat
18	Ozerkheda	659		Vagh	Maharashtra

Krishna River Basin:

Compared to the previous basin, only eight-gauge stations with no major upstream regulation were selected in the Krishna River basin. The area of the basins upstream of the selected stations varied from a minimum of 518 km^2 to a maximum of 5142 km^2. The topography of the study area comprises rolling and undulating terrain, except at the western Ghats which consists of hilly terrain. Krishna basin has a tropical climate with mean annual rainfall of the selected catchments ranging from 564 mm to a maximum of 2831 mm and the average daily temperature in the identified catchments ranges between 21 ^{0}C to 32 ^{0}C. Krishna river basin has different sub-basins that include Ghataprabha, Malaprabha, Tungabhadra, Vedavathi, upper and lower Bhima, upper, middle and lower Krishna, Musi, Palleru and Munneru basins. In this study, one selected station was situated in the Tungabhadra Lower basin, three stations were situated in Krishna Upper, and two stations each were falling in the Tungabhadra Upper and Krishna Lower basin. The details of the location and the name of the river on which the stations are situated are mentioned in Table 3.3.

Table 3.3: Details of the selected stream gauge stations in the Krishna River basin

Sl. No.	Station Name	Area (km^2)	Sub-basin	River Name	Indian-State
1	Kellodu	4271	Tungabhadra Lower	Hagari/ Vedavati	Karnataka
2	Talikot	2411	Krishna Upper	Don	Karnataka
3	Navalgund	3080		Benni Halla	Karnataka
4	Khanapur	518		Malaprabha	Karnataka/ Maharastra
5	Balehonnur	810	Tungabhadra Upper	Bhadra	Karnataka
6	Marol	5142		Vardha	Karnataka
7	Halia	3250	Krishna Lower	Halia	Telangana
8	Naguleru @ Dachepalli	561		Naguleru Vagu	Andra Pradesh

Cauvery River Basin:

Eleven-gauge stations with no major upstream regulation were selected in the Cauvery River basin with the area upstream of the stations varying from a minimum of 332 km^2 to a maximum of 3465 km^2. The topography of the study area comprises hill ranges of the Western Ghats and gently undulating terrain on the other side. The basin has a tropical and sub-tropical climate with mean annual rainfall of the selected catchments ranging from 772 mm to a maximum of 1795 mm and the average daily temperature in the identified catchments ranges between 20 ^{0}C to 32 ^{0}C. The Cauvery River basin is divided into Upper, Middle, and Lower Cauvery basins. Out of eleven selected stations, two are located in the Upper Cauvery Basin, and the remaining are situated in the Middle Cauvery Basin as detailed in Table 3.4.

Table 3.4: Details of the selected stream gauge stations in the Cauvery River basin

Sl. No.	Station Name	Area (km^2)	Sub-basin	River Name	Indian-State
1	Sakleshpura	601	Upper Cauvery	Hemavati	Karnataka
2	K M Vadi	1454		Lakshmantirtha	Karnataka
3	E_Mangalam	3465	Middle Cauvery	Noyil	Tamil Nadu
4	Bendrahalli	1838		Holemadu halla	Karnataka
5	Hogenakkal	1562		Chinnar	Tamil Nadu
6	Kudlur	720		Palar	Karnataka
7	Thevur	1190		Sarabhanga	Tamil Nadu
8	Thoppur	332		Veppadi/ Toppai Ar	Tamil Nadu
9	Nellithurai	1482		Bhavani	Tamil Nadu
10	Thengumarahada	1357		Moyar	Tamil Nadu
11	T. Bekuppe	3336		Arkavathi	Karnataka

East Flowing River Basin:

Only five-gauge stations with no major upstream regulation were selected in the East Flowing River basin with the area upstream of the stations varying from a minimum of 171 km^2 to a maximum of 6930 km^2. The topography of this region comprises hilly regions of the Eastern Ghats, the plateau region, and the coastal region. Similar to Cauvery, this basin also has a tropical and sub-tropical climate with mean annual rainfall of the selected catchments ranging from 1049 mm to a maximum of 1349 mm and the average daily temperature in the identified catchments ranges between 22 ^{0}C to 32 ^{0}C. In this study, four selected stations are situated between the Mahanadi and Pennar basins, and one is located between Pennar and Kanyakumari basins. The details of the selected stations in east flowing river basin are shown in Table 3.5.

Table 3.5: Details of the selected stream gauge stations in the East Flowing River basin

Sl. No.	Station Name	Area (km^2)	Sub-basin	River Name	Indian-State
1	Kashipatnam	171	Between Mahanadi and Pennar (Nagvati and Others)	Gosthani	Andra Pradesh
2	Salur	308		Peddagedda	Andra Pradesh
3	Seedhi	1116	Between Mahanadi and Pennar (Vamsadhara and Others)	Mahendrata naya	Andra Pradesh
4	Gunupur	6930		Vamsadhara	Odisha
5	Ambasamudram	703	Between Pennar and Kanyakumari (Pamba and Others)	Vaigai	Tamil Nadu

Godavari River Basin:

A total of eight stations with no major upstream regulation were selected in the Godavari river basin with the area upstream of the selected stations varying from a minimum of 181 km^2 to a maximum of 2493 km^2. The river basin encompasses diverse landscapes, including forests, plateaus, and valleys. Godavari basin has a tropical climate with mean annual rainfall of the selected catchments ranging from 1057 mm to a maximum of 1480 mm and the average daily temperature in the identified catchments ranges between 21 ^{0}C to 33 ^{0}C. In this study one selected station is situated in the Godavari Lower basin, two stations are located in the Weinganga basin, four stations are in the Indravati basin and one station is located in the Pranhita basin as mentioned in Table 3.6.

Table 3.6: Details of the selected stream gauge stations in the Godavari River basin

Sl. No.	Station Name	Area (km^2)	Sub-basin	River Name	States Covered
1	Pedagedadda	181	Godavari Lower	Bodapetta Vagu	Andra Pradesh
2	Ramakona	2493	Weinganga	Kanhan	Madhya Pradesh
3	Wairagarh	1832		Satti Nadi	Maharastra
4	Amabal	1943	Indravati	Narangi	Chhattisgarh
5	Tumnar	1732		Dantewara	Chhattisgarh
6	Cherribeda	870		Boroda Nadi	Chhattisgarh
7	Sonarpal	1483		Markandi	Chhattisgarh
8	Gandlapet	1396	Pranhita and Others	Pedda Vagu	Telangana

3.3 DATA

In this study, data related to topography, rainfall, temperature, and observed discharge rates were collected from concerned departments/organizations. The collected data was checked for errors, consistency, and uniformity. For example- short period of missing data was filled in by interpolation. However, when long period of data was missing, the records was not used in the analysis. In case of rainfall data, inconsistent records were checked with neighbouring grid points and corrected. The collected hydrological and meteorological data was compiled and analyzed for each of the identified catchments and the corresponding data set was prepared.

3.3.1 Digital Elevation Model (DEM)

All the identified catchments were delineated using 30m resolution Shuttle Radar Topography Mission (SRTM) - Digital Elevation Model (DEM) data obtained from the US Geological Survey's - Earth Resources Observation and Science (EROS) Center (https://earthexplorer.usgs.gov). SRTM-derived Digital Elevation Models were created to delineate the catchment boundaries upstream of the gauging stations. The upstream of the selected gauge station might cover multiple SRTM-DEM tile maps. Hence, all the associated SRTM-DEM tile maps are downloaded, and further mosaicked and projected before delineating the basins using the location of the stream gauge station as its outlet point.

3.3.2 Discharge data

Daily discharge data for each of the identified stream gauge stations was extracted from the India – Water Resources Information System (WRIS) portal (http://india-wris.nrsc.gov.in/HydroObservationStationApp.html) of the Central Water Commission (CWC), Government of India. Depending on the consistency of the available data, the longest discharge record was from 1991 to 2018 (28 years), and the shortest was from 2000 to 2009 (10 years) for the selected catchments.

3.3.3 Climate data (Rainfall and Temperature)

Information relating to climate is a critical input to hydrological modeling with the quality of historical records relating to rainfall, temperature, etc., largely determining the accuracy of model predictions. Procuring climate data from ground-based measurements is challenging due to the limited number of gauge stations and also because of the low spatial coverage of such stations over large regions. Given the fact that the present study focused on the hydrological modeling of 50 catchments located over the extensive South Indian peninsula, procuring historical records of rainfall and climate variables from ground-based gauges operated by diverse government agencies proved to be difficult.

However, data accessibility has improved in recent times with advancements in remote sensing, in-situ sensors, and better computer facilities. A few of the well-known satellite datasets include the National Center for Environmental Prediction Climate Forecast System Reanalysis (NCEP-CFSR), Global Land Data Assimilation System (GLDAS), Tropical Rainfall Measuring Mission (TRMM), and Climate Hazards Group Infrared Precipitation with Station data (CHIRPS) (Mitra et al. 2013; Setti et al. 2020). For India specifically, the India Meteorological Department, Government of India provides historical gridded climate data that have been developed using spatial interpolation techniques. Among these, the IMD gridded data products pertaining to rainfall and air temperature are the most commonly used data for hydrological modeling, reanalysis, and inter-comparison studies for the Indian region. A comparative study conducted by Meher and Das (2019), indicates that the IMD data are preferred over the Climate Research Unit (CRU) global climate data of the UK for estimating missing climate records.

Therefore, in the present study, India Meteorological Department's (IMD) gridded data products (https://www.imdpune.gov.in/cmpg/Griddata/) of historical daily rainfall (0.25^0x0.25^0) for the period 1981-2018 (Pai et al. 2014) were obtained. Using grid points lying within and close to each catchment, the areal rainfall values were derived using the Thiessen polygon method. Average annual rainfall values 'Rain' (mm) for this historical period were obtained for each catchment and included with the other attributes.

Similarly, daily minimum and maximum temperature data for all the selected catchments were obtained from the IMD's 1^0x1^0 gridded temperature product (Srivastava et al. 2009) for the period 1981 to 2018. Similar to rainfall, the average minimum and maximum temperature at each of the catchments were estimated using the Thiessen polygon method explained later in chapter 6.

3.4 CATCHMENT ATTRIBUTES

Catchment attributes are essential in the development of regional models for the FDCs and also for cluster analysis that aids in determining the hydrological homogenous basin groups. Clustering attributes have a strong influence on the results. Hence appropriate cluster attributes should be identified before grouping the homogenous catchments (Yu and Yang 1996). The DEMs were also used to derive 15 physiographic attributes for each of the catchments that include: Maximum elevation 'MAX_e', Minimum elevation 'MIN_e' (km), Relief 'ΔH' (km), Relative Relief '$\Delta H/P$', Slope 'S', catchment area 'A' (km^2), basin perimeter 'P' (km), length of the basin 'L' (km), basin width 'W' (km), Longest Flow Path 'L_p' (km), drainage density 'D_d' (km/km^2), form factor 'FF', shape factor 'SF', circulatory ratio 'R_c' and elongation ratio 'R_L'. The catchment attributes utilized for this research study were inferred from the research studies conducted by Venkatesan (2014), Sukristiyanti et al. (2018), and, Islam and Deb Barman (2020). A brief description of each attribute follows.

i. Maximum elevation (MAX_e): DEM was utilized to determine the maximum elevation in each of the identified basins. The higher elevation often helps in understanding the drainage path, a potential area of accumulation, and the rate at which the river flows downhill. The maximum elevation of the delineated basins varied between 587m to 2634m.

ii. Minimum elevation (MIN_e): Similar to the previous case, the DEM was utilized to define the minimum elevation of basins. The lower elevation influences various hydrological processes such as runoff, infiltration, and evaporation. The minimum elevation of the delineated basins varied between 0 to 889m.

iii. Relief (ΔH): Basin relief can be defined as the distance between the lowest and highest points of the basin. This helps in understanding the denudational behaviour of the basin. High relief may lead to increased runoff and soil erosion. The relief value in the identified basins varied from 252m to 2621m.

iv. Relative Relief ($\Delta H/P$): Relative relief may be defined as the ratio of basin relief to the basin perimeter. This aspect reflects the overall slope steepness of a drainage basin, serving as an indicator of the erosion intensity occurring on the basin's slope. High values are typical for hilly regions, while lower values are indicative of pediplains and valleys. The average value of relative relief was about 0.004.

v. Slope (S): Basin slope is the ratio of basin relief to the length of the basin. A steeper slope tends to generate more rapid and intense runoff during heavy rainfall. The slope aspect also impacts the erosion rates and channel geometry. The average slope of the selected basins was found to be 0.022.

vi. Catchment area (A): The area of the catchment is calculated using the ArcGIS tool. The area of the delineated basins varied from 171 to 6930 km^2. The catchment area influences the amount of surface runoff and infiltration. This is crucial for analyzing the flood risks and groundwater recharge.

vii. Basin perimeter (P): The basin perimeter of the delineated basins varies from 106km to 801km. The aspects define the geographical extent of the available water resources for allocation/distribution.

viii. Length of the basin (L): The length of the delineated basins as determined using the ArcHydro tool varies from 27km to a maximum of 152km. The longer the basin length longer the time taken for water to move from the upper reaches to outlets. This influences the timing and peak of river flows.

ix. Basin width (W): The width of a basin influences hydrological response to rainfall events. Broader basins typically exhibit longer flow paths, leading to delayed peak flows and modifications in hydrograph patterns. The width of the delineated basins varies from 6km to a maximum of 59km.

x. Longest Flow Path (L_p): The longest flow path often referred to as the longest flow distance is a crucial hydrological parameter that influences the time of concentration that is critical in flood risk assessment. The longest flow path of the delineated basins as estimated using the ArcHydro tool varies from 35km to 205km.

xi. Drainage density (D_d): It can be defined as the ratio of the total length of all streams within the catchment divided by the total area of the catchment. Higher density indicates the presence of a well-established stream network, contributing to effective runoff. The density of the delineated basins in this study varies from 0.29 to 2.38 km/km^2.

xii. Form factor (FF): The form factor is calculated as the ratio of the catchment area to the square of the length of the catchment. Higher form factors indicate a basin with lower flow paths compared to basins with low form factors. The form factor of the delineated basins in this study varied from a minimum of 0.11 to 0.84.

xiii. Shape factor (SF): Basin shape can be categorized as circular, rectangular, or triangular. The shape has a direct influence on both the magnitude and the timing of peak discharge at the basin outlet. For instance, in comparison to an elongated basin, a circular basin tends to exhibit an earlier arrival time for peak discharge. The shape factor can be defined as the ratio of length to width of the basin. The shape factor of the delineated basins of this study varies from 1.2 to 8.7.

xiv. Circulatory ratio (R_c): This factor is affected by the basin length, geological formations, land use/land cover, climate, and the slope of the basin. The low value of the circulatory ratio indicates lesser structural disturbances or control on the watershed behaviour. The circulatory ratio is estimated using the equation 3.1.

$$R_C = \frac{4\pi A}{P^2} \tag{3.1}$$

xv. Elongation ratio (R_L): The elongation ratio in this study was evaluated using equation 3.2. Normally, the elongation ratio is categorized into two classes: a low value indicating an elongated watershed, and a high value signifying a circular watershed characterized by high relief and a rapid response to rainfall. The elongation ratio of the delineated basins in this study varied from 0.3 to 0.65.

$$R_L = \frac{2}{L\left(\frac{A}{\pi}\right)^{0.5}} \tag{3.2}$$

The derived attributes for each of the 50 selected catchments are listed in Appendix – 1.

CHAPTER 4

HYDROLOGIC DELINEATION OF BASINS

4.1 DELINEATION OF CATCHMENTS

As discussed in Article 3.3.1, the boundaries of the identified basins were delineated using 30 m resolution SRTM-DEM data obtained from the USGS. All the associated SRTM-DEM tile maps are mosaicked and projected before delineating the basins using the location of the stream gauge station as its outlet point. The process and delineation of the basins were carried out utilizing the ArcHydro tool in ArcMap 10.1 software (Figure 4.1). ArcHydro consists of toolsets to analyze, edit, and convert data to support GIS implementations in the water resources sector. In order to delineate a catchment, different terrain processing including- flow direction, accumulating flow, defining the stream, segmenting streams, delineating the catchment grid, processing catchment polygons and the drainage line, and processing of drainage point is carried out in the ArcHydro tool.

Further, various catchment characteristics for each of the delineated basins were also worked out using the ArcHydro tool and subsequently utilized in multiple regression analysis. A total of fifteen catchment characteristics/attributes (as discussed in Article 3.4) were determined for each of the 50 identified unregulated basins (Appendix – 1).

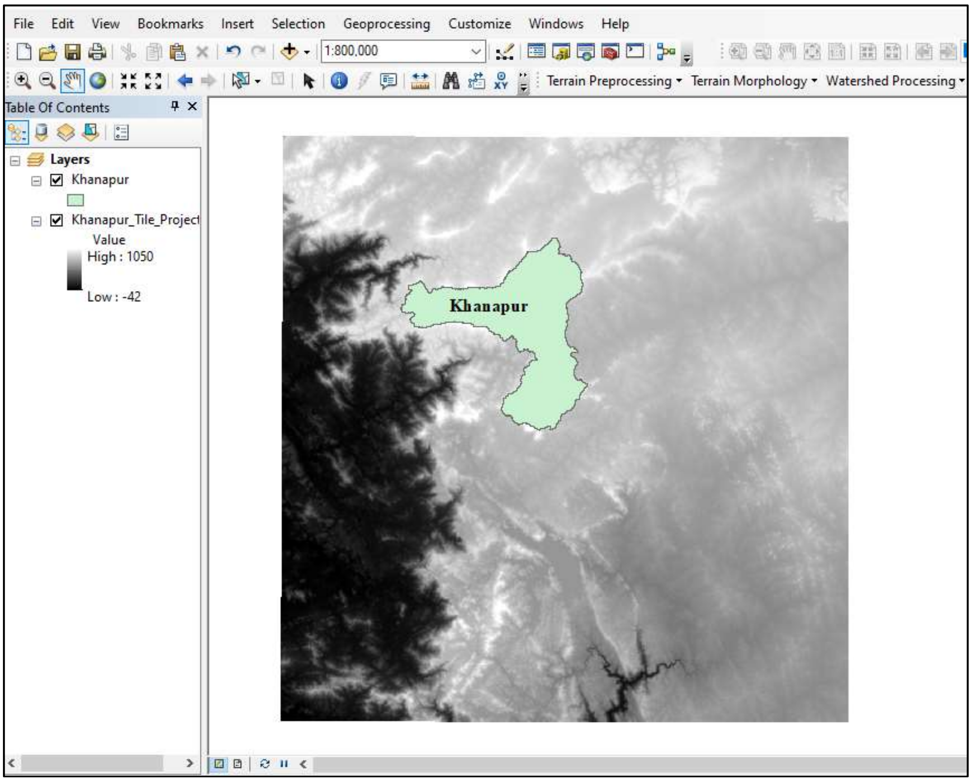


Figure 4.1: Snapshot of Catchment delineation in ArcMap 10.1 software

4.2 CLUSTER-ANALYSIS

4.2.1 Hierarchical Cluster Analysis

Hydrologic delineation refers to the process of classifying the selected gauged catchments into groups which are hydrologically homogeneous. Clustering is a multivariate technique to identify such hydrologically homogenous regions using hydrological or catchment characteristics. Prior to regionalization, a homogeneous group of catchments is created by grouping the catchments into clusters according to the characteristics of the variables within the clusters (Isik and Singh 2008; Yu and Yang 1996; Rao and Srinivas 2006). Clustering techniques are generally classified into hierarchical and flat clustering. The usefulness of hybrid cluster analysis (combination of hierarchical and flat clustering) in regionalization was demonstrated by Rao and Srinivas (2006) for watersheds in Indiana, USA. Hierarchical clustering is preferred over flat clustering when the number of clusters is unknown. Hierarchical clustering works by merging smaller clusters into bigger ones, known as the agglomerative technique, or dividing bigger clusters into smaller ones, known as the divisive technique (Rao and Srinivas 2006; Javadinejad 2021). Hierarchical clustering is typically displayed using a tree-like figure recognized as a “dendrogram” of clusters that explains the organization of the clusters. As per Demirel (2004), the user must decide the number of clusters to be formed, as the dendrogram will not provide the cluster assignment details. Since the length of the dendrogram’s limb denotes the proximity of points, data can be clustered by cutting the dendrogram at a desired level (Isik and Singh 2008; Boscarello et al. 2016).

Previous research has used hierarchical clustering techniques, such as single linkage, complete linkage, centroid, average distance, and Ward’s minimum variance technique for hydrologic regionalization (Li et al.2018). For example, Tasker (1982) adopted a complete linkage algorithm to regionalize watersheds in Arizona. Comparison of different algorithms, including single, complete, and average linkage, centroid, median, and Ward’s method, was demonstrated by Nathan and Mcmahon (1990) using the Statistical Package for the Social Sciences (SPSS) tool. Similarly, Burn et al. (1997) used an agglomerative hierarchical clustering algorithm to regionalize watersheds in Canada. Ward's method outperformed the other methods in terms of separation so that clusters are relatively dense with low variability within groups (Boscarello et al. 2016). Hence Ward’s method was adopted in the present study to group the catchments into hydrologically homogenous clusters using hierarchical agglomerative algorithms.

Irrespective of the type of clustering scheme, a similarity measure is required to classify individual catchments into homogenous groups. Out of the different similarity measuring techniques, the most commonly adopted Euclidean distance (Tasker 1982; Yu and Yang 1996; Isik and Singh 2008) is used as a similarity measure in this study. The Euclidean distance is defined as

$$D_{P,Q} = \sqrt{\sum_{i}^{j} (X_{Pi} - X_{Qi})^2} \tag{4.1}$$

Where $D_{P,Q}$ is the distance between two catchments, P and Q. X_{Pi} and X_{Qi} is the ith attribute at catchments P and Q, and j is the total number of selected attributes.

4.2.2 Identification of Clustering Variables

Clustering attributes/variables have a strong influence on the results. Hence appropriate cluster attributes/variables should be identified before grouping the homogenous catchments (Yu and Yang 1996). The cluster analysis can categorize groups based on discharge data, and topographical and meteorological characteristics (Rao and Srinivas 2006; Boscarello et al. 2016). Further, it is sensible to include those attributes that are not highly correlated with each other (Boscarello et al. 2016). In this context, cluster analysis of 50 catchments in the present study was carried out with the identified clustering attributes in the SPSS statistical package tool using Euclidean distance as a similarity measure and Ward’s technique for linkages.

4.2.3 Homogeneity Test and Discordance Measure

As per the method demonstrated by Nobert et al. (2011), the homogeneity within the catchment groups derived from cluster analysis is assessed using the Coefficient of Variation (CV) test in this study. This test involves the calculation of mean, standard deviation, and CV of daily rainfall (times series data used for cluster analysis) at each catchment of the study area. The regional average coefficient of variation (CV_{Avg}) and standard deviation of CV (σ_{CV}) of the river flow information is given as

$$CV_{Avg} = \sum_{i=1}^{N} \frac{CV_i}{N} \tag{4.2}$$

$$\sigma_{CV} = \sqrt{\frac{\sum_{i=1}^{N} (CV_i - CV_{Avg})^2}{N}} \tag{4.3}$$

Where CV_i is the coefficient of variation at the ith catchment, and N is the number of catchments used in the study. A region is regarded to be homogenous if the homogeneity measure (CC) defined by Equation 4.4 is less than or equal to 0.3.

$$CC = \frac{\sigma_{CV}}{CV_{Avg}} \tag{4.4}$$

Averaging the CV across catchments can help smooth out extreme outliers, which might otherwise skew the clustering process. Outliers or highly variable data points can distort the true relationships between catchments, making it harder to identify meaningful cluster. The average CV captures the typical level of variability across different variables and catchments, allowing the clustering algorithm to focus on the most significant patterns of similarity between catchments, rather than on isolated extremes. The clusters so formed are more likely to represent the overall hydrological behavior rather than being overly sensitive to random fluctuations in variability.

Further, the discordance measure aims to recognize discordant catchment within the group that need to be adjusted or excluded to improve their homogeneity. In order to do so, a discordance method proposed by Hosking and Wallis (1993) has been adopted in this study. The discordance measure utilizes the advantages offered by sampling properties of L-moment ratios. L-moments represent specific linear combinations of order statistics, offering greater resilience to outliers compared to conventional moments. They are less affected by sample variability and prove beneficial in deriving meaningful insights about the underlying probability distribution, especially when dealing with small sample sizes (Hosking 1990). This study used the R-Studio software package to determine the initial four L-moments (L_1, L_2, L_3, and L_4) for each identified catchment. The L-moment ratios are defined as

$$\tau = \frac{L_2}{L_1} \tag{4.5}$$

$$\tau_r = \frac{L_r}{L_2}, when \ldots r \geq 3 \tag{4.6}$$

Where τ is the measure of scale and dispersion, $\tau_3$and $\tau_4$are measures of skewness and kurtosis, respectively.

If $v_i=[\tau^{(i)}, \tau_3^{(i)}, \tau_4^{(i)}]^T$ be the vector containing the L-CV, L-Skeweness, and L-Kurtosis value related to the catchment 'i'. Then the discordancy measure for the catchment 'i' is given as:

$$D_i = \frac{1}{3} N_k (v_i - \bar{v})^T S^{-1} (v_i - \bar{v}) \tag{4.7}$$

$$\bar{v} = \frac{1}{N} \sum_{i=1}^{N} v_i \tag{4.8}$$

Where S is the sample covariance matrix given as

$$S = \sum_{i=1}^{N} (v_i - \bar{v}).(v_i - \bar{v})^T \tag{4.9}$$

Higher values of D_i indicate the most discordant catchment in the group. As per Hosking and Wallis (1993), the catchment identified as discordant should be examined thoroughly, as discordancy may result from sampling variability or changes in the attribute values due to localized extreme events. Irrespective of the discordance value, the statistical parameters of the catchment need to be compared with other catchment stations within the group before declaring a particular catchment as discordant.

4.3 RESULTS AND DISCUSSIONS

4.3.1 Delineation of Catchments

A total of 50 basins were identified as unregulated after reviewing the CWC basin reports and cross-verifying with the WRIS web portal. The selected basins were delineated using SRTM-DEM data on the ArcMap platform. Next, the catchment characteristics were estimated for each of the delineated basins using the ArcHydro extension tool in ArcMap software. The computed catchment characteristics were utilized for the regionalization process.

4.3.1.1 Delineation in West-Flowing River Basin

Eighteen unregulated gauged basins in west flowing river were delineated as shown in Figure 4.2. The average elevation of the delineated basins varied from a maximum of 1475m to a minimum of 15m. The delineated basins had a varied length ranging from 101112m to 26806m and basin width ranging from 37481 to 7084 m. The average value of the form factor in the west-flowing river is 0.37, with the majority of basins indicating rectangular to irregular forms of basin leading to delayed flow intensity in general. The average value of the shape factor is 3.4, the circulatory ratio is 0.18, and the elongation ratio is 0.43 representing an asymmetric or less compact basin. The average drainage density of the west-flowing basins (1.38 km/km^2) is the highest among all other basins indicating a denser network of water courses that respond quickly to rainfall events.

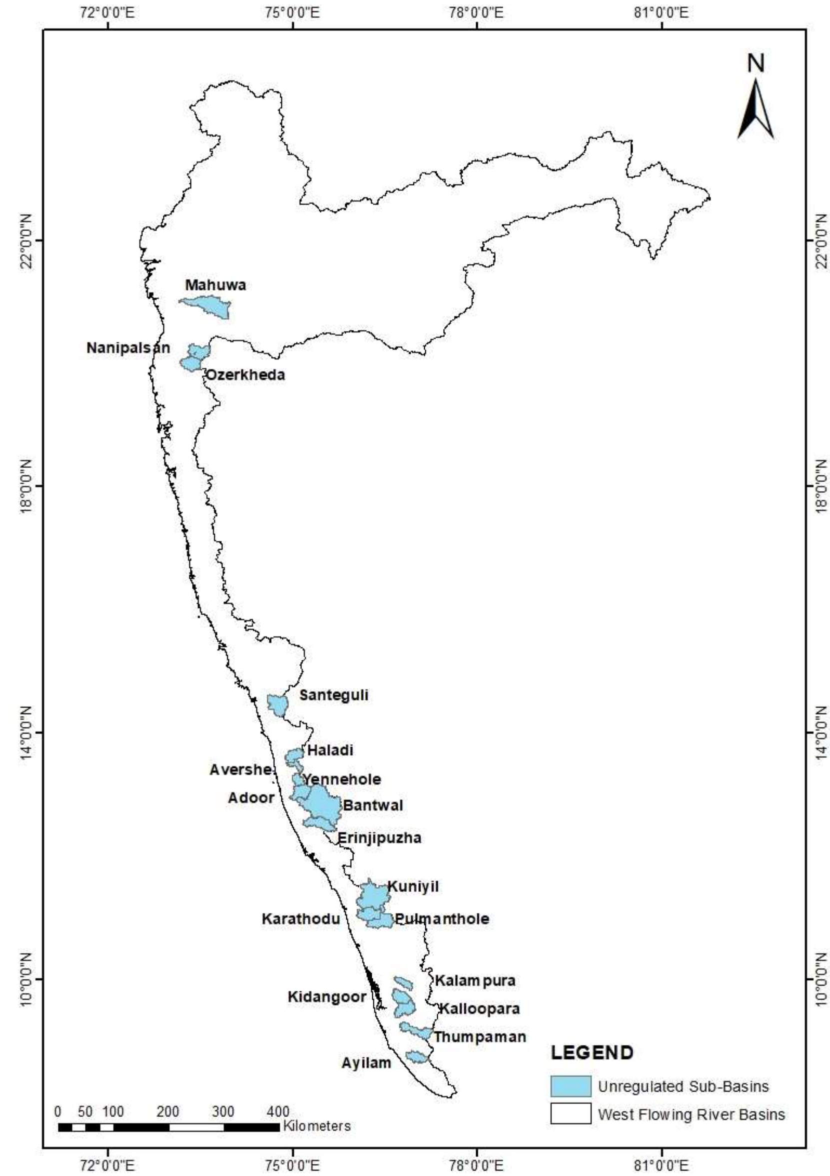


Figure 4.2: Delineated unregulated basins in the west-flowing river basin

4.3.1.2 Delineation in Krishna River Basin

The delineated basins in the Krishna River encompassed eight gauged areas, depicted in Figure 4.3. Their average elevation ranged from a maximum of 1072 m to a minimum of 470 m. These basins displayed diverse dimensions, with lengths spanning from 144565 m to 39215 m and widths ranging between 45492 m to 11844 m. The average form factor of the basins in Krishna River measured 0.31 indicating an irregular form of basin leading to delayed flow intensity in general. Their average shape factor stood at 3.75, while the circulatory ratio was 0.15, and the elongation ratio was calculated at 0.485, signaling an asymmetric or moderately elongated basin structure. Furthermore, the average drainage density of these basins was recorded at 0.98 km/km^2, suggesting a moderately dense network of watercourses that typically respond in a normal manner to rainfall events.

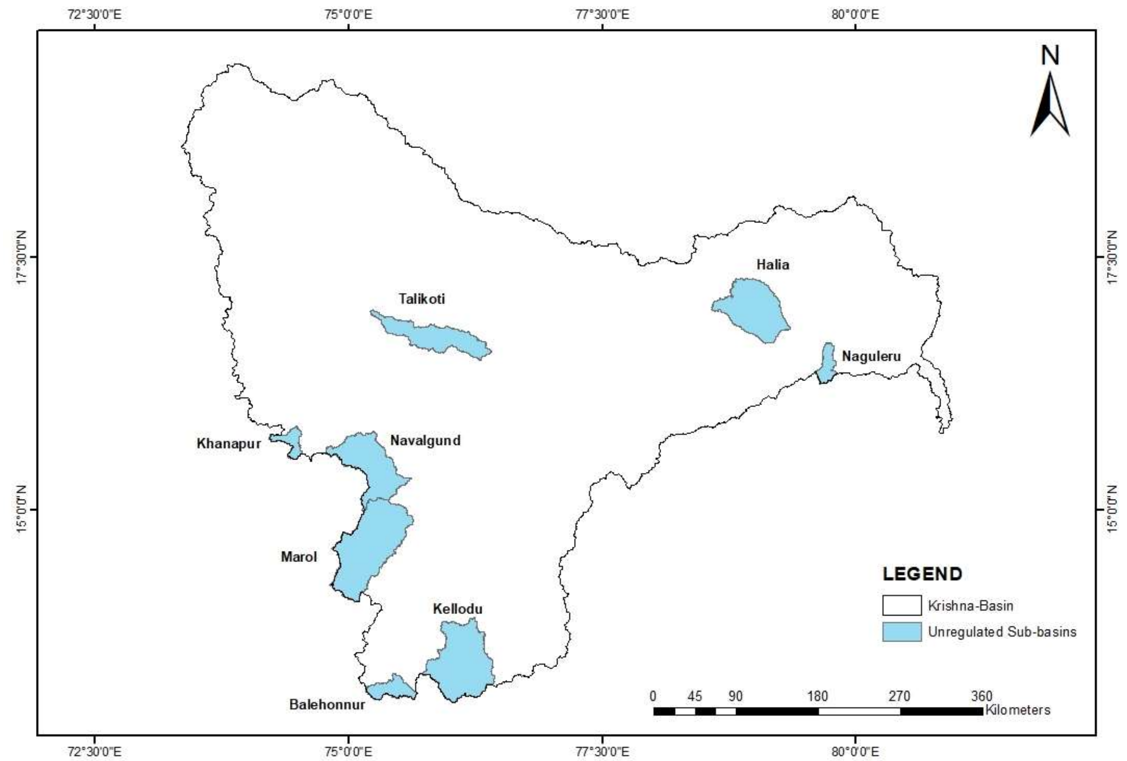


Figure 4.3: Delineated unregulated basins in the Krishna river basin

4.3.1.3 Delineation in Cauvery River Basin

Eleven gauged basins within the Cauvery River were delineated, as illustrated in Figure 4.4. The average elevation of these delineated basins ranged from a maximum of 1805 m to a minimum of 444 m. These basins exhibited diverse dimensions, with lengths spanning from 151769 m to 32717 m and widths ranging between 28292 m to 10151 m. The average form factor of these basins was measured at 0.33, with the majority indicating rectangular to irregular basin shapes. The average values for the shape factor, circulatory ratio, and elongation ratio were 3.47, 0.17, and 0.45, respectively, portraying an asymmetric basin structure. Additionally, the average drainage density of these basins was recorded at 0.95 km/km^2, suggesting a moderately dense network of watercourses that generally responds in a normal manner to rainfall events.

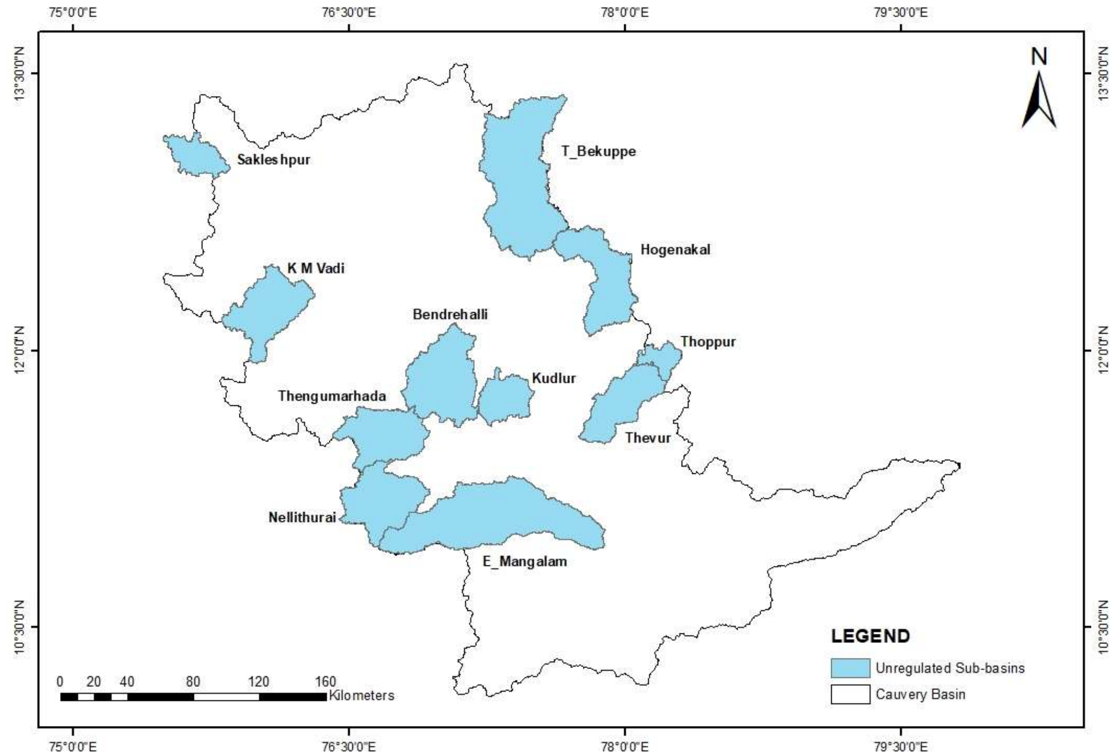


Figure 4.4: Delineated unregulated basins in the Cauvery river basin

4.3.1.4 Delineation in East Flowing River Basin

Only five gauged basins within the east-flowing river were delineated, as depicted in Figure 4.5. The average elevation of these delineated basins varied from a maximum of 1600 m to a minimum of 133 m. These basins exhibited diverse dimensions, with lengths ranging from 117708 m to 27597 m and widths varying from 59026 m to 6199 m. The average form factor of these basins in the westward flow of the river measured 0.33. The average values for the shape factor, circulatory ratio, and elongation ratio were 3.55, 0.18, and 0.46, respectively, with the majority of basins depicting an asymmetric and elongated basin structure. Furthermore, the average drainage density of these basins was recorded at 0.78 km/km^2, indicating a less dense network of watercourses that generally responds more slowly to rainfall events.

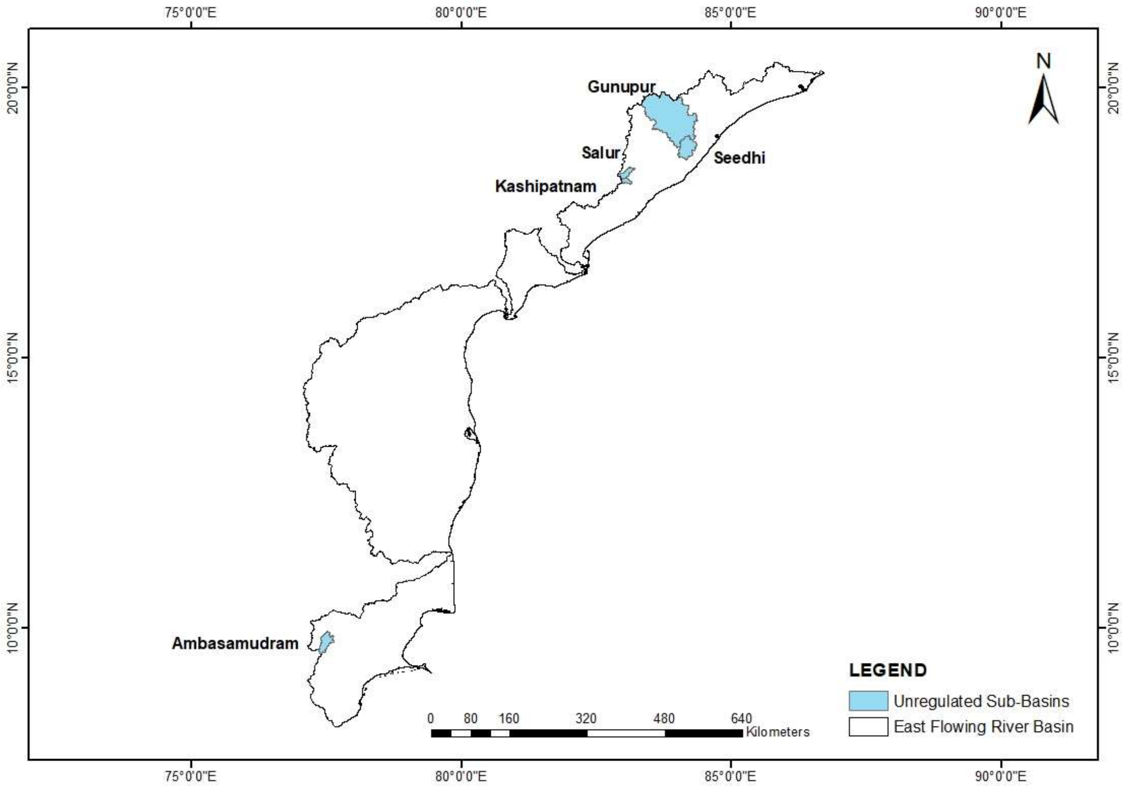


Figure 4.5: Delineated unregulated basins in East Flowing River Basin

4.3.1.5 Delineation in Godavari River Basin

Eight gauged basins within the Godavari River were delineated and are depicted in Figure 4.6. The average elevation of these delineated basins ranged from a maximum of 906 m to a minimum of 375 m. These basins displayed varying dimensions, with lengths ranging from 103357 m to 29783 m and widths varying from 31044 m to 6075 m. The average values for the form factor, shape factor, circulatory ratio, and elongation ratio were 0.33, 3.4, 0.17, and 0.48, respectively, with the majority of basins signifying an elongated and less oval-shaped basin structure. Similar to previous east-flowing river basins, the average drainage density of these basins was recorded at 0.77 km/km^2, indicating a less dense network of watercourses that generally responds more slowly to rainfall events.

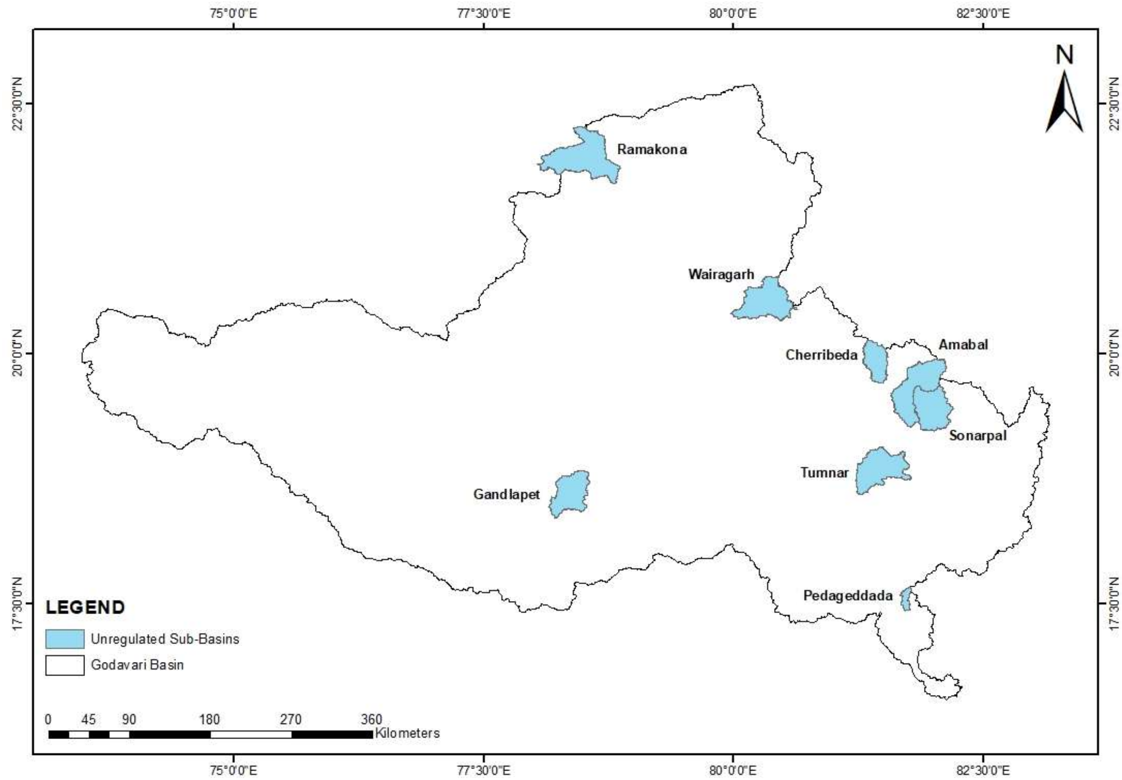


Figure 4.6: Delineated unregulated basins in the Godavari River basin

4.3.2 Cluster Analysis

A necessary first step in hydrologic regionalization is to group the gauged catchments into hydrologically homogeneous groups/clusters so as to ensure accurate information transfer to ungauged basins located within them. Therefore, a hierarchical agglomerative cluster analysis utilizing Ward's linkage method was implemented using carefully selected catchment attributes as clustering variables.

4.3.2.1 Clustering Variables

In order to reduce bias, it is crucial to choose a limited set of variables for cluster analysis. The cluster variables shown in Table 4.1 were identified based on literature review and analysis of the correlation matrix of the 16 catchment attributes as shown in Table 4.2. As mentioned earlier, the clustering of 50 basins was carried out with the SPSS software tool. Each of the nine cluster variables was standardized to give equal importance and avoid the issues related to the usage of different measuring units.

Table 4.1: Cluster variables selected for study

Attribute	Units	Abbreviation
Maximum Elevation	km	MAX_e
Minimum Elevation	km	MIN_e
Slope	--	S
Basin Area	km^2	A
Shape Factor	--	SF
Circularity Ratio	--	R_c
Elongation Ratio	--	R_L
Drainage Density	km/km^2	DD
Rainfall	m	Rain

Table 4.2: Correlation matrix of the cluster variables

Variables	**MAX_e** (km)	**MIN_e** (km)	**ΔH** (km)	**ΔH/P**	**S**	**A** (km^2)	**P** (km)	**L** (km)	**W** (km)	**L_p** (km)	**FF**	**SF**	**R_c**	**R_L**	**DD (km/ km^2)**	**Rain (m)**
MAX_e (km)	1.00															
MIN_e (km)	-0.07	1.00														
ΔH (km)	0.90	-0.49	1.00													
ΔH/P	0.52	-0.52	0.68	1.00												
S	0.59	-0.52	0.74	0.96	1.00											
A (km^2)	-0.01	0.21	-0.10	-0.58	-0.54	1.00										
P (km)	-0.02	0.26	-0.13	-0.67	-0.63	0.95	1.00									
L (km)	-0.03	0.24	-0.13	-0.62	-0.65	0.82	0.91	1.00								
W (km)	0.05	0.20	-0.04	-0.55	-0.41	0.89	0.81	0.56	1.00							
Lp (km)	-0.03	0.19	-0.11	-0.61	-0.61	0.81	0.89	0.96	0.60	1.00						
FF	0.00	-0.05	0.02	-0.04	0.18	0.09	-0.03	-0.33	0.48	-0.20	1.00					
SF	-0.06	-0.04	-0.03	0.01	-0.19	-0.06	0.09	0.42	-0.45	0.32	-0.87	1.00				
R_c	0.01	-0.17	0.08	0.37	0.44	-0.33	-0.53	-0.58	-0.08	-0.49	0.58	-0.59	1.00			
R_L	0.01	0.25	-0.10	-0.19	-0.06	0.31	0.19	-0.12	0.56	-0.20	0.62	-0.72	0.31	1.00		
DD (km/km^2)	0.01	-0.31	0.15	0.13	0.10	-0.14	-0.01	0.10	-0.27	0.05	-0.28	0.40	-0.29	-0.23	1.00	
Rain (m)	-0.01	-0.48	0.20	0.37	0.41	-0.37	-0.42	-0.45	-0.26	-0.39	0.21	-0.18	0.28	-0.07	0.31	1.00

The dendrogram chart (Figure 4.7) obtained from cluster analysis displayed the distribution of catchments into different groups arranged based on the hierarchical agglomerative concept. Three distinctive clusters were identified for regionalization by utilizing the proximity points (just above 10 scale) along the dendrogram limb as shown in Figure 4.7. The details of the catchments associated with derived clusters are provided in Table 4.3, and their region is depicted in Figure 4.8.

Table 4.3: Group memberships of catchments derived from Cluster Analysis

Cluster Number	No. of Catchments	Catchment Station Name	Associated River Basin
Cluster 1	17	Thumpaman, Ayilam, Pulamanthole, Kuniyil, Nanipalasan, Ozerkheda	West Flowing
		Naguleru	Krishna
		Nellithurai, Thengumarahada, Thoppur, Thevur, Kudlur	Cauvery
		Ambasamudram, Kashipatnam, Salur, Seedhi	East Flowing
		Pedagedadda	Godavari
Cluster 2	11	Karathodu, Kalampur, Kalloopara, Avershe, Erinjipuzha, Yennehole, Addoor, Santeguli, Kidangoor, Haladi, Bantwal	West Flowing
Cluster 3	22	Mahuwa	West Flowing
		Balehonnur, Halia, Khanapur, Kellodu, Navalgund, Talikot, Marol	Krishna
		KMVadi, Bendrahalli, Sakleshpura, Hogenakkal, E_Mangalam, T. Bekuppe	Cauvery
		Gunupur	East Flowing
		Ramakona, Amabal, Wairagarh, Tumnar, Sonarpal, Gandlapet, Cherribeda	Godavari

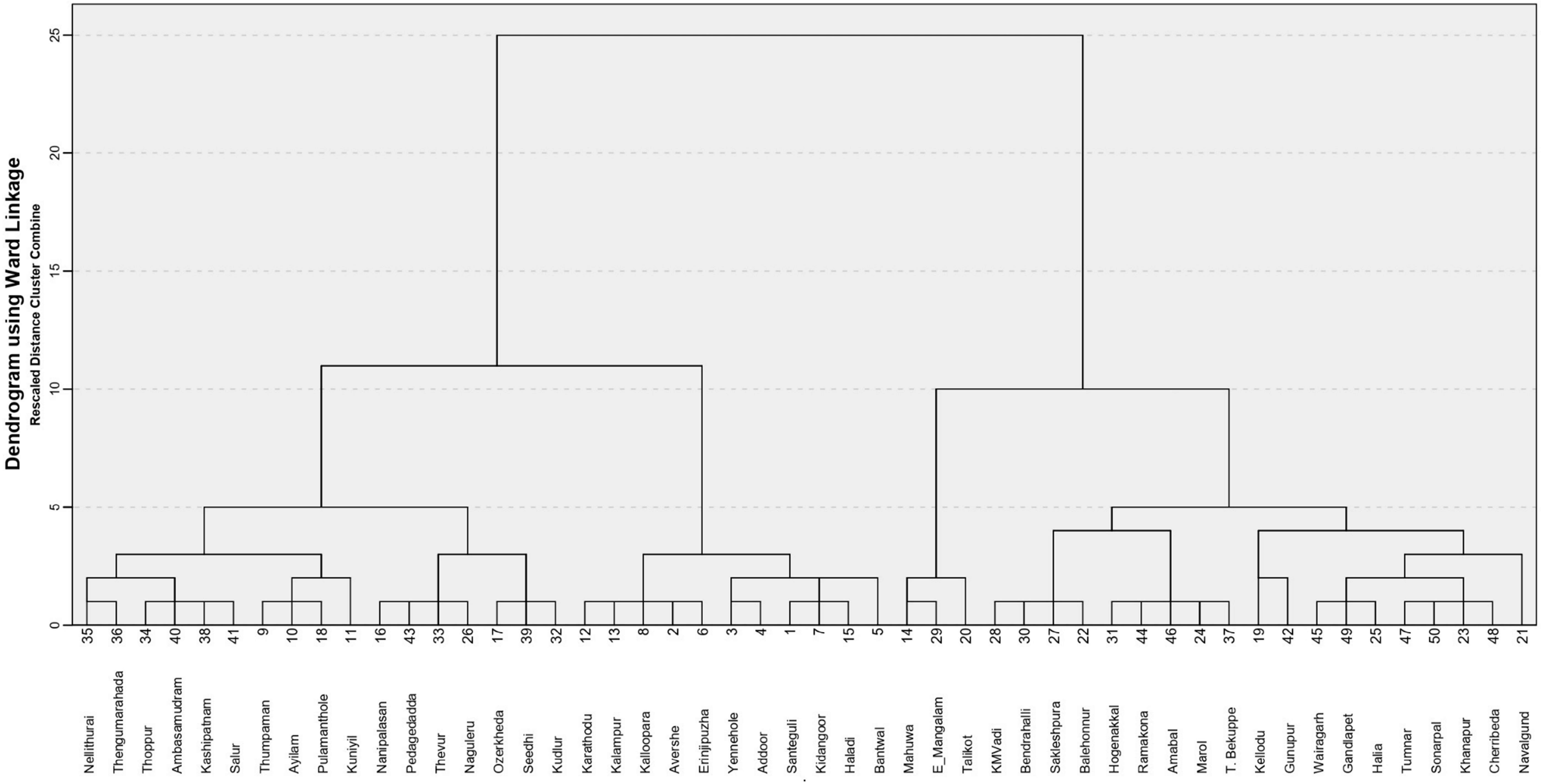


Figure 4.7: Dendrogram chart obtained from the cluster-analysis of 50 catchment using hierarchical agglomerative concept

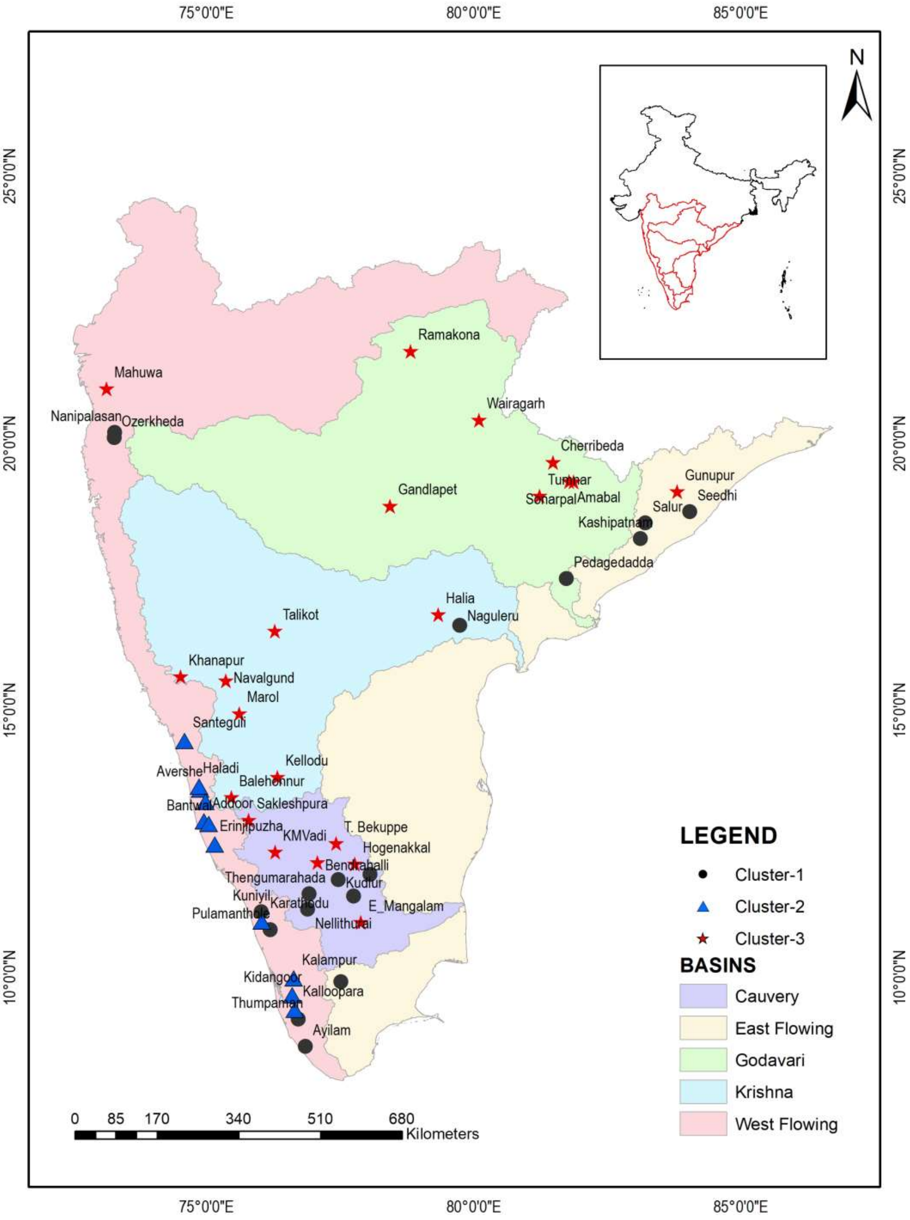


Figure 4.8: Map showing the cluster-wise location of the selected catchment in the Peninsular River System of India

The first cluster contains 17 catchments, with the majority of them spread across West flowing, Cauvery, and East-flowing river basins. The catchment areas of the stations in this cluster vary from a maximum of 1998 km^2 to a minimum of 171 km^2, with average elevation varying from 1710 m to 154m. The mean annual rainfall varies from a maximum of 2153mm to a minimum of 805mm. All the catchments in the second cluster are located within the West flowing river basin with an area varying between 3204 km^2 to 276 km^2. Most of the identified catchments in West flowing rivers are bounded by the Western Ghats Mountains with an average elevation between 1302m to 8m at the coast. The mean annual rainfall varies from a maximum of 4029mm to a minimum of 2280 mm. The last cluster-3 is the biggest, containing 22 catchments, with most of them being located in the Krishna, Cauvery, and Godavari basins. The majority of the catchments in this cluster have a larger area compared to other clusters (a maximum of 6930 km^2 to a minimum of 601 km^2). The average elevations in this cluster vary from 1220 m to 450m and mean annual rainfall ranges from 2831mm to a minimum of 564mm.

4.3.2.2 Homogeneity Test of Cluster Analysis

The study utilized the Coefficient of Variation (CV) test and discordance measure using L-moments to check the credibility of the cluster formations. The homogeneity measure (CC) computed from CV (Equation 4.4) was evaluated considering all 50 catchments as a single homogeneous region and also separately for each of the 3 regions delineated using cluster analysis. Results of this analysis are shown in Table 4.4 from which it is revealed that considering a single region fails the homogeneity requirement since the CC value exceeds 0.3. On the other hand, all 3 delineated clusters yield CC values less than 0.3 and hence they may be considered to be hydrologically homogeneous.

Table 4.4: Results of CV homogeneity test

Regions	No. of catchments	Homogeneity Measure (CC)	Region Type	Test Criteria
All catchments	50	0.622	Non-Homogenous	The region is declared homogenous if CC is less than 0.3 (Nobert et al. 2011)
Cluster-1	17	0.193	Homogenous	
Cluster-2	11	0.067	Homogenous	
Cluster-3	22	0.134	Homogenous	

As discussed previously, the discordance measure test identifies discordant catchments within the homogenous groups. To check the discordancy, 38 years of daily rainfall information derived from the Indian Meteorological Department, Government of India, was utilized for the study. L-moments were determined using the R-studio software tool, and L-ratios were calculated as per Equations (4.5) and (4.6) for each of the identified catchments. The mean L-CV value for the identified catchments was estimated to be 0.84 with lower and upper quartiles varying between 0.80 to 0.87 as represented in Figure 4.9. Similarly, the mean L-Skewness was found to be 0.705 with quartiles ranging from 0.65 to 0.77, and the mean L-Kurtosis was 0.44 with quartiles ranging from 0.36 to 0.53 (Figure 4.9). It is evident that L-CV values exhibited the smallest variability across the 50 catchments while L-Kurtosis exhibited the highest variability.

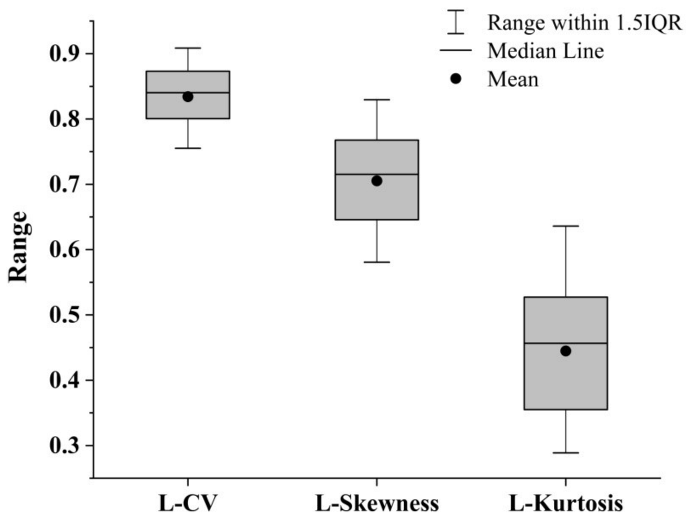
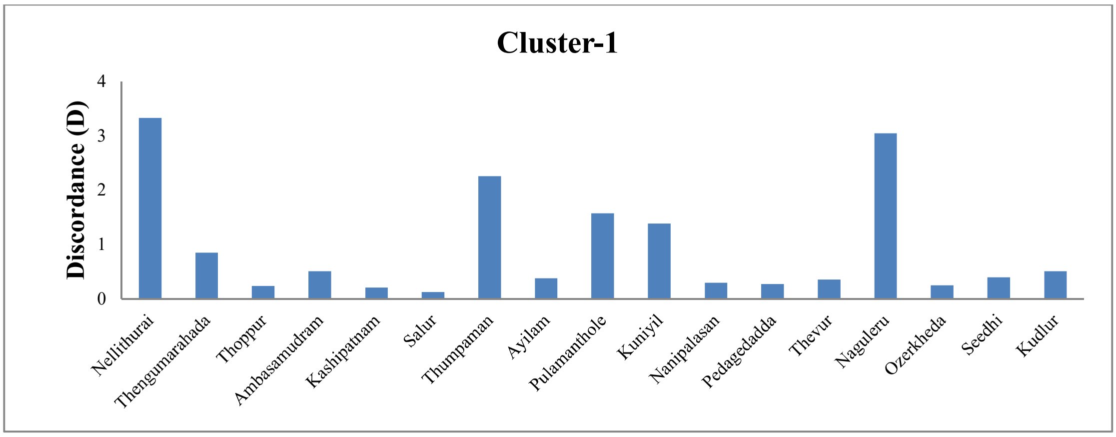
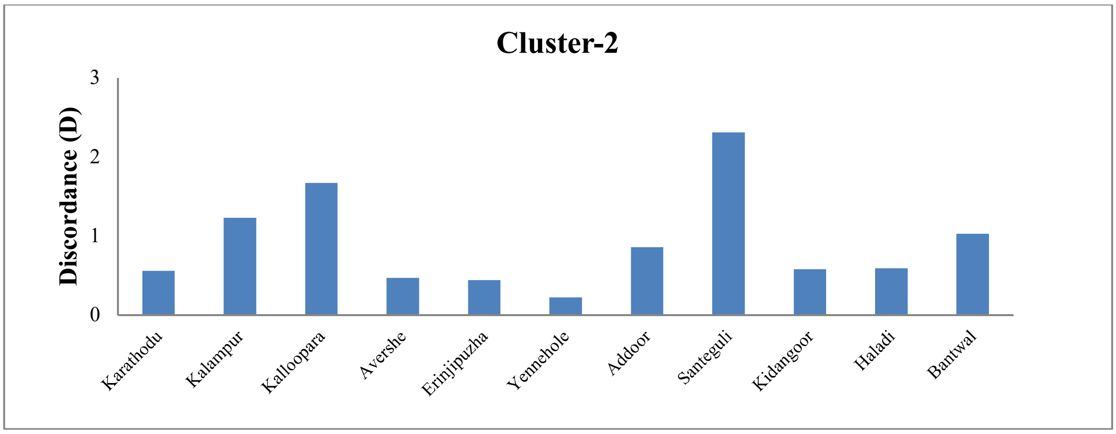


Figure 4.9: Variation of L-CV, L-Skewness, and L-Kurtosis for the identified catchments

Next, the discordance value (D_i) was calculated cluster-wise using Equations (4.7) to (4.9). Figure 4.10 represents the results of the discordance measure test executed for all catchments in a cluster-wise manner. The test reveals no discordancy was observed in the cluster-2 and cluster-3 catchments as most of the D_i values are less than 3 (Hosking and

Wallis 1993). However, one catchment named Nellithurai in cluster-1 has D_i slightly more than 3 making it discordant from others within the region. As suggested by Hosking and Wallis (1993), the discordant catchment was re-verified with respect to the variation of its statistical parameters, such as L-CV, L-Skeweness, and L-Kurtosis and catchment-climatic characteristics with other catchments within the cluster group. It was evident from the cross-verification exercise that- the discordant catchment has a similar kind of statistical and catchment-climatic behaviour as that of the remaining catchments within the same group. Hence it is not worthwhile to shift this catchment to another region as it has dissimilar behaviour with other cluster sites and might affect the regionalization process. Based on these observations, catchment-Nellithurai was retained in Cluster-1 for further analysis.





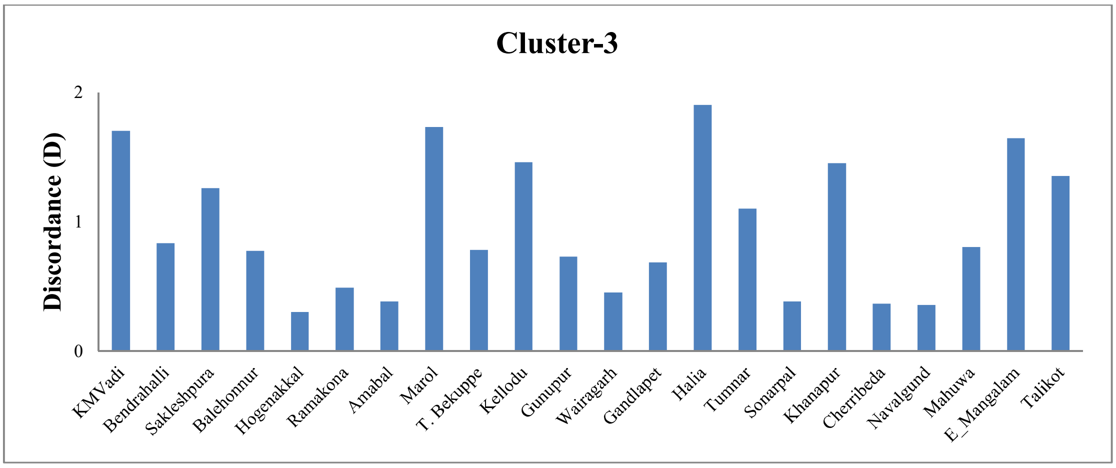


Figure 4.10: Results of Discordance Measure test for Cluster-1, 2 and 3

CHAPTER SUMMARY

- *After examining the basin reports from the Central Water Commission (CWC) and utilizing 30 m resolution SRTM DEM data, a total of 50 basins were identified as unregulated. The delineation of the boundaries all these basins and the determination of catchment characteristics were carried out using the ArcHydro tool. The delineation process revealed the utilization of a broad spectrum of catchments with diverse physiographic and climatic conditions for the study.*

- *For efficient regionalization, a hierarchical agglomerative cluster analysis utilizing Ward's linkage method was implemented, resulting in the delineation of the study area into three homogeneous clusters. Cluster 1 comprised 17 catchments, followed by 11 catchments in Cluster 2, and 22 catchments in Cluster 3. Through the CV test and L-Discordancy measure using the L-Moment ratio, it was confirmed that all three clusters exhibited homogeneity without any discordant catchments.*

CHAPTER 5

REGIONALIZED FLOW DURATION CURVES

5.1 FLOW DURATION CURVE (FDC)

Flow Duration Curve (FDC) is one of the common tools used in hydrological studies that provide concise information about the river flow variability in the study basin (Quimpoet al. 1983; Yu et al. 2002; Castellarin et al. 2004; Boscarello et al. 2016; Burgan and Aksoy 2020). A Flow Duration Curve (FDC) depicts the relationship between the percentage of time (or duration) for which a particular magnitude of discharge is equalled or exceeded at a particular gauging site. It is a valuable hydrological tool in the planning and design of water resources projects and therefore the present study focused on its estimation at ungauged locations. An FDC is said to be the complement of the cumulative distribution function of daily streamflow. FDC for each of the gauge stations can be developed using a standard non-parametric approach that involves counting the number of occurrences of historical flows falling within class intervals of descending flow magnitudes q_i with i = 1, 2...,n and subsequent calculation of percent exceedance probability using an appropriate plotting position formula (Fennessey and Vogel 1990; Vogel and Fennessey 1994; Sugiyama et al. 2003; Castellarin et al. 2004; Isik and Singh 2008; Li et al. 2010; Shu and Ouarda 2012).

The Weibull plotting position formula is most commonly adopted and is given by,

$$p_i = P(Q \geq q_i) = \frac{m}{n+1} \times 100\% \tag{5.1}$$

Where p_i= Probability of the discharge being greater than or equal to a specified value q_i of ordered stream flows, 'm' is the number of counts of the flow values falling within the specified interval, and 'n' is the number of events on records (Shu and Ouarda 2012). Subsequently, nine flow quantiles representing the discharge magnitudes at durations of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% were extracted from the FDCs of each catchment.

5.2 REGIONALIZATION

5.2.1 Cluster Wise Regionalization through Multiple Linear Regression Technique

Once the homogenous regions are delineated using cluster analysis, the regionalization concept is applied in each cluster. Among different regionalization techniques, the multiple linear regression (MLR) analysis is one of the earliest and most widely used techniques globally for regionalization (Li et al. 2010; Bao et al. 2012, Swain and Patra 2017). This technique aims to develop a relationship between identified catchment characteristics and stream flow information corresponding to gauged catchment through a multiple linear regression equation.

The review of the literature revealed that with regard to the characteristics of the observed FDC to be reproduced in the ungauged basin, two broad approaches have been used – a 'parametric' approach in which a function (empirical or probabilistic) is fitted to the observed FDCs and the optimized parameters of the function are predicted for the ungauged basin using the transfer function. On the other hand, in the 'point' approach, flow quantiles (Q_D) corresponding to specific values of duration (D) (for example, Q_{10}, Q_{20},….Q_{90}) are extracted from the observed FDCs and predicted for the ungauged basin using the transfer function. Using the data available, both the parametric and point approaches for the regionalization of the FDCs were implemented. However, preliminary results for the parametric approach (not shown here for brevity) indicated poorer performance in comparison to the point approach, and therefore the latter approach was adopted.

Accordingly, separate MLR equations relating each of the 9 flow quantiles (Q_{10}, Q_{20},….Q_{90}) extracted from the observed FDCs to the identified catchment attributes of the gauged catchments were established. The general form of the MLR equation used was,

$$Q(D) = \psi_0 + \psi_1 X_1 + \psi_2 X_2 + \ldots\ldots\ldots\ldots + \psi_n X_n + \varepsilon \tag{5.2}$$

Where Q(D) is the flow quantile of specific percentage duration(D) (10%, 20% …90%) for each of the gauged catchments; X_1, X_2, ……X_n are the selected catchment attributes for the gauged catchments; ψ_1, ψ_2,…….ψ_n are regression coefficients obtained through the least squares criterion and ε represents the random error. The optimal regression coefficients obtained from Equation (5.2) for identified gauged catchments were then utilized to estimate Q(D) values for the ungauged catchments by substituting their catchment attributes (Yu 2002; Shu and Ouarda 2012; Nruthya and Srinivas 2015; Silva 2019). As per He et al. (2011), the performance of the MLR approach mainly depends on the appropriate choice of attributes.

For the regionalization approach, catchment attributes were selected and utilized for multiple regression analysis in each of the 50 unregulated catchments. The correlation matrix of the catchment attributes was studied to reduce multi-collinearity problems (Mohamoud 2008). The correlation matrix analysis coupled with the stepwise regression procedure facilitated the identification of the most influential and irredundant catchment attribute for explaining the observed variability in the flow quantiles (Nathan and McMahon 1990; Vogel et al. 1999; Yu 2002). The backward stepwise regression analysis tool available in the SPSS statistical package was used in the present study.

5.2.2 Leave-One-Out Jackknife Cross Validation

In an effort to evaluate the reliabilities of the developed MLR models when applied in ungauged catchments, a cluster-wise leave-one-out jackknife cross validation procedure was implemented. Leave-one-out jackknife cross-validation is a commonly used validation technique for evaluating the uncertainties between the developed model and input data

(Efron 1981; Shao and Tu 1995) and is especially useful with small data sets. In a comparative study between jackknife and split-sampling methods, McCuen (2005) found that the jackknife test was less sensitive to the sample size variation and provided better model prediction accuracy than the split-sampling technique. Literature has also indicated that the model precision obtained using the jackknife technique is independent of calibration data (McCuen 2005; Shu and Ouarda 2012).

In the jackknife cross-validation technique as applied to the regionalization of FDCs, one-gauge station is assumed to be ungauged and the information from the remaining (n-1) gauge stations is utilized to develop the regression model associated with specified flow quantiles. The catchment attributes of the withheld station are then used in the developed regression to synthesize the flow quantile in the assumed ungauged catchment. Similarly, the station withheld in the first jackknife run is replaced, and the next gauged station is assumed to be ungauged for the second run. This procedure is continued until all the gauge stations have been utilized to make predictions (McCuen 2005; Shu and Ouarda 2012; Nruthya and Srinivas 2015). The model's reliability is then tested by comparing the predicted jackknife estimates (model predictions) with those of the observed values. Such cross-validations help to derive the reliability of the regional regression model (Castellarin et al. 2004).

5.2.3 Performance Evaluation

The model efficiency in predicting the flow quantile was determined by comparing the model results with the observed flow values corresponding to the station under investigation. The following three efficiency measures were used to assess the performance of the MLR models:

- Coefficient of Determination (R^2):

$$R^2 = \left(\frac{\sum_{i=1}^{N} (O_i - \overline{O})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{N} (O_i - \overline{O})^2} \sqrt{\sum_{i=1}^{N} (P_i - \overline{P})^2}} \right)^2 \tag{5. 3}$$

- Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (P_i - O_i)^2}{N}} \ (\mathrm{m^3/s}) \tag{5.4}$$

- Percentage Bias (PBIAS):

$$PBIAS = \frac{\sum_{i=1}^{N} (O_i - P_i)}{\sum_{i=1}^{N} O_i} 100 \tag{5.5}$$

In the above equations, 'N' denotes the number of selected catchments, 'O' and 'P' are the observed and predicted flow values, and, $\overline{O}$ and $\overline{P}$ indicate the average values of observed and predicted flow rates. Root Mean Square Error (RMSE) displays the extent of a typical error. For an ideal model, the RMSE value should be zero. The coefficient of determination (R^2) determines the percentage variation in the observations which is explained by the model with a value of 1 signifying ideal model performance (Domínguez et al. 2010; Shu and Ouarda 2012). Further, the model's tendency to consistently under-predict or over-predict the observed values is measured using percent bias (PBIAS). Model performance can be judged acceptable if the PBIAS is within ±15% for flow simulations (Moriasi et al. 2015).

5.3 RESULTS AND DISCUSSIONS

5.3.1 Results and Discussions of Flow Duration Curve (FDC)

The observed flows of individual gauge stations were plotted against the corresponding exceedance probability by means of the Weibull plotting position formula (Equation 5.1) to obtain the corresponding flow duration curves for the gauge stations. As mentioned previously, a total of nine flow quantile values viz., 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% were obtained by interpolation for each of the identified gauge stations. The extracted flow quantiles were utilized for regression analysis and subsequent generation of FDCs for ungauged stations. The maximum average flow at 10 % flow quantile (Q_{10}) from all the basins combined was 105.95 m^3/s and the minimum average flow at Q_{90} was about 0.68 m^3/s.

FDC for identified catchments in West Flowing River Basins

The FDCs for all the delineated gauged basins in West Flowing Rivers were developed using the flow information downloaded from the WRIS web portal of the Government of India. Inconsistent and missing data were not considered for analysis to avoid any possible uncertainty. The maximum duration of the daily time series flow record utilized for developing FDC was 28 years (1991 to 2018) and the minimum flow record was 15 years (2004 to 2018). The maximum average flow at 10 % flow quantile (Q_{10}) was estimated to be 224.4 m^3/s and the minimum average flow at Q_{90} was about 1.02 m^3/s. Illustrative flow duration curves for the largest and smallest gauge basins in all the identified catchments in west flowing river basin are shown in Figure 5.1.

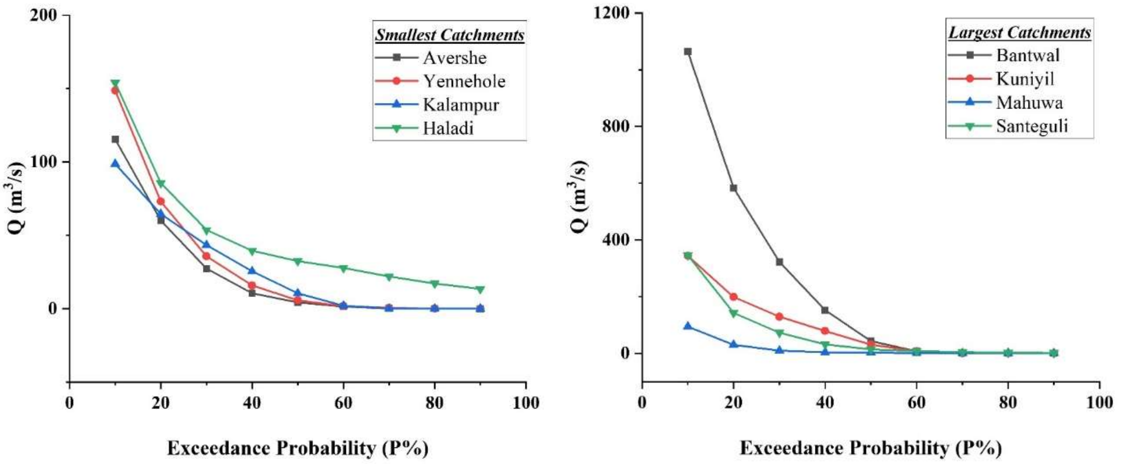


Figure 5.1 FDC for the identified smallest catchments (left) and largest catchments (right) in West Flowing River basins

FDC for identified catchments in Krishna River Basin

The FDCs were developed for eight gauge stations within the Krishna river basin. The FDC development process involved utilizing daily time series flow records, with a maximum duration of 21 years (1995 to 2015) and a minimum flow record duration of 10 years (1999 to 2008). The estimated maximum average flow at 10 % flow quantile (Q_{10}) was 74.85 m^3/s, while the minimum average flow at Q_{90} was about 0.61 m^3/s. Figure 5.2 illustrates the flow duration curves for the largest and smallest gauge basins among all the identified catchments in the Krishna river basin.

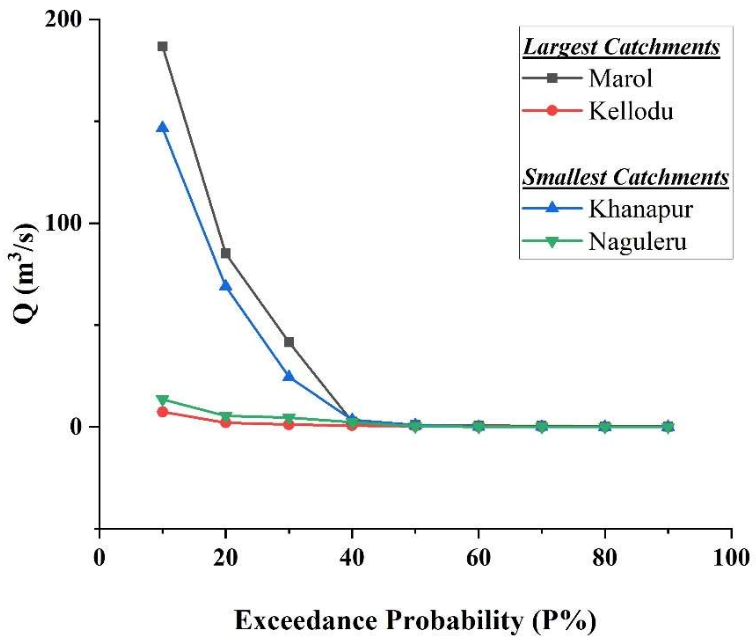


Figure 5.2 FDC for the identified smallest and largest catchments in Krishna River Basin

FDC for identified catchments in the Cauvery River Basin

FDCs were created for eleven gauge stations within the Cauvery river basin. The daily time series flow record utilized for FDC development ranged from a maximum duration of 26 years (1991 to 2016) to a minimum flow record was 11 years (2008 to 2018). In comparison with other basins, the selected catchments within the Cauvery basin exhibited the smallest maximum average flow at 10 % flow quantile (Q_{10}), estimated to be 30.5 m^3/s, while the minimum average flow at Q_{90} was 0.77 m^3/s. Figure 5.3 depicts the flow duration curves for both the largest and smallest gauge basins across all identified catchments in the Cauvery river basin.

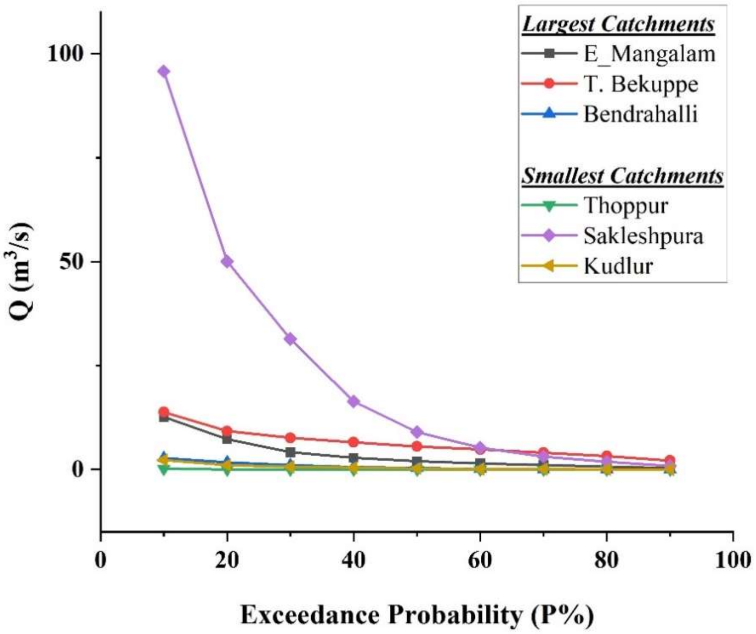


Figure 5.3 FDC for the identified smallest and largest catchments in the Cauvery River Basin

FDC for identified catchments in East Flowing River Basin

FDCs were developed for only five identified stations within the east-flowing river basin. The daily time series flow record used for FDC development ranged from a maximum duration of 18 years (2001 to 2018) to a minimum of 10 years (2000 to 2009). The estimated maximum average flow at 10 % flow quantile (Q_{10}) was 47.65 m^3/s and the minimum average flow at Q_{90} was about 0.61 m^3/s. Illustrative flow duration curves for the largest and smallest gauge basins among all the identified catchments in east flowing river basin are shown in Figure 5.4.

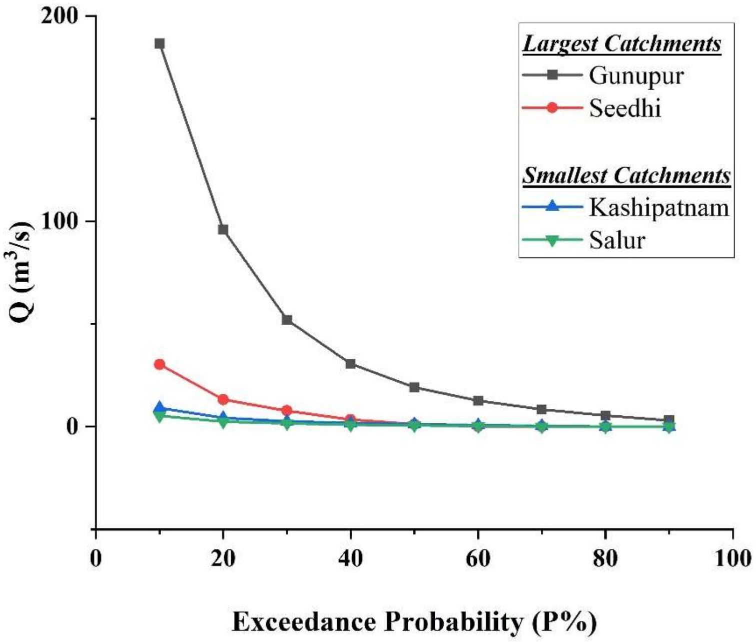


Figure 5.4 FDC for the identified smallest and largest catchments in the East Flowing River Basin

FDC for identified catchments in the Godavari River Basin

Eight FDCs were established for identified stations within the Godavari river basin. The maximum duration of the daily time series flow record utilized for developing FDC was 26 years (1993 to 2018) and the minimum flow record was 11 years (1991 to 2001). The estimated maximum average flow at 10 % flow quantile (Q_{10}) was 53 m^3/s, while the minimum average flow at Q_{90} was 0.11 m^3/s. Figure 5.5 illustrates the flow duration curves for the largest and smallest gauge basins across all the identified catchments in the Godavari river basin.

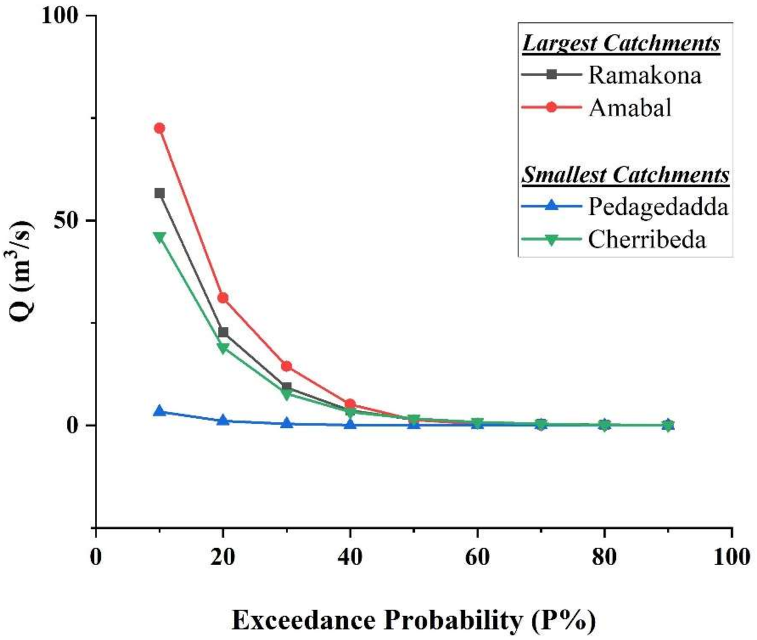


Figure 5.5 FDC for the identified smallest and largest catchments in the Godavari River Basin

5.3.2 Results and Discussions of Regionalization

The regionalization process in this study was carried out for individual cluster groups separately, resulting in catchment-wise regression relationships in the form of Equation (5.2). The stepwise regression method (in the SPSS statistical tool) was adopted to develop the separate MLR models for each of the 9 flow quantiles (Q_{10}, Q_{20},…,Q_{90}) as response variables and catchment attributes as predictor variables (Equation 5.2). From among the catchment attributes (Article 3.4), only 15 of the physiographic catchment attributes were considered as potential predictor variables and backward stepwise regression analysis was carried out to identify the most significant catchment attributes. To compare the model results, the regression analysis was initially executed by considering all 50 catchments to constitute a single region and subsequently considering catchments in each of the 3 delineated homogeneous clusters.

Single region analysis

In this analysis, all 50 catchments were considered as one group for step-wise regression with the observed flow quantile values in each catchment as the response variable and the 15 catchment physiographic attributes as potential predictor variables in Equation 5.2. The resulting final MLR models (Equation 5.2) for each of the 9 flow quantiles represented in terms of the intercept term (ψ_0) and regression coefficients (ψ_1, ψ_2,…. ψ_N) for the most significant predictor variables as determined through step-wise regression are listed in Table 5.1. From the results shown therein, certain inferences, albeit in a statistical sense only, can be drawn regarding the influence of catchment attributes on the flow quantiles and thereby on the shape of the FDCs. For instance, it can be seen that the Maximum elevation (MAX_e) is a significant predictor variable for almost all the flow quantiles and seems to have a major influence on the shape of the FDC. On the other hand, while Minimum elevation (MIN_e) and circulatory Ratio (R_c) have an effect only on the upper half of the FDC, several other attributes such as Relative relief ($\Delta H/P$), catchment Area (A), basin Width (W) and Form Factor (FF) appear to influence only the lower part of the FDC. Attributes Longest Flow Path (L_p) and elongation Ratio (R_L) influence some parts of the upper and lower portions of the FDC whereas basin perimeter (P) affects only the median flow quantiles. Slope (S) is significant only in the case of the two largest flow quantiles and Drainage Density (D_d) for only the lowest flow quantile.

The performances of the optimal MLR models were evaluated using the performance statistics described in Article 5.2.3, i.e., R^2, RMSE, and PBIAS. However, it was found that all the MLR models yielded negligible values of PBIAS and accordingly the values of only R^2 and RMSE are shown in Table 5.2 The performances of the MLR models developed considering all 50 catchments to be in a single region when evaluated in terms of R^2 indicated that the results were quite poor for all the flow quantiles (R^2 between 0.16 to 0.31) with the performances being slightly better for the high flow quantiles in comparison to the low flow quantiles. RMSE values which depend on the magnitude of flows, ranged between 1.90 m^3/s and 133.80 m^3/s with the lower values being associated with low flow quantiles and vice versa.

Table 5.1: Regression coefficients associated with different predictor variables for final MLR models (Equation 5.2) for nine flow quantiles considering all catchments in a single region

Flow Quantile	Constant ψ_0	Regression Coefficients (ψ_i) for the predictor variable														
		MAX_e (km)	MIN_e (km)	ΔH (km)	$\Delta H/P$	S	A (km^2)	P (km)	L (km)	W (km)	L_p (km)	D_d (km/km^2)	FF	SF	R_c	R_L
Q_{10}	-33.3	+93.9	-293.4	=	=	-4078.2	=	=	-8.5	=	+6.3	=	=	=	-909.7	+688.7
Q_{20}	-19.6	+61	-158.4	=	=	-2235	=	=	-4.2	=	+3	=	=	=	-474.4	+349.4
Q_{30}	-27.4	+18.3	-68	=	=	=	=	=	-2.1	=	+1.7	=	=	=	-358.5	+209.7
Q_{40}	+23.8	+23.7	-49	=	-4206.1	=	=	+0.1	-0.5	=	=	=	=	=	=	=
Q_{50}	+9.7	+9.6	-15.4	=	-1530.2	=	=	+0.03	-0.2	=	=	=	=	=	=	=
Q_{60}	+29.5	+6.7	=	=	-1295.5	=	+0.01	=	=	-0.7	-0.2	=	+32.9	=	=	-49.1
Q_{70}	+21.6	+4.8	=	=	-1085.8	=	+0.01	=	=	-0.7	-0.1	=	+29.3	=	=	-35
Q_{80}	+12.5	+2.9	=	=	-658.1	=	+0.005	=	=	-0.5	-0.1	=	+21.7	=	=	-21.3
Q_{90}	+0.4	=	=	=	=	=	+0.002	=	=	-0.3		-0.7	+11.1	=	=	=

Table 5.2: Performance statistics of final MLR models for nine flow quantiles developed considering catchments in a single region, Cluster -1, Cluster – 2, and Cluster – 3

MLR for Flow Quantile	Single Region		Cluster – 1		Cluster – 2		Cluster – 3	
	R^2	RMSE (m^3/s)	R^2	RMSE (m^3/s)	R^2	RMSE (m^3/s)	R^2	RMSE (m^3/s)
Q_{10}	0.31	133.80	0.86	32.73	0.98	9.09	0.83	26.64
Q_{20}	0.29	74.73	0.86	18.39	0.98	6.59	0.85	14.23
Q_{30}	0.26	43.50	0.98	4.74	0.98	4.73	0.83	9.64
Q_{40}	0.27	22.31	0.98	1.61	0.98	2.28	0.76	7.59
Q_{50}	0.20	9.88	0.98	0.43	0.98	0.47	0.77	5.15
Q_{60}	0.18	5.95	0.92	0.80	0.96	1.39	0.79	3.53
Q_{70}	0.19	4.23	0.79	0.67	0.96	1.20	0.79	2.44
Q_{80}	0.22	2.66	0.56	0.79	0.96	0.84	0.81	1.13
Q_{90}	0.16	1.90	0.72	0.48	0.96	0.71	0.77	0.55
Average	0.23	33.22	0.85	6.74	0.97	3.03	0.80	7.88

Cluster-wise analysis

The catchment attributes and the flow quantiles related to 17 catchment of Cluster-1 (Table 4.3) were used to carry out stepwise regression analysis. The forms of the resulting final MLR models for Cluster – 1 are listed in Table 5.3. From these results, it is immediately apparent that unlike in the previous case of considering a single region (Table 5.1), in this case, all the catchment attributes except MAX_e, have an influence on some or all flow quantiles. The attribute relief (ΔH) is a significant predictor for all flow quantiles, and catchment length (L) is significant for all except one flow quantile (Table 5.3). Attributes S, A, W, and FF appear to have a significant effect on only the low-flow quantiles. Also, it is interesting to note that more number of predictor variables are involved in the models for medium flow quantiles in comparison to the high flow quantiles and that the least number of predictors are required for predicting the low flow quantiles. From a hydrological perspective, the flow characteristics in the majority of catchments within Cluster-1 primarily rely on geometric aspects, such as S, L, A, and W. Additionally, the relief factors, particularly MAXe, S, and ΔH, as well as the areal aspect FF, play a significant role in shaping these flow patterns. Notably, within Cluster-1, a high relief value (ΔH), which indicates the overall steepness of the terrain, was observed in the majority of catchments. Furthermore, FF which signifies the intensity of flow, exhibited higher values in six specific catchments (Kuniyil, Ozerkheda, Kudlur, Nellithurai, Thengumarahada, and Seedhi). The shapes of these catchments were more rounded/circular leading to concentrated flows. Conversely, FF values in the remaining catchments indicated an elongated catchment behaviour. The performance statistics for the final MLR models for Cluster – 1 are listed in Table 5.2. Results indicate that grouping catchments into homogeneous regions leads to significant improvement in the performances of the developed MLR models in comparison to using a single region. High values of R^2 for all flow quantiles and more so for the Q_{30}, Q_{40}, and Q_{50} quantiles were obtained indicating excellent predictive capabilities of the developed MLR models. Also, significantly lower values of RMSE were obtained across all flow quantiles (Table 5.2).

The final forms of the MLR models derived using step-wise regression for catchments in Cluster – 2 are listed in Table 5.4. In this case, it can be seen that 4 (MIN_e, ΔH, S, and SF) out of the potential 15 catchment attributes considered as predictors are significant in the models of all the flow quantiles. Attributes FF, R_c, and R_L too are significant for all but a few flow quantiles. The catchment area (A) and length of the basin (L) are significant for the high flow and low flow quantiles respectively. Overall, the number of significant predictors for this cluster is smaller than that for Cluster – 1 but as in the earlier case, models do not seem to be more parsimonious for low flows. Hydrologically, the flows in the majority of catchments within Cluster-2 rely on geometric aspect L, relief aspects MINe, S, and ΔH, and areal aspect FF, SF, R_c, and R_L. Similar to the preceding cluster, most catchments in this cluster exhibit a moderate relief value (ΔH). Moreover, four catchments (Santeguli, Yennehole, Bantwal, and Haladi) display moderate values of form factor and elongation ratio, and lesser values of shape factor and circulatory ratio, indicating that these catchments have moderately elongated shapes and are associated with a low flow response. Among all the cases considered, grouping catchments into Cluster – 2 yielded the most accurate MLR models for all flow quantiles as indicated by the results of the performance analysis shown in Table 5.2. Extremely high values of R^2 (0.96 to 0.98) and extremely low values of RMSE (0.47 m^3/s – 9.09 m^3/s) were recorded for the derived MLR models (Table 5.2).

Derived MLR models for Cluster – 3 catchments are listed in Table 5.5. It is evident that in this case all 15 potential catchment attributes are involved in one or more models for the 9 flow quantiles. In particular, attributes MAX_e, MIN_e, A, and W turn out to be significant predictors for all flow quantiles, and attributes ΔH/P, P, and L_p are important in all but one of the flow quantile models (Table 5.5). Among the remaining attributes, L and FF are significant for the high flow quantiles and DD and R_c for low flow quantiles. In terms of hydrology, the flow patterns in most catchments within Cluster-3 are mainly influenced by geometric aspects, including A, P, W, and L_p. Furthermore, the relief factors, particularly MAXe, MINe, and ΔH/P, along with the areal aspect DD, play a crucial role in shaping these flow patterns. Notably, Cluster-3 exhibits relatively lower values of relative relief and drainage density, indicating an elongated catchment characterized by highly permeable soils and a coarse drainage texture. Performance statistics of the MLR models for Cluster – 3 shown in Table 5.2 indicate that although the accuracies of the derived models are not as good as for the other two clusters, the performances still are far superior to the case of single region case. Values of R^2 in the range of 0.76 - 0.85 and RMSE values between 0.55 m^3/s and 26.64 m^3/s for the derived MLR models are indicative of good performances.

Table 5.3: Regression coefficients associated with different predictor variables for final MLR models (Equation 5.2) for nine flow quantiles considering all catchments in Cluster-1

Flow Quantile	Constant ψ_0	Regression Coefficients (ψ_i) for the predictor variable														
		MAX_e (km)	MIN_e (km)	ΔH (km)	ΔH/P	S	A (km^2)	P (km)	L (km)	W (km)	L_p (km)	D_d (km/km^2)	FF	SF	R_c	R_L
Q_{10}	+480	=	-252.4	+226.5	=	-9177.3	+0.5	=	-8.8	-40.8	=	=	+1006.9	=	-730.1	=
Q_{20}	+280.4	=	-138.3	+137.3	=	-5439.4	+0.3	=	-5.2	-24.3	=	=	+592.7	=	-445.4	=
Q_{30}	+15.6	=	-44.5	+65.2	+28452.1	-9378.4	+0.1	+1.9	-7.6	-18.9	=	=	+466.8	=	=	=
Q_{40}	-16.5	=	-17.2	+33.4	+17715.1	-5535.7	+0.1	+1.2	-4.4	-11.7	=	+8.3	+280.3	=	=	+29.3
Q_{50}	-35.3	=	-11.7	+19.2	+6133.9	-2103.7	+0.03	+0.4	-2.2	-5.9	+0.6	+4.5	+105.9	=	=	+93.3
Q_{60}	-24.2	=	-9.9	+11.1	=	-302.9	+0.01	=	-0.8	-2.3	+0.6	=	+21.1	=	-20.5	+89.1
Q_{70}	-38.8	=	-5.6	+3	=	=	=	=	-0.7	-0.9	+0.6	-3.7	=	+1.8	-21.7	+87.1
Q_{80}	-0.28	=	=	+1.69	=	=	=	=	=	=	=	-1.95	=	=	=	=
Q_{90}	+1.4	=	=	+3.2	=	-87.1	=	-0.03	-0.1	=	+0.1	-1.2	=	=	-21.1	+11.6

Table 5.4: Regression coefficients associated with different predictor variables for final MLR models (Equation 5.2) for nine flow quantiles considering all catchments in Cluster-2

Flow Quantile	Constant ψ_0	Regression Coefficients (ψ_i) for the predictor variable														
		MAX_e (km)	MIN_e (km)	ΔH (km)	ΔH/P	S	A (km^2)	P (km)	L (km)	W (km)	L_p (km)	D_d (km/km^2)	FF	SF	R_c	R_L
Q_{10}	-1008.4	=	-4213.5	-121.9	=	+9197.1	+0.4	=	=	=	=	=	+596.6	+115.5	=	+543.6
Q_{20}	-418.6	=	-2858	-87.6	=	+5007.3	+0.2	=	=	=	=	=	+160.4	+50.3	=	+298.4
Q_{30}	-186.4	=	-2693.4	-82.2	=	+3836.7	+0.1	=	=	=	=	=	+49.8	+28.7	+270	=
Q_{40}	-93.3	=	-2267.1	-134	=	+4306.9	+0.04	=	+2.6	=	=	=	=	+7.4	+86.3	=
Q_{50}	-110.2	=	-2471.5	-150.6	=	+5353.6	=	=	+3.5	=	=	=	+50	+7.9	+108.4	-96.6
Q_{60}	-183.9	=	-2502.8	-156.9	=	+6238.9	-0.02	=	+4.4	=	=	=	+111.3	+10.5	=	-68.8
Q_{70}	-143.4	=	-2253.5	-107.4	=	+4768.8	=	=	+2.4	=	=	=	+93.1	+14.6	+187.5	-145
Q_{80}	-117.3	=	-1806.9	-87.5	=	+3963.4	=	=	+1.9	=	=	=	+76.63	+11.8	+146.8	-122.1
Q_{90}	-93.3	=	-1444.9	-70	=	+3164	=	=	+1.5	=	=	=	+59.1	+9.3	+115.5	-95.1

Table 5.5: Regression coefficients associated with different predictor variables for final MLR models (Equation 5.2) for nine flow quantiles considering all catchments in Cluster-3

Flow Quantile	Constant ψ_0	Regression Coefficients (ψ_i) for the predictor variable														
		MAX_e (km)	MIN_e (km)	ΔH (km)	ΔH/P	S	A (km^2)	P (km)	L (km)	W (km)	L_p (km)	D_d (km/km^2)	FF	SF	R_c	R_L
Q_{10}	+298.8	-184	+103.5	=	+78998.7	=	+0.2	=	-3.6	-23.4	+1.1	=	+765.8	=	-869.3	=
Q_{20}	+35.7	-151.2	+101	=	+70198.6	=	+0.1	+0.4	-1.6	-15.6	=	-14.9	+435.2	=	=	=
Q_{30}	+63.4	-118.2	+99.2	=	+53307.6	=	+0.1	+0.2	=	-5.4	-0.8	=	+208.6	=	=	-184
Q_{40}	+46.5	-64.8	+54.8	=	+31833.4	=	+0.02	+0.1	=	-0.8	-0.6	=	=	=	=	-110.2
Q_{50}	-6.6	-49.7	+37.9	=	+25301	=	+0.01	+0.1	=	-1.2	-0.3	-5.8	=	=	=	=
Q_{60}	-24.6	-37.8	+28.1	=	+19403.3	=	+0.006	+0.1	=	-1.1	-0.2	-5.2	=	=	+88.3	=
Q_{70}	-17.3	-25.9	+19.2	=	+13390.9	=	+0.004	+0.1	=	-0.8	-0.1	-3.9	=	=	+62.5	=
Q_{80}	-4.1	-12.2	+9.4	=	+6284.3	=	+0.004	+0.1	=	-0.7	-0.1	-1.7	+19.3	=	+53	-22.4
Q_{90}	-4.9	-4.6	+2.8	=	=	+535.8	+0.002	+0.03	=	-0.4	-0.04	-0.8	=	=	+32.9	=

5.3.3 Leave-One Out Jackknife Cross Validation

The leave-one-out jackknife cross-validation process described in Article 5.2.2 was implemented to check the reliability of the flow quantile MLR models developed for the three clusters. The flow quantile estimate from the validation test was compared with the observed values and performance was evaluated in terms of R^2 (Equation 5.3), RMSE (Equation 5.4), and PBIAS (Equation 5.5).

The range in the values of each of these performance statistics in each cluster for each flow quantile is shown separately in Figure 5.6. Examination of R^2 values reveals that in Cluster – 1, high values are recorded for the median flow quantiles, reasonably better values for high flow quantiles, and low values for the low flow quantiles. This implies that the MLR models developed for this cluster (Table 5.3) may be considered reliable for the prediction of medium to high-flow quantiles in ungauged basins. Examination of results for individual catchments revealed that the unsatisfactory performance for low flow quantiles was due to over-prediction of Q_{70} and Q_{90} quantiles at six catchments, namely Nanipalsan, Kudlur, Thoppur, Seedhi, Ambasamudram, and Salur. Conversely, under-prediction of these quantiles was seen at four catchments (Naguleru, Thengumarhada, Kashipatnam, and Pedagedada). For Cluster – 2, the reliability of the models in terms of R^2 values is reasonably high for high-flow quantiles (except Q_{20}) but poor for the low-flow quantiles (Figure 5.6). Performance was unsatisfactory due to the under-prediction of low flow quantiles at four catchments, namely, Yennehole (Q_{30}, Q_{50}, andQ_{70}), Bantwal (Q_{10}, Q_{30}, andQ_{60}), Adoor (Q_{60}), and Haladi (Q_{60} to Q_{90}). Over-prediction of low flows was observed at Yennehole (Q_{60}), Adoor (Q_{30} and Q_{70}), Kidangoor (Q_{10} and Q_{20}) and Karathodu (Q_{10} and Q_{30}). In Cluster – 3, R^2 values in the range 0.4 – 0.5 were recorded for high flow quantiles but were extremely low for all other flow quantiles. The poor performance of the model is due to over-prediction at catchments - Talikoti (Q_{60} to Q_{90}), Navalgund (all flow quantiles), Halia (all quantiles), K.M.Vadi (Q_{10}), Bendrehalli (all quantiles), Gogenakal (all quantiles), T.Bekuppe (Q_{10} to Q_{20}) and Gadlapet(Q_{10} to Q_{20}). Conversely, under-prediction was observed at catchments Mahuwa (Q_{10} to Q_{20}), Kellodu (Q_{10} to Q_{30}), Balehonnur (all quantiles), E-Mangalam (Q_{10} to Q_{50}), Amabal (all quantiles), Cherribeda (Q_{10}), and Sonarpal (all quantiles).

Values of RMSE on the other hand, indicate higher reliabilities in the jackknife validation. For instance, Figure 5.6 shows that in Cluster – 1, low values of RMSE are evident for all flow quantiles except Q_{10} and Q_{20}. For the MLR models in Cluster–2, RMSE values are high for high-flow quantiles, moderate for median flow quantiles, and low for the low flow quantiles. Low values of RMSE were recorded for all flow quantiles (except Q_{10}) for Cluster–3. The PBIAS values shown in Figure 5.6 indicate that they were less than 50% for all flow quantiles in all three clusters except for Q_{60} in Cluster–2.

Overall, the Jackknife cross-validation procedure provided a mixed response regarding the reliability of the MLR models for the regionalization process. The regression models designed for all three clusters performed very well for high-flow quantiles but were unsatisfactory in predicting the low-flow quantiles. Such weakening of the model for low flow quantiles indicates the possibilities of uncertainties introduced in the low flow values on account of zero flows. Further, leaving out one station at a time during Jackknife cross-validation might affect the model statistics if a highly influential attribute (comparatively higher area, elevation, width, length, etc.) is left out during analysis. A study by Arsenault and Brissette (2016) indicated that the poor performance of regionalization is due to uncertainties in data measurement or incorrect selection of catchment attributes. Hence it is essential to review these stations' information for errors, uncertainties, and other discrepancies for applying appropriate corrections.

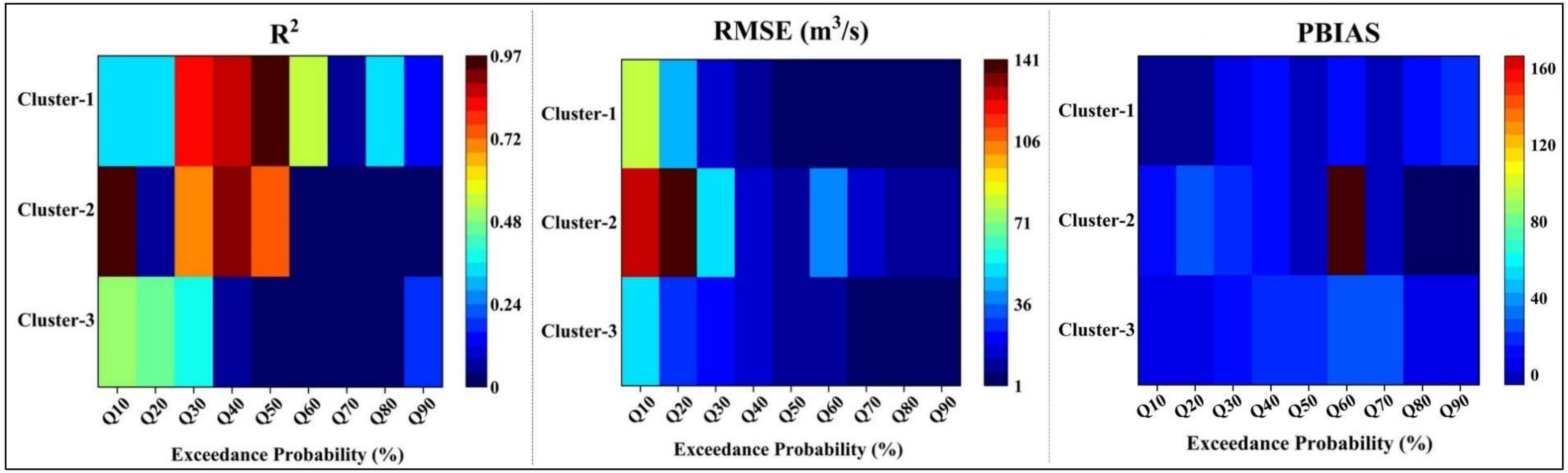


Figure 5.6: Jackknife performance statistics for flow duration MLR models in Clusters – 1, 2, and 3

CHAPTER SUMMARY

- *The final structures of the Multiple Linear Regression (MLR) models, featuring only the most significant predictor variables, were determined through a step-wise regression technique. In the initial phase, MLR models were created assuming that all 50 catchments belonged to a single region. Subsequently, models were developed separately for each of the three clusters, taking into account the catchments within each cluster. Compared to a single region, the models developed for the individual clusters demonstrated strong performance, with average R^2 values for nine flow quantiles reaching 0.85 for Cluster 1, 0.97 for Cluster 2, and 0.80 for Cluster 3.*
- *The implementation of a Jackknife cross-validation technique to assess the reliability of the MLR models indicated consistently strong to satisfactory performance for high flow quantiles across all three clusters. However, the models were found to be unsatisfactory in predicting low-flow quantiles in each cluster.*

CHAPTER 6

CONCEPTUAL HYDROLOGICAL MODELING

6.1 GENERAL

Since historical records of streamflow time series are essential at ungauged sites, the present study also focused on the regionalization of conceptual rainfall-runoff (RR) models. For this purpose, the Australian Rainfall-Runoff Library (RRL) toolkit was identified, from which three popular conceptual RR models namely, AWBM, SIMHYD, and Tank were selected for use in this study.

6.2 COMPILATION AND ANALYSIS OF CLIMATIC DATA

6.2.1 Compilation of Climate Data

In the present study, the gridded daily maximum and daily minimum air temperature data (1^0x1^0) and gridded daily rainfall data (0.25^0x0.25^0) for the period 1981-2018 were downloaded from the Climate Data Service Portal of the India Meteorological Department (IMD), Government of India for grid points located in and around the identified 50 catchments. The downloaded daily rainfall, and daily maximum and minimum air temperature data were used to estimate the daily mean areal values of these variables for the identified catchments using the Thiessen Polygon method. For the case of uniformly spaced grid points, application of the Thiessen method results in square polygons centred around each grid but limited to the catchment boundary as shown in Figure 6.1.

The weighted average of climatological data ($\overline{D}$) over a catchment was computed using Equation 6.1 (Subramanya 2008).

$$\overline{D} = \sum_{i=1}^{m} C_i \left(\frac{A_i}{A_t} \right) = \sum_{i=1}^{m} C_i W_i \tag{6.1}$$

Where m is the number of grid points, A_i is an area of Theissen polygon around each grid point, A_t is the total area of the basin, C_i is the climatological data at the corresponding grid points within that area and W_i is the Thiessen weights at each grid point.

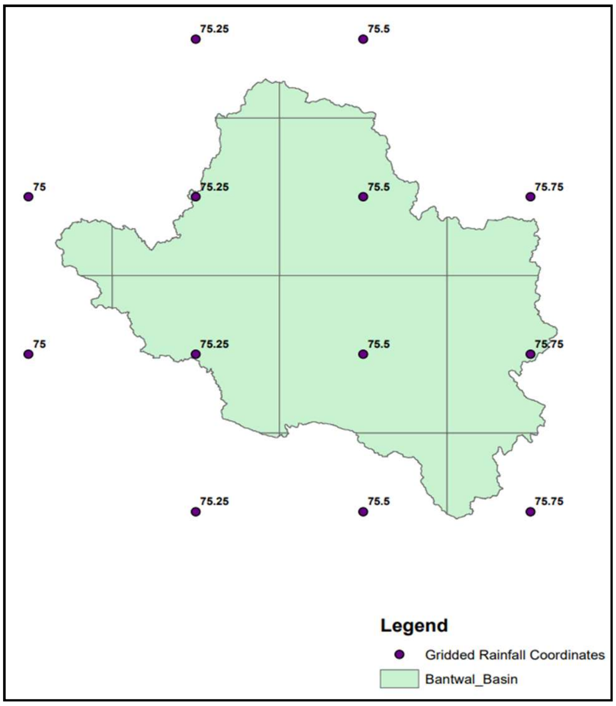


Figure 6.1: Snapshot of Theissen's polygon developed for Bantwal catchment using ArcMap

Theissen polygon method was used to determine the daily average values of all four climate variables: rainfall, maximum and minimum temperature, and reference evapotranspiration. The maximum and minimum temperature data along with rainfall was used for carrying out the trend analysis, while the reference evapotranspiration (derived from surface temperature) and rainfall data were used as an input in rainfall-runoff modeling explained later in this Chapter.

6.2.2 Trend Analysis of Climate Data (Rainfall and Temperature)

While implementing RR models over long historical periods, the question arises as to whether climate change has resulted in changes in the time-series of climate data which are provided as inputs to the models. Specifically, presence of trends in input records can influence the predictions of the RR models and this aspect assumes significance when interpreting the model predicted time-series of runoff. Therefore, in the present study, trend analysis of rainfall and temperature data was carried out using the non-parametric Mann–Kendall test and the Sen Slope Technique.

Mann–Kendall Trend Test

In this investigation, the non-parametric Mann–Kendall test was employed to examine the annual trends in rainfall/ temperature for the identified catchment. This test evaluates the comparative magnitudes of data rather than their specific values (Gilbert 1987; Bisht et al. 2018; Zamani et al. 2018; Anshuman et al. 2019). The advantage of employing this test is that it doesn't require the data to adhere to any specific distribution (Gocic and Trajkovic 2013; Kumar and Nandagiri 2017; Frimpong et al. 2022; Bharath and Venkatesh 2022). The Mann-Kendall test statistic 'M' (Kendall 1975; Mann 1945) can be calculated as follows;

$$M = \sum_{i=1}^{n} \sum_{j=i+1}^{n} sgn\left(x_j - x_i\right) \qquad (6.2)$$

where n is the total number of data points. X_i and X_j are climate data values in time series i and j (where j > i), and sgn ($X_j - X_i$) is the sign function.

$$sgn(x_j - x_i) = \begin{bmatrix} +1, if\ x_j - x_i > 0 \\ 0, if\ x_j - x_i = 0 \\ -1, if\ x_j - x_i < 0 \end{bmatrix} \qquad (6.3)$$

The variance of the data values was computed as:

$$Var(M) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{q} t_i(t_i - 1)(2t_i + 5)}{18} \qquad (6.4)$$

Here, q represents the count of tied groups, and t_i represents the number of data points in the i^{th} group. A tied group refers to a collection of sample data points with identical values. The standard normal test statistic Z_M was calculated utilizing the following equation:

$$Z_M = \begin{bmatrix} \frac{M-1}{\sqrt{Var(M)}}, if\ M > 0 \\ 0 \qquad if\ M = 0 \\ \frac{M+1}{\sqrt{Var(M)}}, if\ M < 0 \end{bmatrix} \tag{6.5}$$

Positive Z_M values signify rising trends, whereas negative Z_M values indicate declining trends. If Z_M exceeds $Z_{\alpha/2}$, where α denotes the selected significance level (e.g., 5%), then the null hypothesis is rejected, suggesting a significant trend (Kumar and Nandagiri 2017).

Sen's slope

The Sen's slope method is a non-parametric technique adopted to quantify the slope of the observed trend in climate data (Gocic and Trajkovic 2013; Kumar and Nandagiri 2017; Frimpong et al. 2022; Bharath and Venkatesh 2022). The Sen slope was calculated using;

$$Q_i = \frac{X_j + X_k}{j-k}\ for\ i = 1,2 \ldots\ldots N \tag{6.6}$$

Where X_j and X_k represent the values of the data at times j and k respectively (where j > k). If there are N values in the time series, then there will be N = n(n-1)/2 slope estimates. Sen's slope is determined as the median slope among these N slope values calculated as:

$$Q_{med} = \begin{Bmatrix} Q_{\left[\frac{N+1}{2}\right]} & If\ N\ is\ odd \\ Q = \frac{1}{2}\left[Q_{\left[\frac{N}{2}\right]} + Q_{\left[\frac{N+2}{2}\right]}\right] & If\ N\ is\ even \end{Bmatrix} \tag{6.7}$$

A positive Q value signifies an upward trend, while a negative value indicates a downward trend in the time series (Kumar and Nandagiri 2017; Bharath and Venkatesh 2022). Sen's slope offers an advantage over the regression slope because it is less influenced by significant data errors and outliers (Sen 1968; Helsel and Hirsch 2002).

Trend analysis of the climate data (rainfall and temperature) in the present study was computed using XLSTAT tool which is an add-in tool in Excel.

6.2.3 Computation of Reference Evapotranspiration

As per Podger (2004), most of the models available in RRL toolkit use the Potential Evapotranspiration (PET) as an input for modelling. However, later research developments have led to the concept of 'reference evapotranspiration' as an alternative to PET and the same was adopted in this study. Reference Evapotranspiration can be estimated by various methods like the Penman-Montieth method, Priestley-Taylor method, Hargraves method, Thornthwaite method, and Blaney Criddle method (Yates et al. 1994). Although the Penman-Monteith method is considered to be the most accurate, it is data-intensive and requires information related to wind speed, air temperature, relative humidity, and solar radiation. Since climate data pertaining to these variables was unavailable in the present study, the Hargreaves equation was used since it requires input data on only minimum and maximum air temperature (Allen et al. 1998). The Hargreaves equation for estimating potential or reference crop evapotranspiration is as follows:

$$ET_o = 0.0023 \times R_a \times (T_{max} - T_{min})^{0.5} \times (T_{mean} + 17.8) \tag{6.8}$$

Where ET_o is reference crop evapotranspiration (mm/d)

R_a is the extraterrestrial radiation in $MJm^{-2}day^{-1}$

T_{max} and T_{min} is a maximum and minimum temperature (^0C)

T_{mean} is the mean temperature (^0C)

Extraterrestrial radiation (R_a) can be computed using the formula,

$$R_a = \frac{24(60)}{\pi} G_{sc} d_r [\omega_s \sin\phi \sin\delta + \cos\phi \cos\delta \sin\omega_s] \tag{6.9}$$

Where G_{sc} is a solar constant (0.0820 MJ m^{-2} min^{-1})

d_r is a contrary space between Earth and the Sun

ω_s is the sunset hour angle in radian

φ is latitude in radian

δ is a solar declination

Here Ra is obtained in units of MJ m^{-2} day^{-1} and is multiplied by 0.408 to yield the equivalent evaporation in mm/day. The latitude is positive for the Northern Hemisphere and negative for the Southern Hemisphere as expressed in radiance. Contrary space between Earth and the Sun d_r can be estimated using the below formula

$$d_r = 1 + 0.033\cos\left(\frac{2\pi}{365}j\right) \tag{6.10}$$

Where j is a Julian day (the number of days in a year between the start of the year and the end of the year)

- Solar declination δ can be computed using the equation mentioned below

$$\delta = 0.409\sin\left(\frac{2\pi}{365}j - 1.39\right) \tag{6.11}$$

- Sunset hour angle ω_s is given by the equation

$$\omega_s = \cos^{-1}\left[-\tan\phi.\tan\delta\right] \tag{6.12}$$

In the present study, the daily reference evapotranspiration data was estimated for the period of records using Equations 6.8 to 6.12 with daily maximum and minimum temperature at the corresponding grid points within the catchment area as inputs. Next, the daily average values of reference evapotranspiration values for the identified catchments were determined using the Thiessen Polygon method as explained earlier in Article 6.2.1.

6.3 RAINFALL-RUNOFF MODELING USING CONCEPTUAL MODELS

As discussed in Article 2.3, distributed hydrological models require input data to be provided at fine spatial resolutions (sub-basin, grid cell) and therefore their applicability in data-scarce regions is impracticable. On the other hand, lumped models consider the entire catchment as a single spatial unit and use spatially averaged inputs. Yet several previous studies have shown that lumped models can simulate the hydrological components with acceptable accuracy (Croke et al.2006; Kunnath and Eldo 2019). Paul et al. (2021) provide a comprehensive review of several available hydrological models and suggest a framework for selecting an appropriate spatial resolution (lumped vs. distributed) for a given purpose.

Lumped conceptual models have been adopted by various water engineers, researchers, and hydrologists for water resources management studies (Boughton and Chiew 2007; Yu and Zhu 2015; Vaze et al.2011; Anshuman et al. 2019). Several such models have been developed across the world for rainfall-runoff modelling; the Tank model (Sugawara et al. 1983), ARNO model (Todini 1996), Australian Water Balance Model (AWBM) (Boughton 2004), SIMHYD (Chiew et al. 2002); Sacramento (Brazil and Hudlow 1981), etc.

Rainfall-Runoff Library (RRL) is a catchment runoff model available under the eWater Toolkit development by the Co-operative Research Center for Catchment Hydrology (CRCCH), eWaters Australia. eWaters is a government-owned not-for-profit enterprise dedicated for promoting sustainable water resources management around the world. The RRL includes five models namely, AWBM, SIMHYD, Tank, Sacramento and SMAR to simulate the runoff using rainfall and evapotranspiration data of the basin. A snapshot of the RRL model is shown in Figure 6.2. Apart from calibration and validation activity, the RRL allows evaluation of different model types, display wettest and driest years, and can carry out parameter sensitivity analysis (Podger 2004).

For the purpose of this study, three models in the RRL toolkit namely, AWBM, SIMHYD and Tank were selected owing to their wide applicability in a variety of world-wide hydroclimatic conditions. The three lumped rainfall-runoff models were individually applied to each of the 50 gauged catchments, utilizing input data comprising average rainfall and reference evapotranspiration. The models were applied for the periods for which historical discharge records were available using a daily time step.

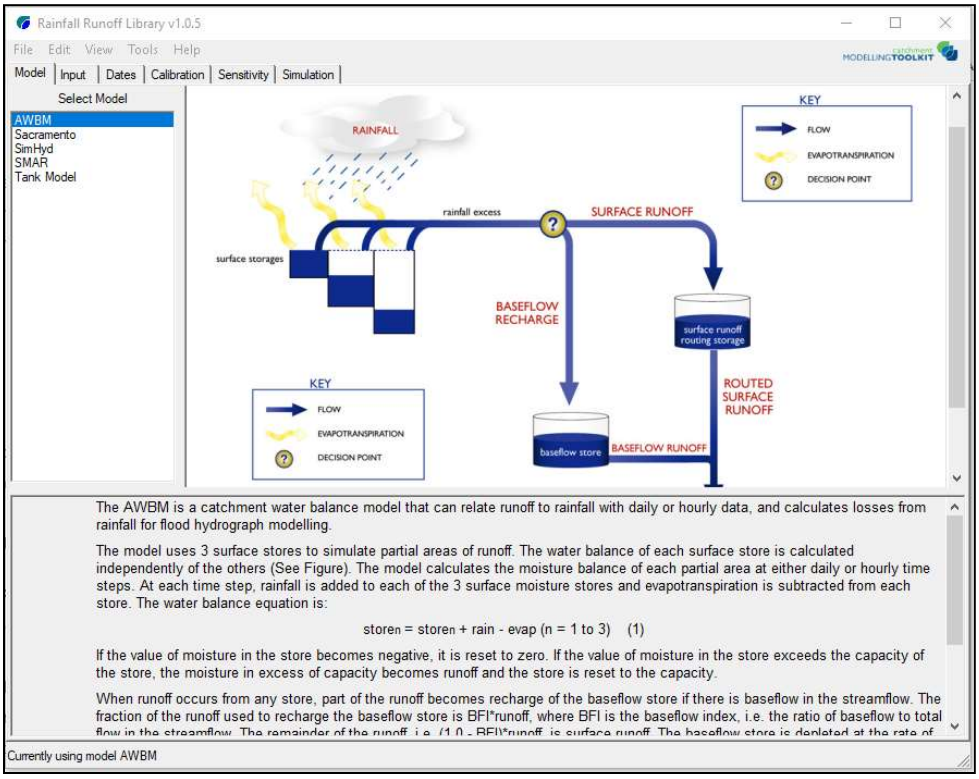


Figure 6.2: Snapshot of RRL model

6.3.1 Australian Water Balance Model (AWBM)

AWBM is a conceptual lumped model having three independent surface stores with capacity C_1, C_2, C_3, and partial areas- A_1, A_2, A_3, and a subsurface store of unlimited capacity as shown in Figure 6.3. The model evaluates the soil moisture of each partial area by adding rainfall and deducting the evapotranspiration at each time step using Equation 6.13.

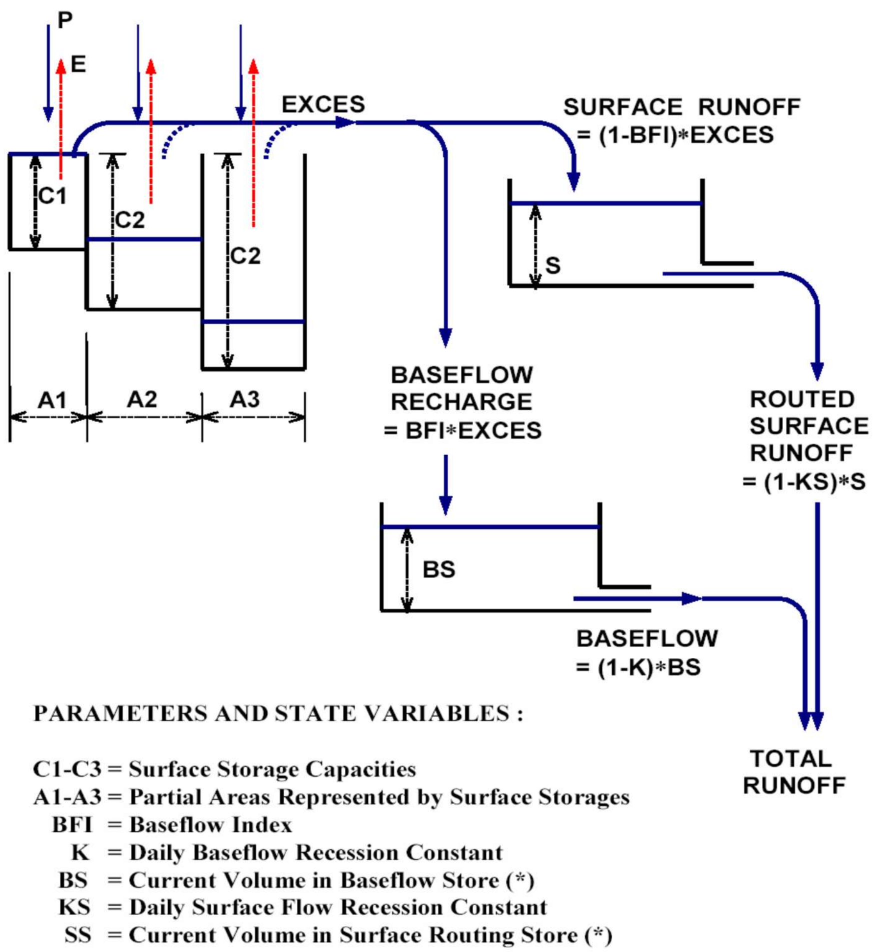


Figure 6.3: Model Structure of AWBM rainfall-runoff model in RRL (Source: Podger 2004)

$$S_n = S_{n-1} + R_n - ET_n \tag{6.13}$$

Where S_{n-1} and S_n = Storage on n-1 and n^{th} day

R_n and ET_n= Rainfall and Evapotranspiration on the n^{th} day

On exceeding the capacity, runoff excess is produced from one or more surface stores. The portion of this excess runoff is converted into baseflow recharge and stored in a baseflow store known as BFI runoff, whereas the remaining (1-BFI)*runoff is drained as surface runoff as shown in Figure 6.3. Here the Base Flow Index (BFI) is the ratio of baseflow to the total flow in the river flow under study.

The surface runoff and the baseflow run down on hourly or daily timescale with the help of surface flow recession co-efficient (KS) and baseflow recession co-efficient (K) as shown in Figure 6.3 (Boughton, 2004; Anshuman et al. 2019; Kunnath and Eldo 2019; Jaiswal et al. 2020). The routed surface runoff and the baseflow from the subsurface store provide the total runoff from the basin. There are a total of eight model parameters in the AWBM model as shown in Table 6.1.

Table 6.1: Parameters of AWBM rainfall-runoff model

Parameters	Descriptions	Range	Units
A_1	Partial area of store 1	0-1	-
A_2	Partial area of store 2	0-1	-
C_1	Capacity of surface store 1	0-50	mm
C_2	Capacity of surface store 2	0-200	mm
C_3	Capacity of surface store 3	0-500	mm
BFI	Base Flow Index	0-1	-
K_{base}	Base flow recession constant	0-1	-
K_{surf}	Surface flow recession constant	0-1	-

(Source: Podger 2004)

6.3.2 SIMHYD Model

SIMHYD is a lumped conceptual daily rainfall-runoff model developed by Chiew et al. 2002. This model is a modified version of HYDROLOG developed in 1972 and MODHYDROLOG in 1991 (Podger 2004). SIMHYD model has three different stores separately for interception, soil moisture, and groundwater. The daily precipitation first replenishes the interception store, which further gets depleted daily through evapotranspiration. The surplus rainfall after satisfying the interception store is either infiltrated in porous layers or forms surface runoff on impervious strata. The rainfall that surpasses the infiltration capacity transforms into infiltration excess runoff as depicted in Figure 6.4.

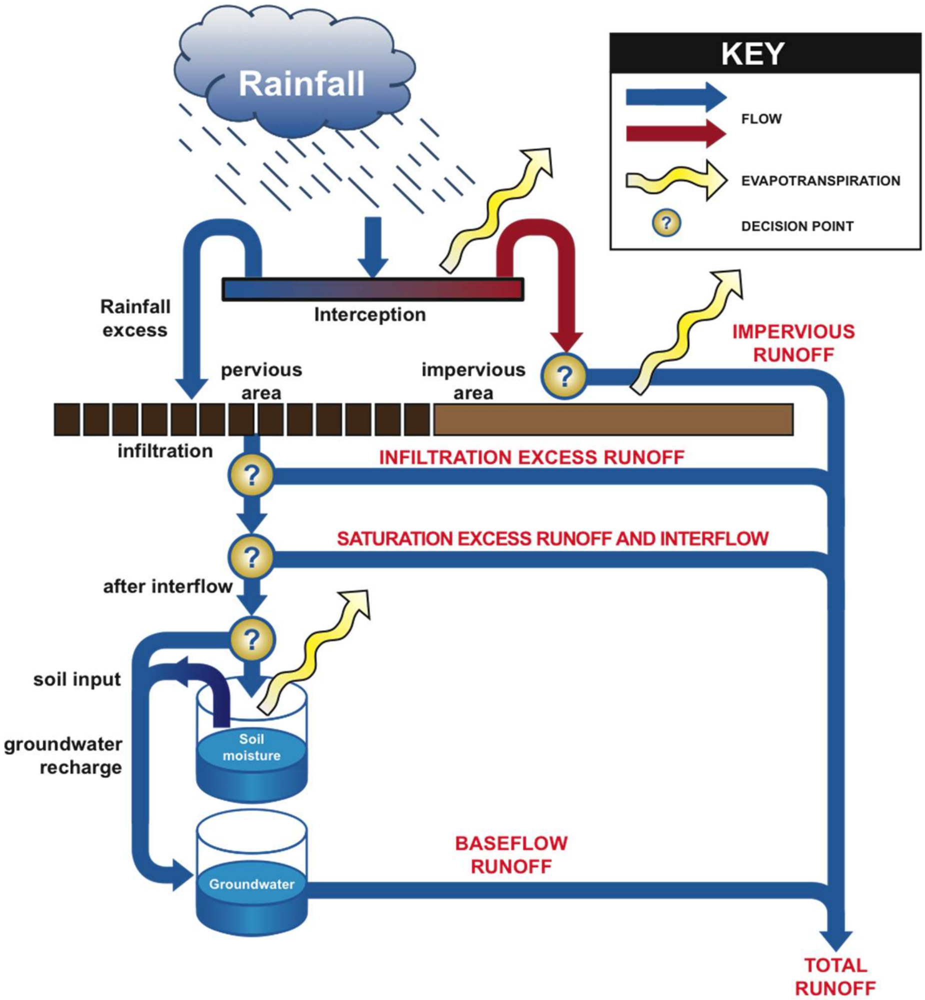


Figure 6.4 Model Structure of SIMHYD rainfall-runoff model in RRL (Source: Podger 2004)

Next, the infiltrated water is further diverted to the stream as interflow, and next to the groundwater store as recharge, and soil moisture store. Initially, the interflow is computed as a linear function of soil wetness (derived from the ratio of soil moisture level to soil moisture capacity). Consequently, the relationship used to simulate interflow attempts to encapsulate both interflow and saturation excess runoff processes.

Similar to interflow, the groundwater recharge component is also computed as a linear function of soil wetness. The residual moisture moves into the soil moisture store. Evapotranspiration from this store is calculated as a linear function of the soil wetness, constrained by the atmospherically controlled rate of areal potential evapotranspiration. The soil moisture store has a limited capacity and any excess overflows into the groundwater store. Base flow from the groundwater store is modeled as a linear recession from the store, contributing to the overall runoff. Therefore, the runoff generation infiltration excess, interflow (and saturation excess runoff), and base flow provide the total runoff from the basin. The basin equations of the model include:

$$IET = \min(pet, (1-PF) * PT, ImI) \tag{6.14}$$

$$IET = \min(PI, pet, RISC) \tag{6.15}$$

$$IC = PF \times ICo * \exp(-IS * SMF) \tag{6.16}$$

$$Infiltration = \min(\text{through fall}, IC) \tag{6.17}$$

$$IR = IntCo * SMF * infiltration \tag{6.18}$$

$$IAF = infiltration - IR \tag{6.19}$$

$$Recharge = RC * SMF * IAF \tag{6.20}$$

$$\text{Soil Input} = IAF - recharge \tag{6.21}$$

Where IET is Impervious ET, PF is Pervious Fraction, PT is Pervious Threshold, ImI is Impervious Incident, IET is Interception ET, PI is Pervious Incident, RISC is Rainfall Interception Store Capacity, IC is Infiltration Capacity, ICo is Infiltration Coefficient, IS is Infiltration Shape and SMF is Soil Moisture Fraction, IR is Interflow Runoff, IntCo is Interflow Coefficient, IAF is Infiltration After Interflow and RC is Recharge Coefficient. SIMHYD model has nine different model parameters as mentioned in Table 6.2.

Table 6.2: Parameters of SIMHYD rainfall-runoff model

Descriptions	Range
Baseflow Coefficient	0-1
Impervious Threshold	0-5
Infiltration Coefficient	0-400
Infiltration Shape	0-10
Interflow Coefficient	0-1
Pervious Fraction	0-1
Rainfall Interception Store Capacity	0-5
Recharge Coefficient	0-1
Soil Moisture Store Capacity	1-500

(Source: Podger 2004)

6.3.3 Tank Model

The tank model is a simple water balance model created by Sugawara et al. 1983 that analyses the daily flow from daily rainfall and evapotranspiration inputs. The RRL Tank model consists of four consecutive vertical tanks. Rainfall and evapotranspiration serve as inputs into the first tank, with subsequent deduction of evaporation from each succeeding tank in series. Here the first tank has two outlets whereas the subsequent tanks have one outlet. The first container yields surface discharge, the second provides intermediate discharge, the third contributes sub-base discharge, and the fourth delivers base discharge as depicted in Figure 6.5. The total runoff is calculated as the summation of runoff from the different tanks using the following Equation 6.22.

$$Q = \sum_{x=1}^{4} \sum_{y=1}^{nx} (C_x - H_x) a_{xy} \tag{6.22}$$

Where Q is the runoff in mm, C_x is the water level in the x^{th} tank, H_{xy} is the height of the y^{th} outlet in the x^{th} tank and a_{xy} is the runoff coefficient.

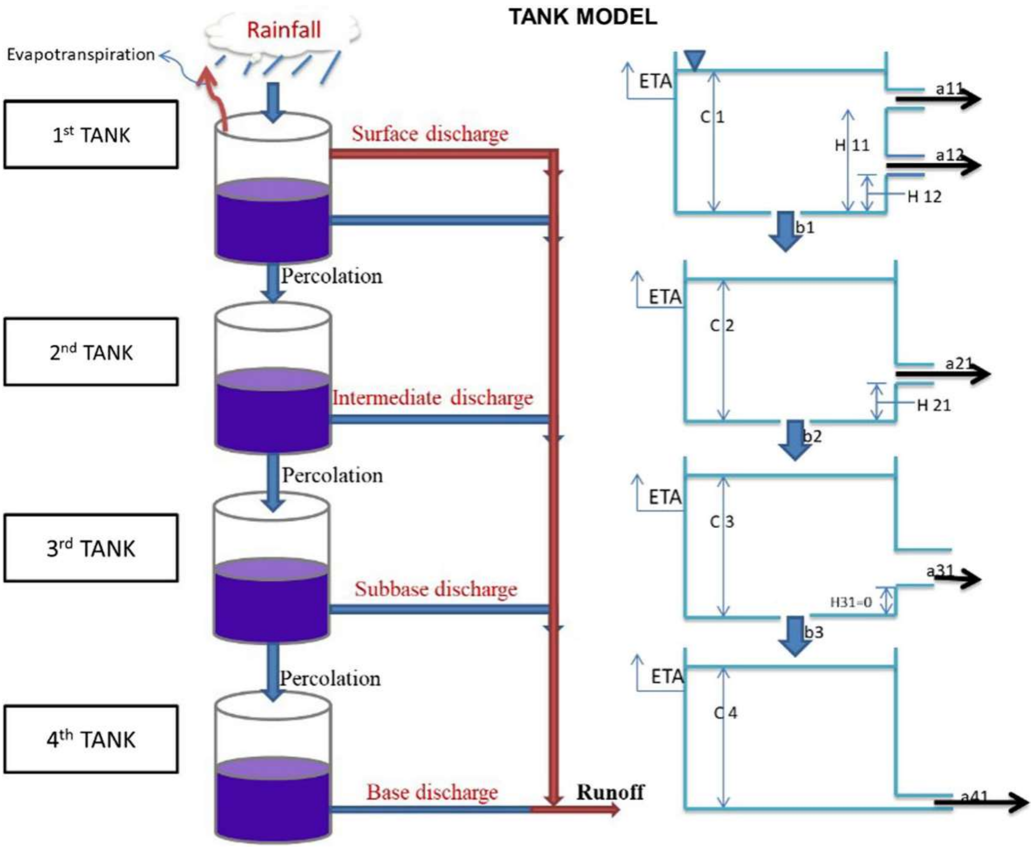


Figure 6.5: Model Structure of Tank rainfall-runoff model in RRL (Source: Podger 2004 and Jaiswal et al. 2020)

The actual evapotranspiration (ET_A) from the individual tank is the loss, which can be computed based on the level of water (C_x) in the x^{th} tank using the following equation given by Bekens (1979).

$$ET_A = ET_p\left(1 - e^{-\alpha\sum C_x}\right) \qquad (6.23)$$

Where ET_p is the potential evapotranspiration in mm and α is the evapotranspiration co-efficient.

The infiltration losses (I_x) in the respective tank can be estimated using the infiltration coefficient (b_x) and water levels (C_x) of the respective tank using Equation 6.24.

$$I_x = C_x \times b_x \qquad (6.24)$$

Compared to AWBM and SIMHYD, the Tank model uses 18 different model parameters as listed in Table 6.3.

Table 6.3: Parameters of Tank rainfall-runoff model

Parameters	Descriptions	Range	Units
H_{11}	Height of initial outlet of first Tank	0-500	mm
H_{12}	Height of the second outlet of the first tank	0-300	mm
H_{21}	Height of outlet of the second tank	0-100	mm
H_{31}	Height of outlet of the third tank	0-100	mm
H_{41}	Height of outlet of the fourth tank	0-100	mm
a_{11}	Runoff coefficient of the first outlet of tank-1	0-1	-
a_{12}	Runoff coefficient of the first outlet of tank-1	0-1	-
a_{21}	Runoff coefficient of the outlet of tank-2	0-1	-
a_{31}	Runoff coefficient of the outlet of tank-3	0-1	-
a_{41}	Runoff coefficient of the outlet of tank-4	0-1	-
α	Evaporation coefficient	0-1	-
b_1	Infiltration coefficient of tank-1	0-1	-
b_2	Infiltration coefficient of tank-2	0-1	-
b_3	Infiltration coefficient of tank-3	0-1	-
C_1	Water level in the tank-1	0-100	mm
C_2	Water level in the tank-2	0-100	mm
C_3	Water level in the tank-3	0-100	mm
C_4	Water level in the tank-4	0-100	mm

(Source: Podger 2004 and Jaiswal et al. 2020)

6.3.4 Model Calibration

RRL tool supports different types of optimizers that include Genetic Algorithm (GA), Pattern Search Multi-Start (PS-Multi), Uniform Random Sampling (URS), Rosenbrock Single/Multi-Start Optimizer and Shuffled Complex Evolution-Developed by University of Arizona (SCE-UA) to assist in calibration of models. These optimizers run using different objective functions that will give the calibration process a bias towards a range of flows that are significant. The model also has eight different objective functions including the Nash-Sutcliffe criterion (Coefficient of efficiency), Sum of square errors, Root mean square error (RMSE), Root mean square difference about bias, Absolute value of bias, Sum of square roots, Sum of square of the difference of square root and Sum of absolute difference of the log (Podger 2004). The calibration starts with some default values of the model parameters. Based on the fitness values of the complexes, they are shuffled and evolved after each iteration. The model runs for a designated number of simulations. Research studies using the RRL toolkit have most commonly adopted Nash–Sutcliffe Efficiency (NSE) as the objective function for the calibration process (Yu 2015; Yu and Zhu 2015; Anushman et al. 2019; Esmaeili-Gisavandani et al. 2021). Hence, the same approach was adopted in all three models for the calibration process.

In a recent paper, Shen et al. (2022) showed that calibrating a hydrological model for a full period produces robust findings compared to the split sampling technique wherein the calibration of the model is done with old data and then validating with recent data resulting in inferior performance. Hence to compare the model results, the rainfall-runoff modeling was initially executed for split sampling technique and then subsequently tested for full period calibration.

6.3.5 Model Performance Evaluation

The performances of the conceptual hydrological models were evaluated by comparing the model predicted flows with the observed flow values corresponding to the catchment under investigation. In addition to the coefficient of determination (R^2) (defined earlier in Article 5.2.3), the Nash-Sutcliffe co-efficient (NSE), and Kling-Gupta efficiency metric (KGE) were also used to evaluate model performances.

$$NSE = 1 - \frac{\sum_{i=1}^{N} (O_i - P_i)^2}{\sum_{i-1}^{N} (O_i - \overline{O})^2} \tag{6.25}$$

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_p}{\sigma_o} - 1\right)^2 + \left(\frac{\mu_p}{\mu_o} - 1\right)^2} \tag{6.26}$$

In the above equations, ‘N’ denotes the number of days, ‘O’ and ‘P’ are the observed and predicted flow values, and, $\overline{O}$ indicate the average values of observed flow rates. NSE metric indicates how well the simulated results match with the observed data. NSE metric ranges from -∞ to 1, where a value of 1 signifies ideal model performance (Moriasi et al. 2015). KGE is a weighted combination of the three constitutive components (i.e., correlation, variability bias, and mean bias) decomposed from the Nash-Sutcliffe efficiency formula. In KGE, r is the linear correlation between observed and simulated flows, σ_O and σ_S are the standard deviations, and μ_O and μ_S are the mean of the observed and simulated flows respectively. The KGE value ranges from −∞ to 1.0 where KGE = 1.0 indicates perfect agreement between simulations and observations, and KGE >0.3 indicates behavioural performance (Wouter et al. 2019; Shen et al. 2022).

6.4 DOMINANT MODEL PARAMETERS

Conceptual hydrological models are classified as grey-box models since the parameters embedded within the model structure can be interpreted with some degree of physical-basis. This is in contrast to empirical models which are black-box in nature and model parameters are purely used as ‘fitting’ coefficients without any physical meaning (Brath et al. 2002). Mechanistic models on the other hand are considered white-box in nature since they rely on physics-based equations and parameters.

Parameters appear in conceptual models in two contexts – 1) as partitioning coefficients which divide a given input into a model component into two hydrologic variables (e.g., rainfall input to the upper soil store into surface runoff and infiltration) and 2) as a representation of the storage of a given component which in turns defines the magnitude of an output hydrologic variable (e.g., soil moisture storage and its effect on the rate of evapotranspiration). While it is true that such parameters in conceptual models also serve as fitting coefficients which are optimized to match model simulations with observations of the final output variable, they can still be interpreted in the contexts defined above. As Sivapalan et al. (2003b) note, parameters in such models can have conceptual meaning but cannot be measured in the field.

Also, not all parameters of a conceptual hydrological model need have the same effect on the output with some parameters being more sensitive than the others. A variety of perturbation-based sensitivity analysis methods ranging from simple trial-and-error to sophisticated probabilistic Monte Carlo techniques are used for the purpose of identifying the relative sensitivities of the model output to the model parameters (Tang et al. 2007; Devak and Dhanya 2017). Evidently, the model output (typically streamflow) will be more sensitive to those parameters which are involved with the dominant model components (e.g., surface runoff, subsurface flow, groundwater flow, evapotranspiration) which affect it most for the given hydrometeorological and topographical conditions existing in the catchment (Atkinson et al. 2002). Therefore, in a typical conceptual hydrological modelling exercise carried for a catchment, parameter sensitivity analysis becomes a necessary and an integral part of the exercise for two reasons – 1) to identify the most sensitive model parameters so that only they need to be calibrated in the subsequent step with the non-sensitive parameters being set to default values. In this manner, the number of model parameters which need to be calibrated gets significantly reduced 2) once the sensitive model parameters are identified and calibrated, the hydrological processes/variables with which they are associated may then be classified as being dominant in that catchment (e.g., Medina and Munoz 2020).

However, when applying more than one model to a large number of catchments the conventional sensitivity analysis approach becomes computationally intensive and also the problem of equifinality may be encountered. Therefore, in the present study, involving application of 3 models to 50 catchments, a more empirical approach involving the use of box-whisker plots for assessment of variabilities of the optimised model parameters across the catchments was used. The assumption was that a model parameter could be considered more sensitive if it exhibited higher variability across the catchments and could be considered insensitive otherwise. Such model parameter variabilities were assessed separately for catchments in each of the 3 clusters created using hierarchical cluster analysis (Chapter 4).

6.4.1 Parameters of the Conceptual Hydrological Models

The conceptual hydrological models AWBM, SIMHYD, and Tank model were applied to all 50 delineated basins for a specified periods of time and calibrated as described in Article 6.3. As mentioned in Article 6.2, the AWBM model has a total of eight model parameters, SIMHYD has nine and Tank has eighteen model parameters that were optimized to obtain the desired results (Figure 6.6). The optimized model parameters from AWBM, SIMHYD, and Tank models were compiled and analysed for their variability and thereby, sensitivity.

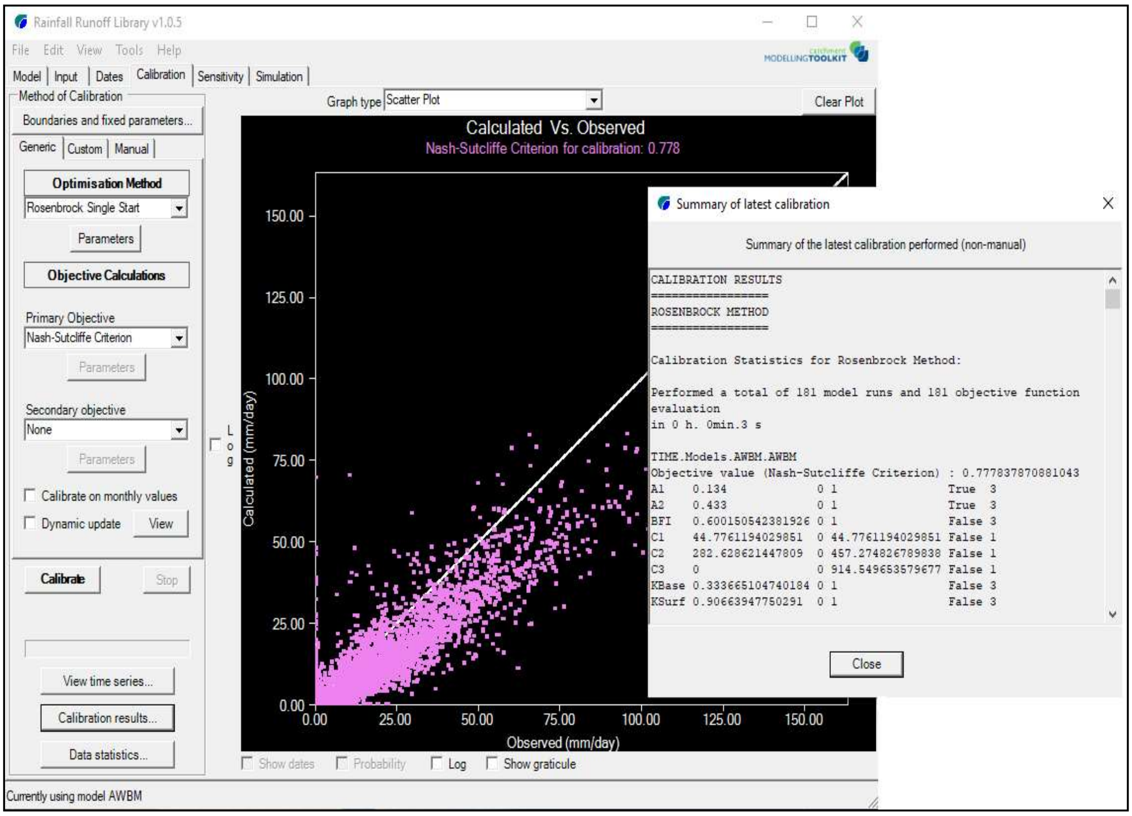


Figure 6.6: Snapshot of calibration statistics of model parameters in the RRL tool

Apart from optimizing the model parameters, the RRL toolkit also provides the values of the dominant model variables that can utilized for further hydrological investigation. The statistical analysis of dominant model parameters obtained from the RRL tool can be carried out cluster-wise to obtain regional model parameters. Such regional model parameters can be utilized in the prediction of flows in ungauged basins located within the associated homogenous cluster.

6.5RESULTS AND DISCUSSIONS

6.5.1 Trend Analysis of Climate Data (Rainfall and Temperature)

Trend analysis of the climate data revealed that the rainfall in majority of the catchments has a statistically insignificant trend (with significance level of 1%) except in four catchments; Thumpaman, Ayilam, Pulmanthole and Balehonnur. The Sen slope estimates indicate that the majority of the basins in the west-flowing river, Krishna and Godavari show a decreasing trend of rainfall. The maximum value of decreasing increment (-50.4mm/year) was observed in Balehonnur in Krishna basin and the maximum increasing trend of 50.5 mm/year was seen in Avershe catchment of west flowing river basin. The increasing and decreasing trend of rainfall at identified catchments is illustrated in Figure 6.7. Overall, the trend was insignificant, therefore rainfall in selected catchments can be assumed to be climatological stationary for conducting hydrological analysis.

Conversely, the maximum and minimum temperature data in the majority of the catchments have statistically significant trends (with significance level of 1%). Overall, the maximum temperature indicates an increasing trend (average 0.013 °C/Year) in all catchments, whereas the minimum temperature shows an increasing trend (average 0.009 °C/Year) in majority of catchments. The rising and decreasing trends were observed only in minimum temperature as shown in Figure 6.8. However, Mann Kendall and Sen slope values for maximum and minimum temperature are quite small and may not affect the overall results significantly.

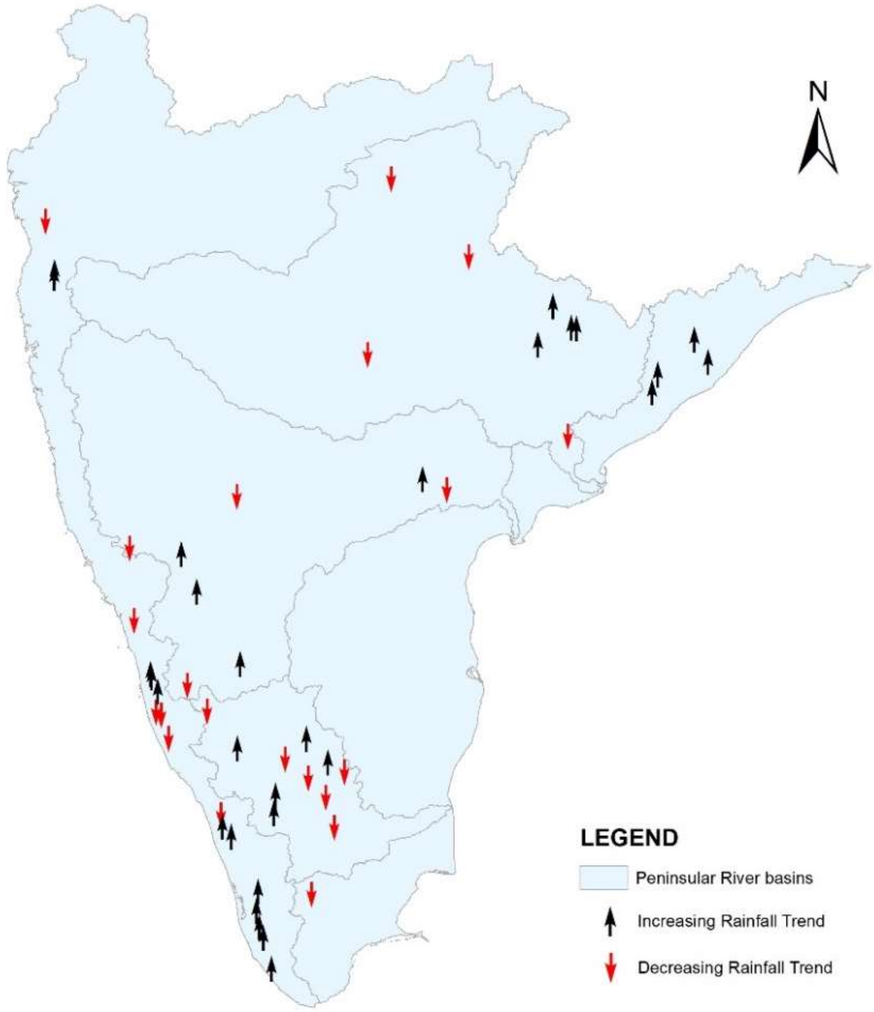


Figure 6.7 Trend analysis of rainfall at identified catchments

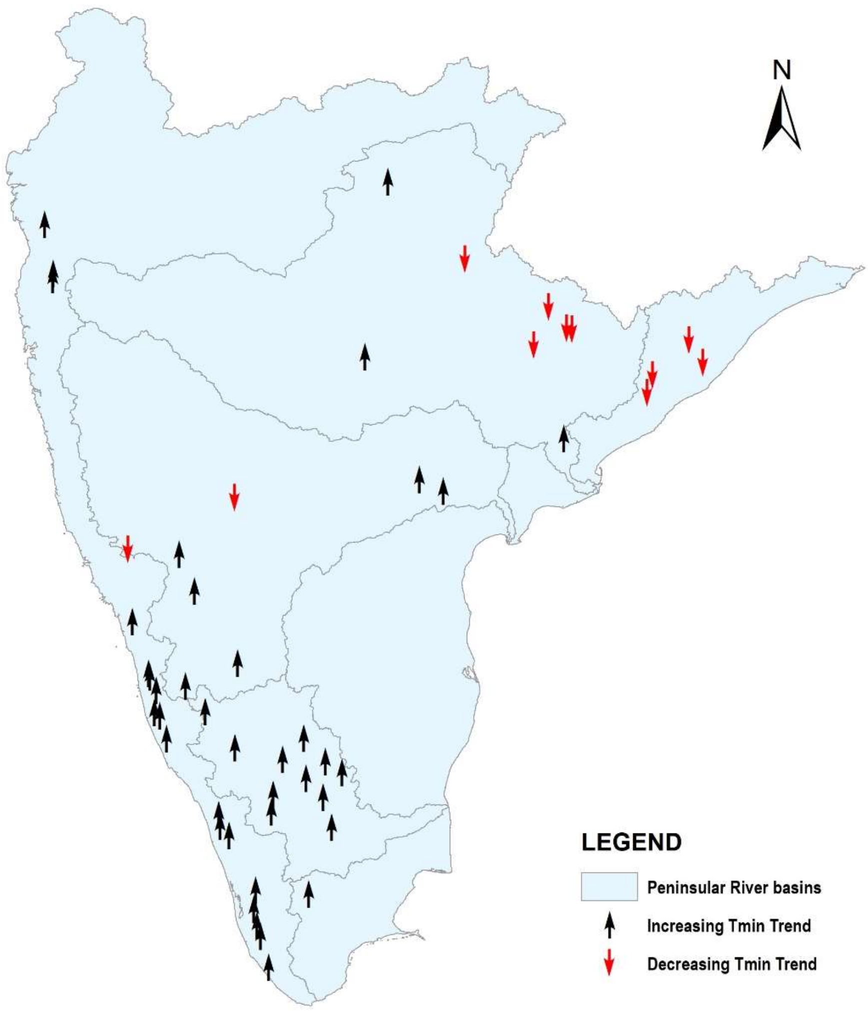


Figure 6.8 Trend analysis of minimum temperature at identified catchments

6.5.2 Conceptual Hydrological Modeling

6.5.2.1 Implementation of the Models

Rainfall-runoff modeling for all 50 gauged basins was carried out using AWBM, SIMHYD, and Tank models. Initially, the area and the description of the catchment are provided to set up the simulation file. Next daily time-series data related to rainfall, evapotranspiration, and flow is uploaded in dot (.) cdt format, and then the time period command is utilized to fix the warm-up, calibration, and validation period of model simulation. Models were set up based on the availability of daily time series data and calibrated with daily Nash Sutcliffe Efficiency (NSE) as the objective function criteria to optimize the simulated values. As discussed in section 6.3.4, RRL offers six distinct optimizers to aid in model calibration. Each chosen model underwent catchment-wise simulation with all six optimizers, and the optimal technique was determined based on the desired calibration outcome. Ultimately, Shuffled Complex Evolution (SCE-UA), Pattern Search Multi-start (PS-Multi), and Rosenbrock techniques emerged as the most effective optimizers in this investigation for attaining the desired results. The results of the simulated model were downloaded for further analysis. Apart from the rainfall-runoff modeling, the RRL tool also aids in the analysis of the dominant model parameters. The flowchart representing the implementation of the conceptual models is illustrated in Figure 6.9.

The conceptual hydrological modeling was carried out separately in each of the identified unregulated catchments for the available period of records. Although the performance of individual models was assessed separately, the overall evaluation of model performance was based on the previously established hydrological homogeneous clusters (outlined in Chapter 4). This approach aimed to offer insights into the variability of optimal model parameters within each hydrologically homogenous cluster. Subsequently this also permitted identification of average model parameters which are applicable for ungauged basins within each cluster.

As discussed in 6.3.4, the rainfall-runoff modeling was initially executed for split sampling technique and then subsequently tested for full period calibration. The performance of each of the models was evaluated in terms of Nash-Sutcliffe Efficiency (NSE), correlation coefficient (R^2), and Kling-Gupta's Efficiency (KGE) metric. The details of results obtained from each of the models are explained cluster-wise.

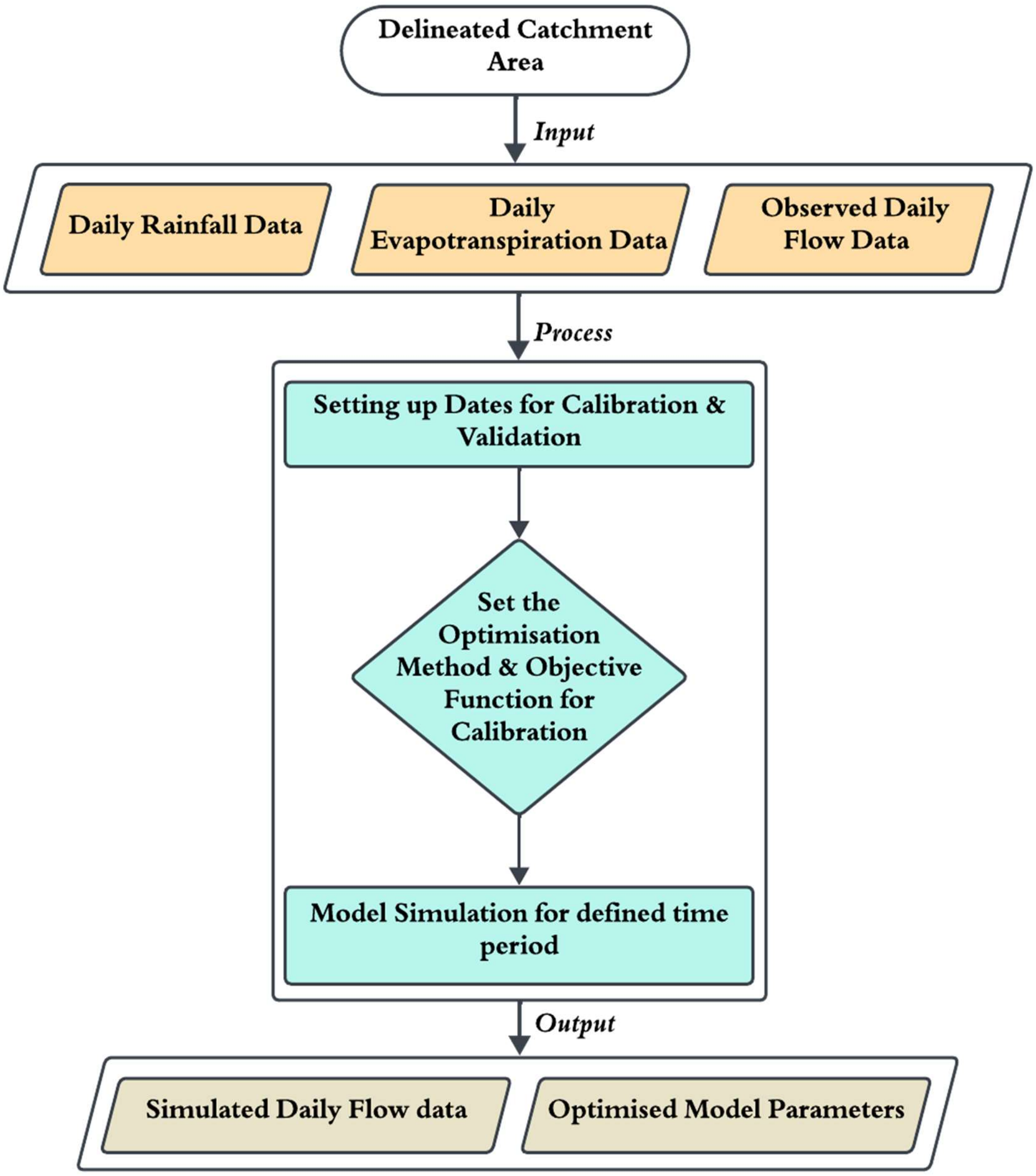


Figure 6.9: Flowchart representing the implementation of the models

6.5.2.2 Conceptual hydrological modeling results of Cluster-1 catchments

AWBM Model:

Out of 17 catchments in Cluster -1, the model performance in 11 catchments (except Naguleru, Nellithurai, Kashipatnam, Ambasamudram, Salur, and Pedagedadda) was found satisfactory in terms of NSE and R^2 during split-sampling calibration period trials (Figure 6.10). In terms of KGE, most of the model simulations except for Naguleru, Thoppur, Nellithurai, Salur, and Pedagedadda were found satisfactory during calibration sub-periods.

However, only four catchments (Naguleru, Nellithurai, Kashipatnam, and Pedagedadda) exhibited lower performance in terms of R^2 and KGE during the full calibration trials (refer to Figure 6.11). Additionally, in terms of NSE, Salur and Ambasamudram catchments, along with the aforementioned four catchments, displayed under-performance during the full calibration trials. Nevertheless, it was noted that all basins within the west-flowing river demonstrated satisfactory model performance, suggesting the potential utility of AWBM model in this region.

SIMHYD Model:

Except for Nine catchments including Naguleru, Kudlur, Thevur, Nellithurai, Thengumarhada, Kashipatnam, Ambasamudram, Salur, and Pedagedadda, the model performance in other catchments was found satisfactory in terms of NSE and R^2 during the split calibration period. Conversely, only five catchments viz., Naguleru, Thoppur, Nellithurai, Salur, and Pedagedda underperformed in terms of KGE during split-sampling calibration period trials (Figure 6.10). Compared to the AWBM model, most of the catchments in Krishna, Cauvery, East flowing and Godavari showed poor model performance in terms of NSE, R^2, and KGE during split-sampling validation period trials.

Throughout the full calibration period, Naguleru, Thevur, Nellithurai, Thengumarhada, Kashipatnam, Ambhasamudram, Salur, and Pedagedadda exhibited inadequate performance in terms of NSE and R^2. Conversely, the SIMHYD model showcased satisfactory to very good performance in terms of KGE across ten catchments excluding Naguleru, Thevur, Nellithurai, Kashipatnam, Ambasamudram, Salur, and Pedagedadda during the full calibration period (as shown in Figure 6.11).

Tank Model:

During the split calibration period, nine catchments—Naguleru, Kudlur, Thevur, Nellithurai, Thengumarhada, Kashipatnam, Ambasamudram, Salur, and Pedagedadda—exhibited inadequate model performance in terms of NSE and R^2. On the contrary, six catchments viz., Naguleru, Thevur, Thoppur, Nellithurai, Salur, and Pedagedadda underperformed in terms of KGE during split-sampling calibration period trials (Figure 6.10). Similar to SIMHYD, most of the catchments showed poor model performance in terms of NSE, R^2, and KGE during split-sampling validation period trials.

During the full calibration period, Naguleru, Thevur, Nellithurai, Thengumarhada, Kashipatnam, Ambhasamudram, Salur, and Pedagedadda showed unsatisfactory performance in terms of NSE and R^2. In line with the split sampling simulation, seven catchments—Naguleru, Kudlur, Thoppur, Nellithurai, Kashipatnam, Ambasamudram, and Pedagedadda—exhibited inadequate performance in terms of KGE during the full calibration trials (refer to Figure 6.11).

Note: For Graphical representation the seventeen catchments are serially numbered along x-axis (Figure 6.10 and 6.11) as: (1)- Thumpaman; (2)- Ayilam; (3)-Kuniyil; (4)-Nanipalsan; (5)-Ozerkheda; (6)-Pulmanthole; (7)- Naguleru; (8)-Kudlur; (9)-Thevur; (10)-Thoppur; (11)-Nellithurai; (12)-Thengumarhada; (13)- Kashipatnam; (14)- Seedhi; (15)- Ambasamudram; (16)- Salur; and (17)-Pedagedadda.

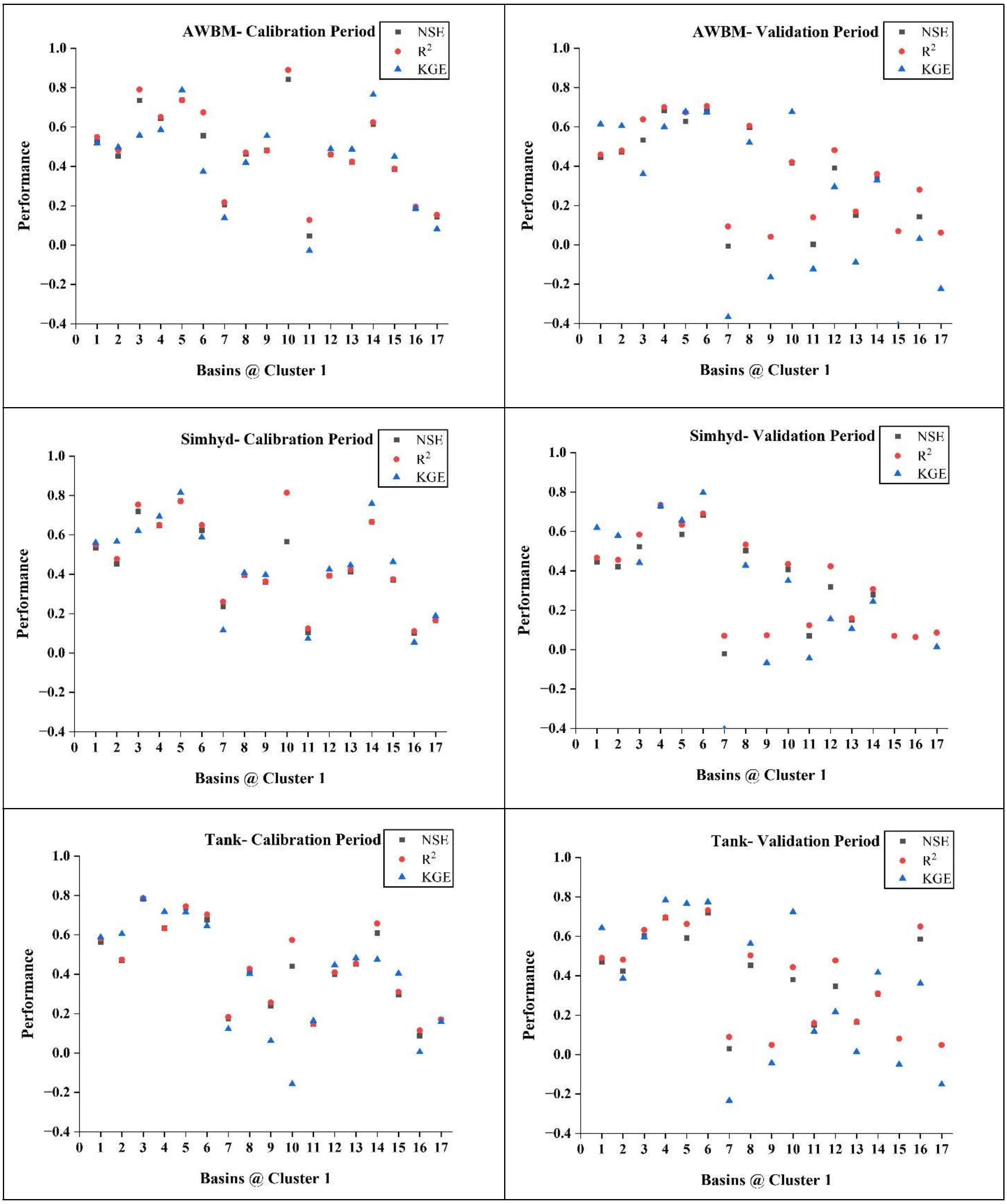


Figure 6.10 Model Performance Results of AWBM/SIMHYD/Tank model for Cluster-1 during the split sampling calibration period

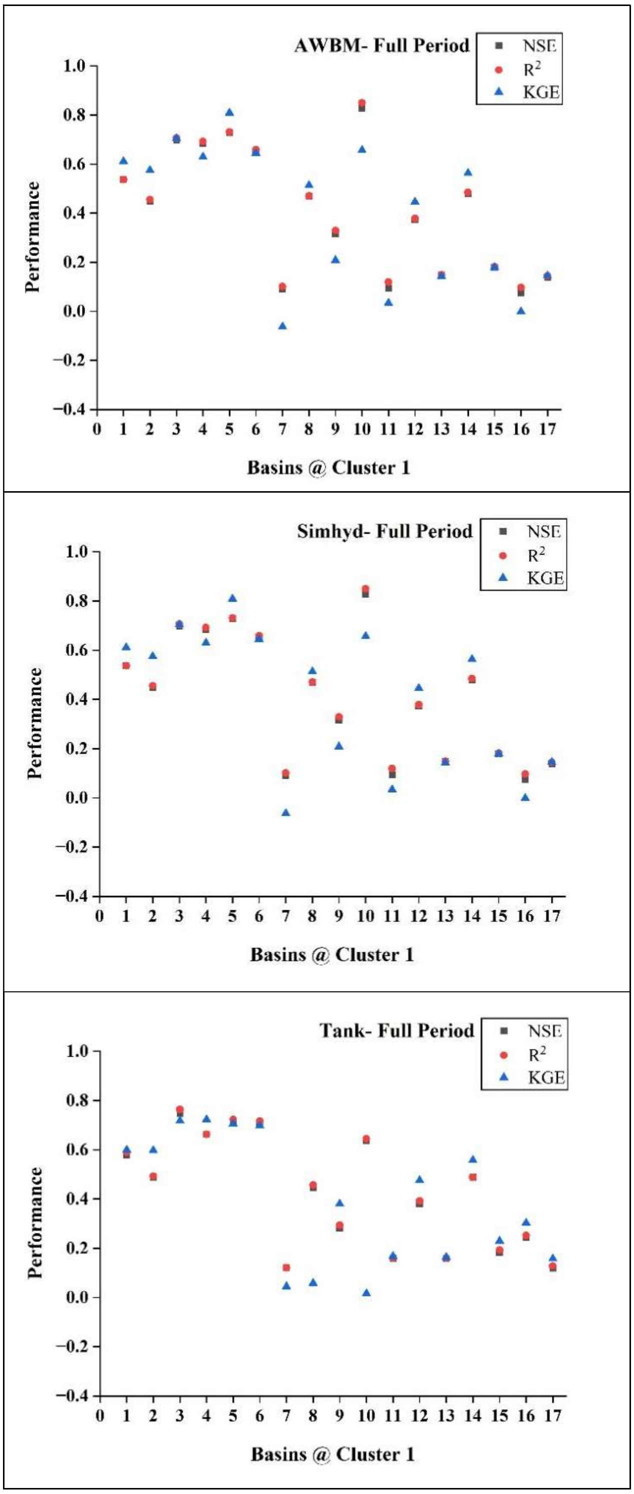


Figure 6.11 Model Performance Results of AWBM/SIMHYD/Tank model for Cluster-1 during the full calibration period

6.5.2.3 Conceptual hydrological modeling results of Cluster-2 catchments

AWBM Model:

All eleven catchments performed efficiently in terms of NSE, R^2, and KGE metrics during split-sampling calibration period trials (Figure 6.12). Similarly, model simulation for all the catchments except Santeguli in Cluster-2 showed very good to excellent performance during full calibration trials as shown in Figure 6.13.

SIMHYD Model:

Similar to the AWBM model, all the catchments performed efficiently in terms of NSE, R^2, and KGE metrics during split-sampling calibration period trials and full calibration period trials (Figures 6.12 and 6.13). Apart from other catchments, the performance of the Santeguli catchment in cluster-2 showed slightly lesser model performance during full calibration simulation trials as shown in Figure 6.13.

Tank Model:

All eleven catchments performed efficiently in terms of NSE, R^2, and KGE metrics during split-sampling and Full-period calibration period trials (Figures 6.12 and 6.13). This suggests that all three models can be reliably utilized for predicting the flows of Cluster-2 catchments, which primarily consist of west-flowing rivers.

Note: For Graphical representation the eleven catchments are serially numbered along x-axis (Figure 6.12 and 6.13) as: (1)- Santeguli; (2)- Avershe; (3)-Yennehole; (4)-Adoor; (5)-Bantwal; (6)-Erinjipuzha; (7)- Kidangoor; (8)-Kalloopara; (9)-Karathodu; (10)-Kalampur; and (11)-Haladi.

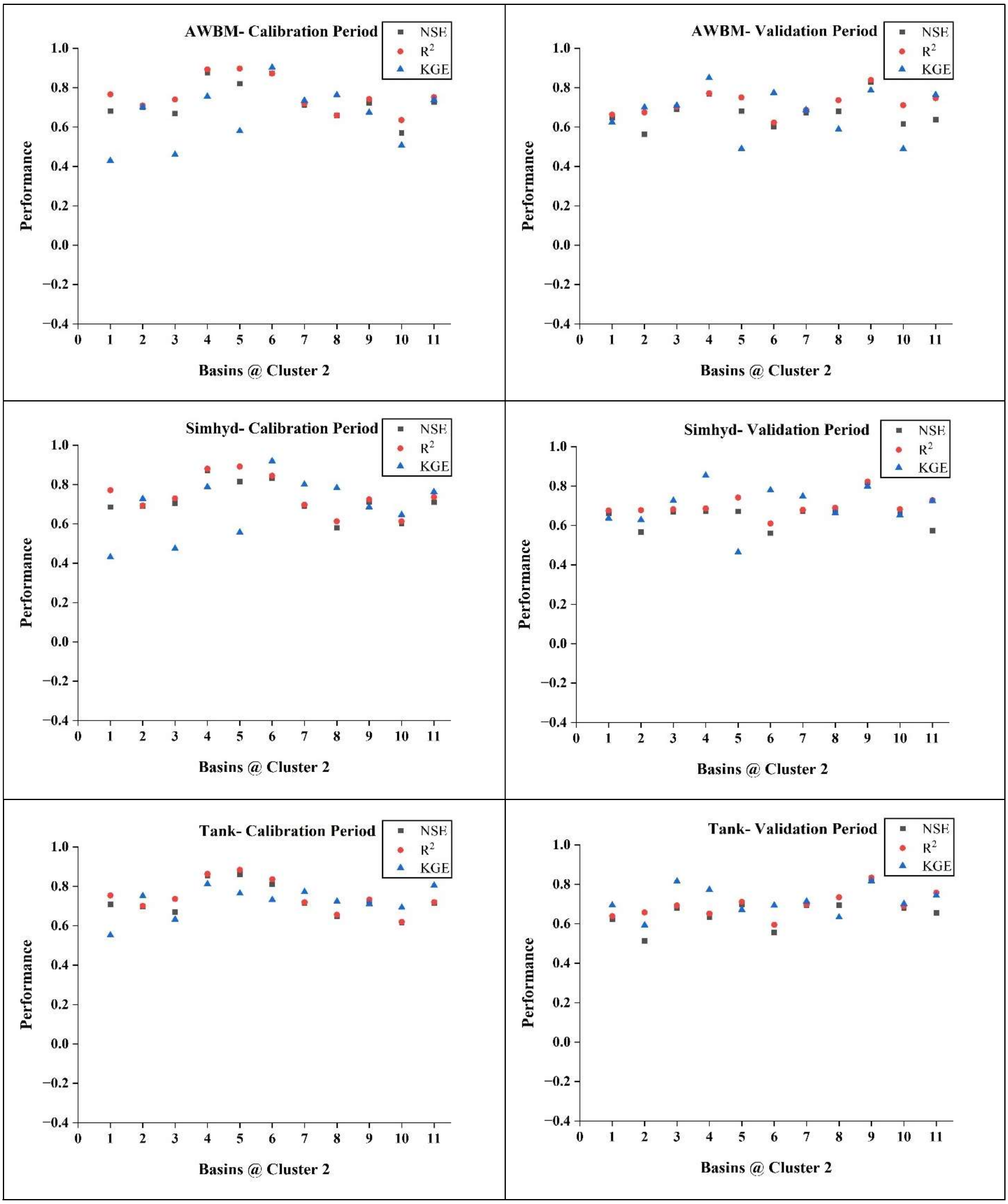


Figure 6.12 Model Performance Results of AWBM/SIMHYD/Tank model for Cluster-2 during the split sampling calibration period

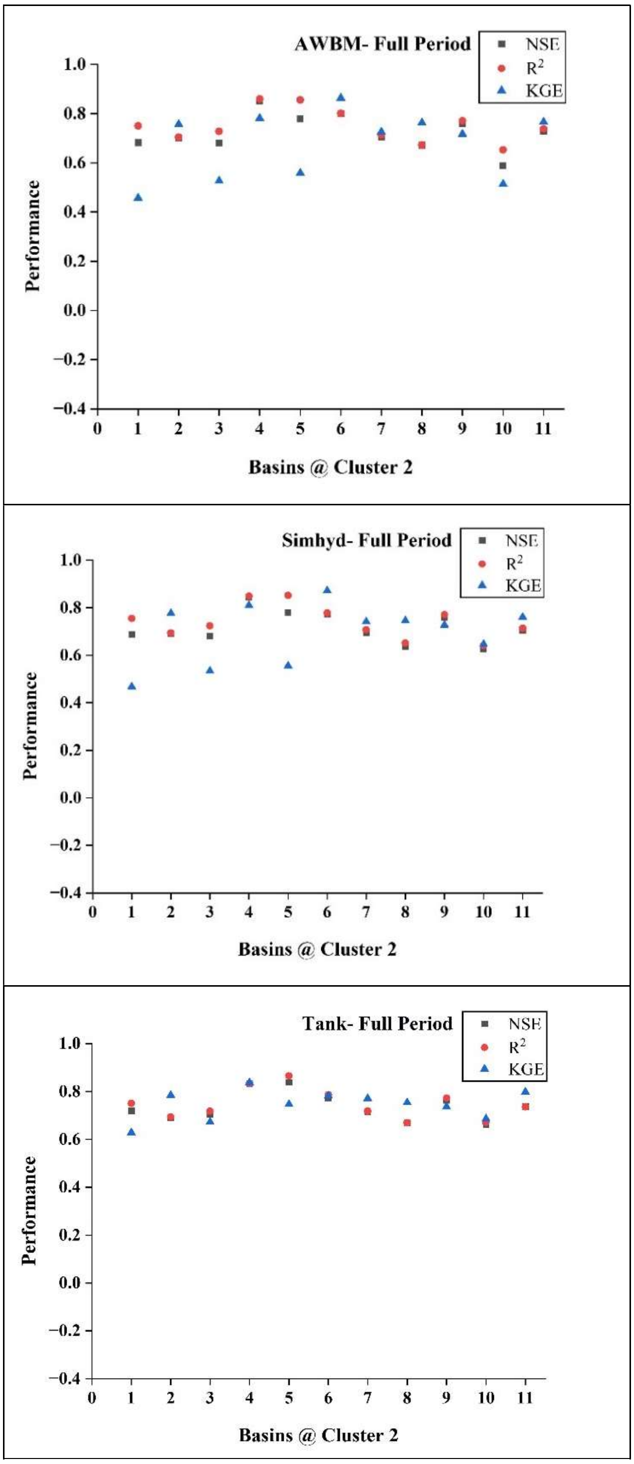


Figure 6.13 Model Performance Results of AWBM/SIMHYD/Tank model for Cluster-2 during the full calibration period

6.5.2.4 Conceptual hydrological modeling results of Cluster-3 catchments

AWBM Model:

Out of 22 catchments, the model performance in 20 catchments (except Bendrehalli and T.Bekuppe) performed satisfactorily in terms of NSE and R^2 during split-sampling calibration period trials (Figure 6.14). On the other hand, the model simulations of the Khanapur and Marol basins showed under-performance in terms of KGE metric during the Split sampling period.

Throughout the full calibration period KM Vadi, E_Mangalam, Bendrehalli, and T. Bekuppe displayed insufficient performance in terms of NSE and R^2. In contrast, the AWBM model demonstrated satisfactory to very good performance in terms of KGE across nineteen catchments, with the exceptions of Khanapur, E_Mangalam, and T_Bekuppe, during the full calibration period (as illustrated in Figure 6.15). This discrepancy might be attributed to the presence of significant outliers in flow values, which hinder the simulation process.

SIMHYD Model:

Except for five catchments including KM Vadi, E-Mangalam, Bendrehalli, Hogenakkal, and T.Bekuppe, the other catchments' model performance was found satisfactory in terms of NSE and R^2 during the split calibration period. Conversely, three catchments viz., Halia, E-Magalam, and T.Bekuppe underperformed in terms of KGE during split-sampling calibration period trials (Figure 6.14).

For the full calibration period, seven catchments including Marol, KM Vadi, E-Mangalam, Bendrehalli, Hogenakkal, T.Bekuppe, and Sonarpal exhibited inadequate performance in terms of NSE and R^2. In contrast, only three catchments—KM Vadi, E-Mangalam, and T. Bekuppe - indicated poor performance during the full calibration period trials, as illustrated in Figure 6.15.

Tank Model:

Compared to all catchments, five catchments including Halia, KM Vadi, E-Mangalam, Bendrehalli, and T. Bekuppe showed poor model performance in terms of R^2 during the split calibration period. Six catchments, including Navalgund and the aforementioned five, showed unsatisfactory performance in terms of NSE during the split sampling process (refer to Figure 6.14). In terms of KGE, all catchments performed satisfactorily during the split sampling period except for two catchments including Kellodu, and Halia.

Seven catchments, namely Navalgund, KM Vadi, E_Mangalam, Bendrehalli, Hogenkkal, T. Bekuppe, and Sonarpal exhibited inadequate performance in terms of NSE and R^2 during full calibration trials. However, in the case of full calibration period simulation trials, all catchments except E-Mangalam, Bendrehalli, and T. Bekuppe displayed satisfactory to very good performance (Figure 6.15).

Note: For Graphical representation, the twenty-two catchments are serially numbered along the x-axis (Figure 6.14 and 6.15) as: (1)- Mahuwa; (2)- Kellodu; (3)-Talikoti; (4)-Navalgund; (5)-Ballehonnur; (6)-Khanapur; (7)- Marol; (8)-Halia (9)-Sakleshpur; (10)-KM Vadi; (11)-E-Mangalam; (12)-Bendrehalli; (13)- Hogenakkal; (14)-T_Bekuppe; (15)-Gunupur; (16)- Ramakona; (17)-Wairagarh; (18)-Amabal; (19)-Tumnar; (20)-Cherribeda; (21)- Gandlapet and (22)-Sonarpal.

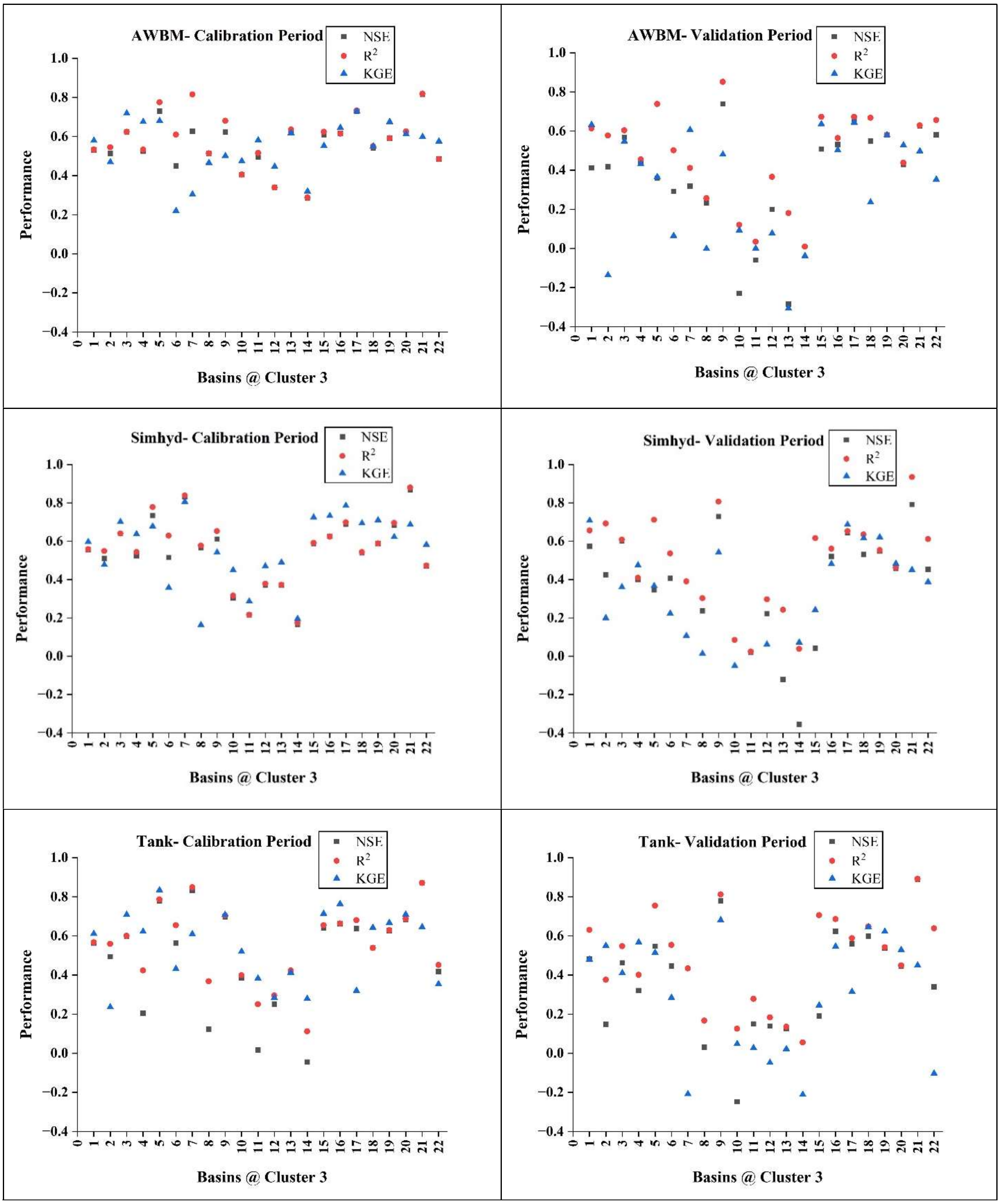


Figure 6.14 Model Performance Results of AWBM/SIMHYD/Tank model for Cluster-3 during the split sampling calibration period

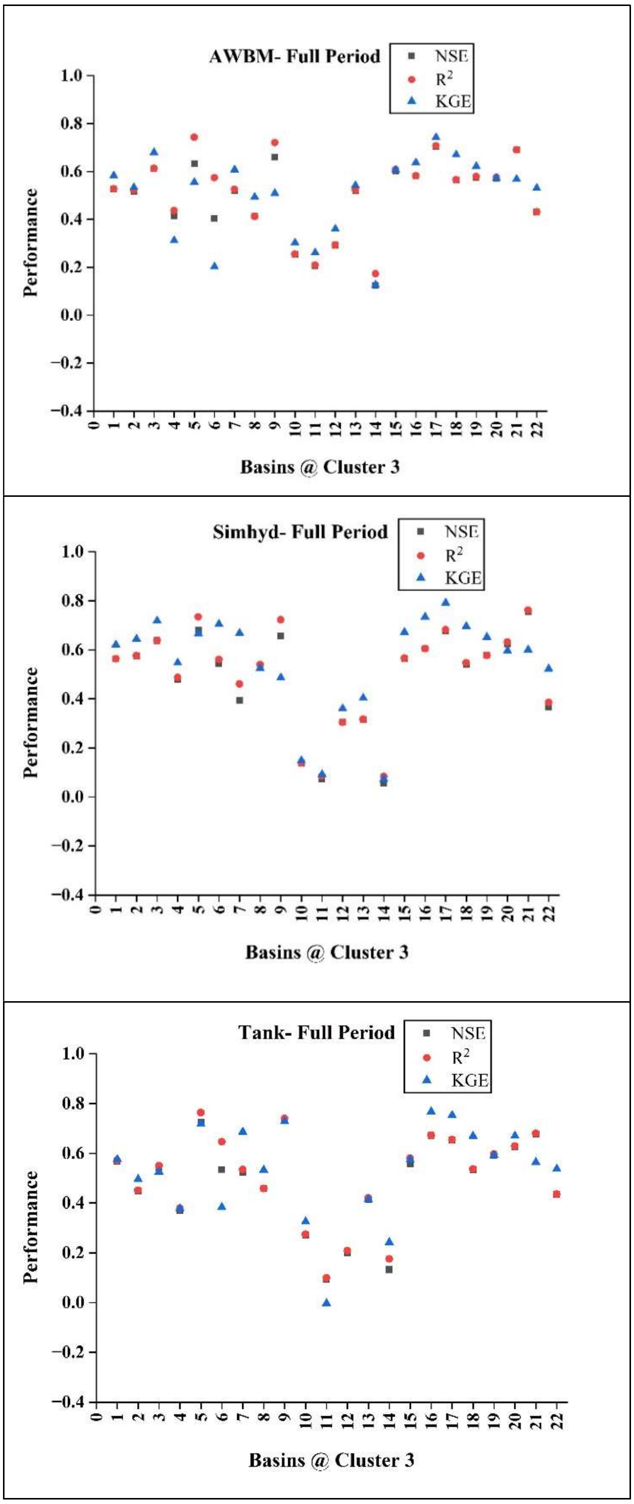


Figure 6.15 Model Performance Results of AWBM/SIMHYD/Tank model for Cluster-3 during the full calibration period

The following inferences can be drawn from the conceptual hydrological modeling using AWBM, SIMHYD, and Tank models:

- The performance statistics of the rainfall-runoff modeling reveal that simulations conducted during the entire calibration period exhibit much better performance compared to those conducted during the split-sampling period. Therefore, the full-period calibration simulations can be used confidently for the prediction of river flows.
- Results indicated that the AWBM model exhibited satisfactory performance (average KGE of 0.50) in 88% of the basins (Figure 6.16), surpassing the SIMHYD (average KGE of 0.53) and Tank (average KGE of 0.35) models, which showed satisfactory performance in 80% of the basins during full period calibration. Figure 6.16 shows the variation in performance of AWBM model in terms of KGE metrics during the full calibration period at different catchments of the study area. This clearly highlights the potential of the AWBM model in the prediction of runoff in ungauged basins.
- Non-satisfactorily performance in a few of the basins especially in Clusters 1 and 3 indicates the possibilities of uncertainties attributed to low flow or zero flow values. This might also be due to the presence of potential outliers in the flow data or it might also be due to erroneous measurement/monitoring of the information. Consequently, it becomes imperative to conduct a thorough review of information from these catchments to identify and rectify any errors and other discrepancies for applying appropriate corrections.

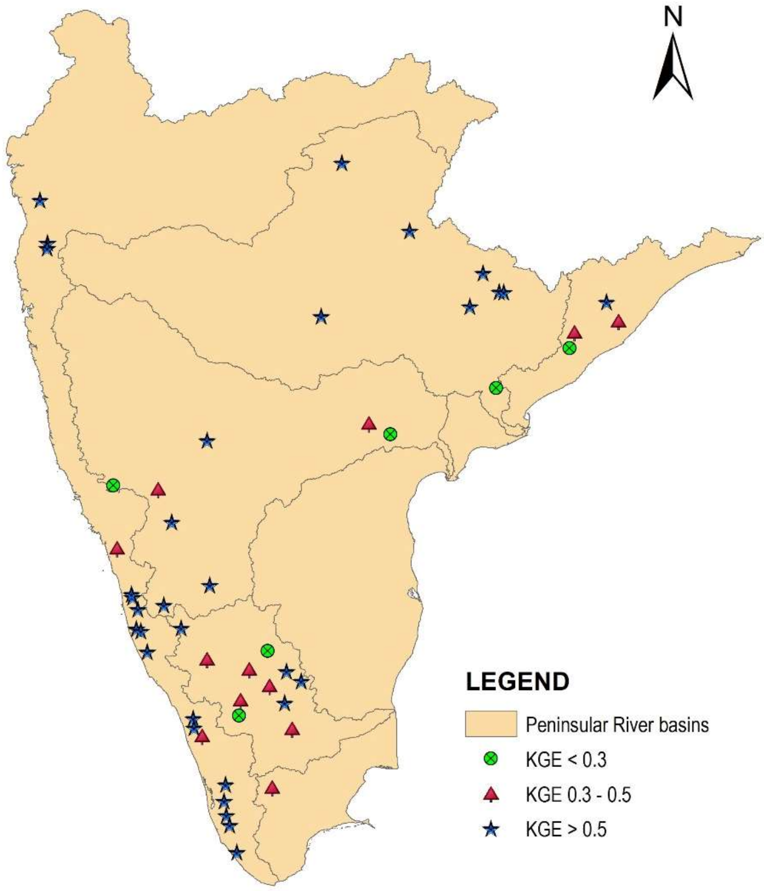


Figure 6.16 Map showing the performance efficiency of AWBM model in terms of KGE metrics in different study catchments

6.5.3 Dominant Model Parameters

The conceptual hydrological modeling of 50 unregulated basins (section 6.5) concluded that the AWBM model outperformed the SIMHYD and Tank model in the prediction of flows. Further, it also revealed that simulations conducted during the full calibration period showcased better performance metrics compared to split sampling trials. Hence the best-forming AWBM model simulation for a full calibration period was considered for analysis of dominant hydrological parameters.

AWBM model has eight model parameters that are calibrated to obtain the desired results. The cluster-wise variation of these model parameters obtained for full calibration period simulation is graphically depicted in Figure 6.17. Across all three clusters, minimal variations were seen in the proportions of surface storage A1, A2, and surface storage capacity C1 value. This implies a consistent behaviour of these parameters across the clusters during simulation.

Surface storage capacity C2 exhibited moderate variability, while a substantial variation was evident in the values of Base flow index (BFI), Base flow recession constant (KBase), Surface Flow Recession Constant (KSurf), and surface storage capacity C3 parameter across all clusters. This significant variability highlights the sensitivity of these parameters to the diverse hydrological characteristics in each cluster. This detailed understanding is crucial for accurately representing and predicting water balance studies in diverse hydrological scenarios.

Variabilities of the parameters of the best performing AWBM model across the catchments in each of the homogeneous clusters were examined and cluster-wise median values were extracted. It was assumed that these could be considered the optimal parameter values for application of the AWBM model in ungauged catchments located in a given cluster and thereby circumvent the need for model calibration using gauged flow records. Therefore, the median value of the model parameters derived from the AWBM model as given in Table 6.4, can be used as regional model parameters.

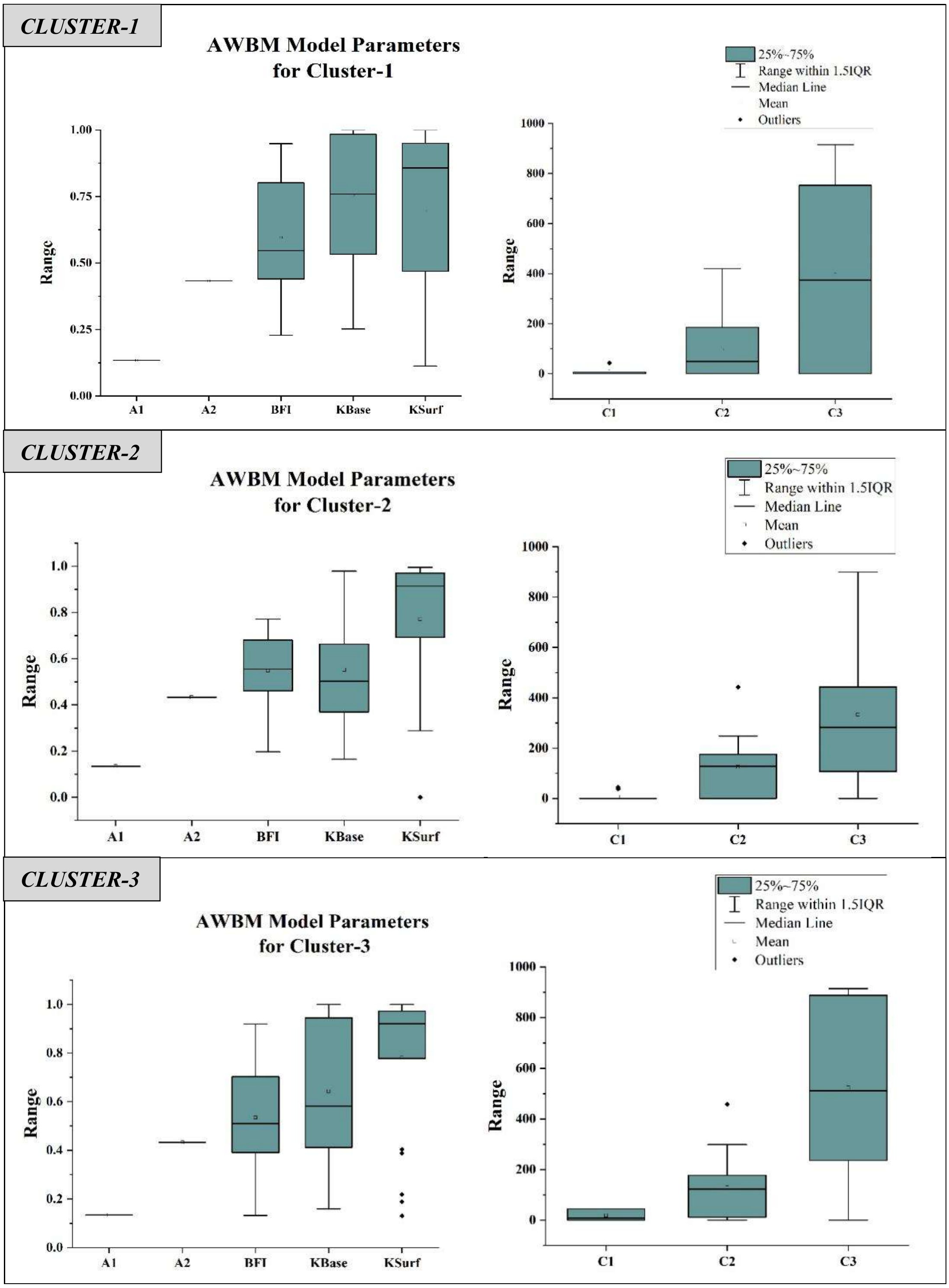


Figure 6.17 Variation of model parameters related to AWBM for Cluster-1, 2, and 3 during full calibration period simulation.

Table 6.4: Median value of Model Parameters of AWBM Model for Cluster -1, 2 and 3

Model Parameters		A1	A2	BFI	C1	C2	C3	KBase	KSurf
Median Value	Cluster-1	0.13	0.43	0.55	0.00	49.72	374.19	0.76	0.86
	Cluster-2	0.13	0.43	0.55	0.00	128.55	282.62	0.50	0.91
	Cluster-3	0.13	0.43	0.51	8.58	123.01	510.85	0.58	0.92

The efficacy of this assumption was tested by applying the AWBM with the median parameter values in 3 gauged catchments located one each in the clusters. Random selection ensures a representative sample, and the simulations help assess how well the model predictions match observed or expected outcomes in the chosen basins. Kuniyil basin in cluster-1, Adoor basin from cluster-2, and Wairagarh basin from Cluster-3 were randomly selected for the validation process.

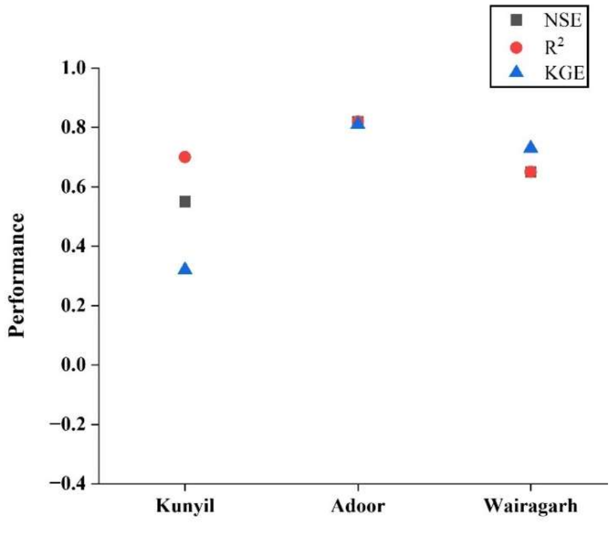


Figure 6.18: Performance of the regional model parameters from the AWBM model

The simulation was carried out to ensure a similar model setup that was adopted during conceptual hydrological modeling and the performance of the regional model parameters was evaluated in terms of R^2, NSE, and KGE metrics as shown in Figure 6.18. The performance of the model (without calibration) was evaluated and found to perform reasonably well in all 3 test catchments (Kuniyil, Adoor and Wairagarh) with the NSE values of 0.55, 0.82, and 0.65 being obtained. This suggests that employing the AWBM model, with its optimal model parameters obtained through regionalization can produce reasonably precise daily streamflow time-series in ungauged catchments within the study area.

The findings from this study reveal that employing hierarchical cluster analysis, hydrological regionalization, and conceptual hydrological modeling in conjunction with predominantly unregulated historical flow records across numerous catchments and diverse catchment attributes, can lead to the creation of accurate and reasonably reliable models for predicting flows in ungauged catchments.

CHAPTER SUMMARY

- *Among all basins, the AWBM model demonstrated acceptable to very good performance in 44 basins, surpassing the SIMHYD and Tank models, which exhibited satisfactory performance in 40 basins during full calibration simulations. The occurrence of unsatisfactory performance in specific basins, particularly in Clusters 1 and 3, suggests potential uncertainties associated with low flow or zero flow values.*
- *Simulations conducted during the entire calibration period exhibit much better performance and can be used confidently for the prediction of river flows.*
- *The adoption of dominant model parameters as regional model parameters in the conceptual hydrological modeling provided satisfactorily results in randomly selected catchments. Therefore, the derived cluster-wise regional model parameters can be used with some degree of confidence for predictions of flows in ungauged basins*

CHAPTER 7

CONCLUSIONS

7.1 SUMMARY AND CONCLUSIONS

The present research was taken up with the specific objective of developing a comprehensive Hydrologic Regionalization-based methodology for deriving the flow duration curve (FDC) and streamflow time-series in ungauged basins located in the peninsular region of South India. For this purpose, historical flow records of 50 largely unregulated catchments located in the region were used. The following conclusions may be drawn from the results obtained in this study:

7.1.1 Delineation of Unregulated Basins and Cluster Analysis

- A total of 50 basins were identified as unregulated after reviewing the CWC basin reports and delineated using 30m resolution SRTM DEM data. Also, a database of 15 catchment attributes was extracted from the DEMs of catchments.
- To enable effective regionalization, a hierarchical agglomerative cluster analysis was implemented using Ward's linkage method, and the study area was delineated into three homogeneous clusters. Cluster – 1 had 17 catchments, followed by 11 catchments in Cluster – 2 and 22 catchments in Cluster – 3. All three clusters were found to be homogenous without any discordant stations as per the CV test and L-Discordancy measure using the L-Moment ratio indicating the efficacy of the hierarchical cluster analysis for regionalization.

7.1.2 Implementation of Cluster-wise Hydrological Regionalization

- Period-of-record flow duration curves for each of the catchments were developed and a total of nine flow quantiles (Q_{10} to Q_{90}) were extracted by interpolation for each catchment.

- As the next step in regionalization, multiple linear regression (MLR) models relating each flow quantile (response variable) to the catchment attributes (predictor variables) were developed. A step-wise regression technique was used to arrive at the final forms of the MLR models containing only the most significant predictor variables. Initially, MLR models were developed considering all 50 catchments to be within a single region and subsequently for each of the three clusters by considering catchments located within them. Performances of the developed MLR models were evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and percentage bias (PBIAS) statistics. Models developed for the clusters performed quite well with average R^2 values for nine flow quantiles being 0.85 for Cluster – 1, 0.97 for Cluster – 2, and 0.80 for Cluster – 3.

- In contrast, considering all 50 catchments a single group was unsatisfactory in predicting the flow quantiles with an average R^2 of 0.23. These results demonstrate the critical need to delineate catchments into homogeneous groups and hierarchical cluster analysis proved to be an efficient technique for doing this.

- A Jackknife cross-validation technique which was adopted to check the reliability of the MLR models, revealed a mixed response for different flow quantiles. Very good to satisfactory performance was recorded for high flow quantiles but was found to be unsatisfactory for low flow quantiles in all three clusters.

7.1.3 Conceptual Rainfall-runoff modeling

- Conceptual rainfall-runoff modeling was carried out on all 50 delineated basins using the RRL toolkit (AWBM, SIMHYD, and Tank models). Apart from discharge data, rainfall, and reference evapotranspiration are two important climate data that are used as input in the RRL model. The performance of each of the models was evaluated in terms of Nash Sutcliffe Efficiency (NSE), correlation coefficient (R^2), and Kling-Gupta's Efficiency Metric.
- Out of all basins, the AWBM model indicated acceptable to very good performance in 44 basins, outperforming the simhyd and tank models, which indicated satisfactory performance in 40 basins. Non-satisfactorily performance in certain basins especially in clusters 1 and 3 indicates the possibilities of uncertainties attributed to low flow or zero flow values. It might also be due to the presence of potential outliers in the flow data or due to erroneous measurement/monitoring of the information.
- Comparative analysis of conceptual rainfall-runoff modeling using split sampling and full-period calibration reveals that simulations conducted during the entire calibration period exhibit much better performance and can be used confidently for the prediction of river flows.

7.1.4 Dominant Model Parameters

- Analysis of the best forming AWBM model simulation for a full calibration period indicated that the model parameters including A1, A2, and C1 showed minimal variations across all clusters implying a consistent behaviour of these parameters. Whereas the parameters including C2, BFI, KBase, and KSurf were sensitive as they exhibited moderate to substantial variation in parameter values.
- The derived median value of the dominant model parameters was adopted as regional model parameter values for application of the AWBM model in ungauged catchments located in a given cluster.

- The model's performance (without calibration) using regional model parameters, was evaluated and demonstrated satisfactory results across all test catchments. This suggest that employing the AWBM model with its optimal model parameters can generate reasonable accurate daily streamflow data in ungauged catchments within the study area.

Overall results of this study demonstrate that the use of hierarchical cluster analysis, regionalization, and conceptual hydrological modeling along with largely unregulated historical flow records for a large number of catchments and a variety of catchment attributes can result in the development of models for the prediction of flow duration curves in ungauged catchments which are very accurate and reasonably reliable. This study is an attempt to solve the issues related to the prediction of flows in poorly gauged or ungauged basins and comprehensive research work is needed to reinforce the outcomes of the present study.

7.2 LIMITATIONS OF THE STUDY

- Deficiency in flow records. More than 150 unregulated basins were identified during this study. However, only 50 were selected owing to the non-availability and/or inconsistency of the flow data.
- The study utilizes gridded climate data (rainfall and temperature) instead of the actual ground-based climate data. Compared to ground-based measurement, the gridded climate data have coarser spatial resolution, suffer from temporal inconsistencies, may introduce uncertainties and biases, and may not capture the local variations.
- The research utilizes lumped hydrological model that offer simplicity and computational efficiency. However, distributed models are also available to analysis spatial variability of hydrological processes, accounting for variations in land cover, soil properties, and topography within a catchment.

- The adopted lumped conceptual models rely solely on rainfall and reference evapotranspiration as driving variables that may lead to inadequate representations of hydrological processes, especially in heterogeneous landscapes.

7.3 FUTURE SCOPE FOR RESEARCH

- The uncertainties attributed to low flow or zero flow values or due to erroneous measurement/monitoring of the information can be cross-verified with concerned department officials for new stations.
- Regionalization process and conceptual hydrological modeling can also be attempted for different time periods namely daily, monthly, and seasonal flow analysis based on the research requirements. However, they suffer from problems like masking the behaviour of the flow regime (peak and low flows) and there might be uncertainties in prediction using Global Hydrological Models (Chouaib et al. 2019).
- Considering the level of complexity and data availability, other watershed attributes like soil, vegetation, geological factors, and other climatic information can also be used in the regionalization process.
- Regionalization using other available techniques can be attempted to make a comparative analysis of the study.
- Hydrological modeling can be tried with other available models including the distributed models to differentiate the findings obtained from the current study.

REFERENCES

Abdolhay, A., Saghafian, B., Soom, M.A. and Ghazali, A.H. (2012). "Identification of homogenous regions in Gorganrood Basin (Iran) for the purpose of regionalization". *Natural Hazards*, 61 (3), 1427–1442.

Abebe, N. A., Ogden, F. L., and Pradhan, N. R. (2010). "Sensitivity and uncertainty analysis of the conceptual HBV rainfall–runoff model: Implications for parameter estimation". *J. Hydrology*, *389*(3-4), 301-310.

Ahmad, S., and Simonovic, S. P. (2005). "An artificial neural network model for generating hydrograph from hydro-meteorological parameters". *J. of Hydrology*, *315*(1-4), 236-251.

Allen, R.G., Pereira, L.S., Raes, D., and Smith, M. (1998). "FAO Irrigation and drainage paper No. 56". Rome: *Food and Agriculture Organization of the United Nations*, *56*(97), e156.

Anshuman, A., Kunnath-Poovakka, A., and Eldho, T.I. (2019). "Towards the use of conceptual models for water resource assessment in Indian tropical watersheds under monsoon-driven climatic conditions". *Environ Earth Sci., 78*, 282.

Arsenault, R. and Brissette, F. (2016). "Analysis of continuous streamflow regionalization methods within a virtual setting". *Hydrological Sciences J.*, *61*(15), 2680–2693.

Atkinson, S. E., Woods, R. A., and Sivapalan, M. (2002). "Climate and landscape controls on water balance model complexity over changing timescales". *Water Resources Research*, *38*(12), 1314.

Bao, Z., Zhang, J., Liu, J., Fu, G., Wang, G., He, R., Yan, X., Jin, J.and Liu, H. (2012). "Comparison of regionalization approaches based on regression and similarity for predictions in ungauged catchments under multiple hydro-climatic conditions". *J. of Hydrology, 466–467*, 37–46.

Bastola, S., Ishidaira, H., and Takeuchi, K. (2007). "Regionalization of parameters of hydrological models: inclusion of model parameter uncertainty". *Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resources Management*, IAHS Publication, (313), 319-328.

Basri, H. (2013). "Development of Rainfall-runoff Modeling Using a Tank Model: Problems and Challenges in Province of Aceh, Indonesia". *Aceh International J. of Science and Technology*, *2*(1), 26-37.

Benke, K. K., Lowell, K. E., and Hamilton, A. J. (2008). "Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model". *Mathematical and Computer Modelling*, *47*(11-12), 1134-1149.

Bergstrom, S. (1991). "Principles and confidence in hydrological modelling". *Hydrology Research*, *22*(2), 123-136.

Beven, K. (1989). "Changing ideas in hydrology-the case of physically-based models". *J. of hydrology*, *105*(1-2), 157-172.

Beven, K. (2006). "A manifesto for the equifinality thesis". *J. of hydrology*, *320*(1-2), 18-36.

Bharath, A. L., and Venkatesh, B. (2022). "Long term drought, trend analysis, and homogeneity analysis for the Belagavi district, Karnataka". *J. of Earth System Science*, *131*(4), 238.

Bisht, D. S., Chatterjee, C., Raghuwanshi, N. S., and Sridhar, V. (2018). "Spatiotemporal trends of rainfall across Indian river basins". *Theoretical Applied Climatology.,* *132*(1–2), 419–436.

Blank, D., and Delleur, J. W. (1968). "Program for Estimating Runoff from Indiana Watersheds, Part I: Linear System Analysis in Surface Hydrology and Its Application to Indiana Watersheds", Purdue University, Indiana Water Resources Research Center, *IWRRC Tech. Report No.4.*

Blöschl, G., and Sivapalan, M. (1995). "Scale issues in hydrological modelling: A review". *Hydrological Processes, 9*(3-4), 251-290.

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenije, H. (2013). "Runoff prediction in ungauged basins synthesis across processes, places and scales". *Cambridge Univ. Press.*

Boscarello, L., Ravazzani, G., Cislaghi, A., and Mancini, M. (2016). "Regionalization of flow-duration curves through catchment classification with streamflow signatures and physiographic–climate indices". *J. of Hydrologic Engineering 21*(3), 05015027.

Boughton, W. (2004). "The Australian water balance model". *Environmental Modelling and Software*, *19*(10), 943–956.

Boughton, W., (2006). "Calibrations of a daily rainfall-runoff model with poor quality data Walter Boughton". *Environmental Modelling and Software*. *21*(8), 1114-1128.

Boughton, W., and Chiew, F. (2007). "Estimating runoff in ungauged catchments from rainfall, PET and the AWBM model". *Environmental Modelling and Software*, *22*(4):476–487.

Bracken, L. J., Wainwright, J., Ali, G. A., Tetzlaff, D., Smith, M. W., Reaney, S. M., and Roy, A. G. (2013). "Concepts of hydrological connectivity: research approaches, pathways and future agendas". *Earth-Science Reviews*, *119*, 17-34.

Brath, A., Montanari, A., and Toth, E. (2002). "Neural networks and non-parametric methods for improving real-time flood forecasting through conceptual hydrological models". *Hydrology and Earth System Sciences Discussions*, *6*(4), 627-639.

Braun, L. N., and Renner, C. B. (1992). "Application of a conceptual runoff model in different physiographic regions of Switzerland". *Hydrological Sciences J.*, *37*(3), 217-231.

Brazil, L. E., and Hudlow, M. D. (1981). "Calibration procedures used with the national weather service river forecast system". In: Haimes YY, Kindler J (eds) *Water and related land resource systems. Pergamon Press, New York*, 457–566.

Burgan, H.I., and Aksoy, H. (2020). "Monthly flow duration curve model for ungauged river basins". *Water 12*(2), 338.

Burn, D. H., Zrinji, Z. and Kowalchuk, M. (1997). "Regionalization of catchments for regional flood frequency analysis". *J. of Hydrologic Engineering, 2*(2), 76–82.

Castellarin, A., Galeati, G., Brandimarte, L., Montanari, A., and Brath, A. (2004). "Regional flow-duration curves: Reliability for ungauged basins". *Advances in Water Resources*, *27*(10), 953–965.

Chiang, S., Tsay, T., and Nix, S. J. (2002). "Hydrologic Regionalization of Watersheds. I: Methodology Development". *J. of Water Resources Planning and Management, 128*(1), 3-11.

Chiew, F. H. S., Stewardson, M. J., and McMahon, T. A. (1993). "Comparison of six rainfall-runoff modelling approaches". *J. of Hydrology*, *147*(1-4), 1-36.

Chiew, F., and McMahon, T. (1994). "Application of the daily rainfall-runoff model MODHYDROLOG to 28 Australian catchments". *J. of Hydrology*, *153*(1-4), 383-416.

Chiew, F.H.S., Peel, M.C., and Western, A.W. (2002). "Application and testing of the simple rainfall–runoff model SIMHYD". In: V.P. Singh and D.K. Frevert, eds. Mathematical models of small watershed hydrology and applications. Littleton, *CO: Water Resources*, 335–367.

Chouaib, W., Alila, Y., and Caldwell, P. V. (2019). "On the use of mean monthly runoff to predict the flow–duration curve in ungauged catchments". *Hydrological Sciences J., 64*(13), 1573–1587.

Crawford, N. H. and Linsley, R. K. (1966). "Digital simulation in hydrology: Stanford Watershed Mode IV". *Technical report 39,* Civil Engineering Department, Stanford University.

Croke, B. F. W., Andrews, F., Jakeman, A. J., Cuddy, S. M., and Luddy, A. (2006). "IHACRES classic plus: a redesign of the IHACRES rainfall-runoff model". *Environmental Modelling and Software, 21*(3):426–427.

Daniel, E. B., Camp, J. V., LeBoeuf, E. J., Penrod, J. R., Dobbins, J. P., and Abkowitz, M. D. (2011). "Watershed modeling and its applications: A state-of-the-art review". *The Open Hydrology J.*, *5*(1), 26-50.

Dawdy, D. R. (1983). "A Review of Rainfall-Runoff Modeling. Proceedings of the Tegucigalpa Hydromath Symposium", *IAHS Publ. No. 152.*

Dawdy, D. R., and Lichty, R. W. (1968). "Methodology of hydrologic model building". *Proceedings, use of analog and digital computers in hydrology*, *2*, 347-355.

Dawdy, D. R., Lichty, R. W., and Bergmann, J. M. (1972). "A rainfall-runoff simulation model for estimation of flood peaks for small drainage basins". *US Government Printing Office*.

Demirel, M. C. (2004). "Cluster analysis of streamflow data over Turkey". *MSc thesis*, ITU, Istanbul, Turkey.

Devak, M., and Dhanya, C. T. (2017). "Sensitivity analysis of hydrological models: Review and way forward". *J. of Water and Climate Change*, *8*(4), 557–575.

Devi, G. K., Ganasri, B. P., and Dwarakish, G. S. (2015). "A review on hydrological models". *Aquatic Procedia*, *4*, 1001-1007.

Dhami, B. S., and Pandey, A. (2013). "Comparative review of recently developed hydrologic models". *J. of Indian Water Resources Society*, *33*(3), 34-42.

Dominguez, E., Dawson, C. W., Ramírez, A., and Abrahart, R. J. (2010). "The search for orthogonal hydrological modelling metrics: A case study of 20 monitoring stations in Colombia". *J. of Hydroinformatics, 13*(3), 429-442.

Duan, Q., Sorooshian, S., and Gupta, V. (1992). "Effective and efficient global optimization for conceptual rainfall-runoff models". *Water resources research*, *28*(4), 1015-1031.

Edijatno, De Oliveira Nascimento, N. I. L. O., Yang, X., Makhlouf, Z., and Michel, C. (1999). "GR3J: a daily watershed model with three free parameters". *Hydrological Sciences J.*, *44*(2), 263-277.

Efron, B. (1981). "Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods". *Biometrika, 68*(3), 589-599.

Esmaeili-Gisavandani, H., Lotfirad, M., Sofla, M. S., and Ashrafzadeh, A. (2021). "Improving the performance of rainfall-runoff models using the gene expression programming approach". *J. of Water and Climate Change*, 12(7), 3308–3329.

Farmer, D., Sivapalan, M., and Jothityangkoon, C. (2003). "Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: Downward approach to water balance analysis". *Water Resources Research*, *39*(2), 1035.

Fennessey, N., and Vogel, R. M. (1990). "Regional Flow-Duration Curves for Ungauged Sites in Massachusetts". *J. of Water Resources Planning and Management, 116*(4), 530-549.

Fernandez, W., Vogel, R. M., and Sankarasubramanian, A. (2000). "Regional calibration of a watershed model". *Hydrological sciences J.*, *45*(5), 689-707.

Fernandez, W., Vogel, R. M., and Sankarasubramanian, A. (2000). "Regional calibration of a watershed model". *Hydrological sciences J.*, *45*(5), 689-707.

Franchini, M., and Pacciani, M. (1991). "Comparative analysis of several conceptual rainfall-runoff models". *J. of Hydrology*, *122*(1-4), 161-219.

Frimpong, B. F., Koranteng, A., and Molkenthin, F. (2022). "Analysis of temperature variability utilising mann–kendall and Sen's slope estimator tests in the Accra and Kumasi Metropolises in Ghana". *Environmental Systems Research*, *11*(24).

Gaviria, C., and Carvajal-Serna, F. (2022). "Regionalization of flow duration curves in Colombia". *Hydrology Research*, *53*(8), 1075–1089.

Gilbert, R.O. (1987). "Statistical Methods for Environmental Pollution Monitoring". *John Wiley & Sons, New York.*

Gocic, M., and Trajkovic, S. (2013). "Analysis of changes in meteorological variables using mann-kendall and SEN's slope estimator statistical tests in Serbia". *Global and Planetary Change*, *100*, 172–182.

Goyal, M.K., and Gupta, V. (2014). "Identification of homogeneous rainfall regimes in northeast region of India using fuzzy cluster analysis". *Water Resources Management, 28*(13), 4491–4511.

Guo, Y., Zhang, Y., Zhang, L., and Wang, Z. (2020). "Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review". *WIREs Water, 8*(1), e1487.

Helsel, D. R., and Hirsch, R. M. (2002). "Statistical methods in water resources". *U.S. Geological survey technique of water resources investigations*, Book 4, Chapter A3, P524. < htpp://water.usgs.gov/pubs/twri/ twri4a3/>.

Hosking, J. R. (1990). "L-moments: Analysis and estimation of distributions using linear combinations of order statistics". *J. of the Royal Statistical Society: Series B (Methodological), 52*(1), 105–124.

Hosking, J. R., and Wallis, J. R. (1993) "Some statistics useful in regional frequency analysis". *Water Resources Research, 29*(2), 271–281.

Huang, C., Wang, G., Zheng, X., Yu, J., and Xu, X. (2015). "Simple linear modeling approach for linking hydrological model parameters to the physical features of a river basin". *Water Resources Management, 29*(9), 3265–3289.

Hrachowitz, M., Savenije, H. H G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C. (2013). "A decade of Predictions in Ungauged Basins (PUB)—a review". *Hydrological Sciences J.*, *58*(6), 1198-1255.

Hydrological Data (Unclassified) Book. (2018). "Hydrological Data Directorate Information System Organization". Water Planning and Projects Wing, Central Water Commission, New Delhi, Government of India.

Integrated Hydrological Data (Non-Classified River Basins) Book. (2017). "Hydrological Data Directorate Information System Organisation". Water Planning and Projects Wing, Central Water Commission. New Delhi.

Isik, S., and Singh, V. P. (2008). "Hydrologic Regionalization of Watersheds in Turkey". *J. of Hydrologic Engineering, 13*(9), 824-834.

Islam, A., and Deb Barman, S. (2020). "Drainage basin morphometry and evaluating its role on flood-inducing capacity of tributary basins of Mayurakshi River, India". *SN Applied Sciences*, *2*(6), 1087.

Jaiswal, R.K., Ali, S., and Bharti, B. (2020). “Comparative evaluation of conceptual and physical rainfall–runoff models”. *Applied Water Science, 10*, 48.

James, L. D., and Burges, S. J. (1982). “Selection, calibration, and testing of hydrologic models, Hydrologic Modeling of Small Watersheds”. CT Haan, HP Johnson, DL Brakensiek Eds, Monograph, *ASAE*, Michigan, USA, *5*, 437–472.

Javadinejad, S. (2021). “A review on homogeneity across hydrological regions”. *Resources Environment and Information Engineering*, *3*(1), 124-137.

Jin, X., Xu, C., Zhang, Q., and Chen, Y. D. (2009). “Regionalization study of a conceptual hydrological model in Dongjiang basin, south China”. *Quaternary International, 208*(1-2), 129-137.

Jothityangkoon, C., Sivapalan, M., and Farmer, D. L. (2001). “Process Controls of Water Balance Variability in A Large Semi-Arid Catchment: Downward Approach to Hydrological Model Development”. *J. of Hydrology*, *254*(1-4), 174-198.

Karki, N., Shakya, N. M., Pandey, V. P., Devkota, L. P., Pradhan, A. M., and Lamichhane, S. (2023). “Comparative performance of regionalization methods for model parameterization in ungauged Himalayan watersheds”. *J. of Hydrology: Regional Studies 47*, 101359.

Kendall, M. G. (1975). “Rank correlation methods”. *London: Charles GriDn & Co*.

Khalid, K., Ali, M. F., Rahman, N. F. A., Mispan, M. R., Haron, S. H., Othman, Z., and Bachok, M. F. (2016). “Sensitivity analysis in watershed model using SUFI-2 algorithm”. *Procedia engineering*, *162*, 441-447.

Klemes, V. (1983). “Conceptualization and scale in hydrology”. *J. Hydrol., 65*, 1–23.

Krishna, M., Prasad, M., Rao, D., Viswanadham, R., Sarma, V., and Reddy, N. (2016). “Export of dissolved inorganic nutrients to the northern Indian Ocean from the Indian monsoonal rivers during discharge period”. *Geochimica Et Cosmochimica Acta., 172*, 430-443.

Kumar Raju, B. C., and Nandagiri, L. (2017). “Analysis of historical trends in hydrometeorological variables in the upper Cauvery Basin, Karnataka, India”. *Current Science*, *112*(03), 577.

Kunnath-Poovakka, A., and Eldho, T.I. (2019). "A comparative study of conceptual rainfall-runoff models GR4J, AWBM and Sacramento at catchments in the upper Godavari river basin, India". *J Earth System Science, 128*, 33.

Latt, Z.Z., Wittenberg, H., and Urban, B. (2014). "Clustering hydrological homogeneous regions and neural network based index flood estimation for ungauged catchments: An example of the Chindwin River in Myanmar". *Water Resources Management, 29*(3), 913–928.

Li, M., Shao, Q., Zhang, L., and Chiew, F. H. (2010). "A new regionalization approach and its application to predict flow duration curve in ungauged basins". *J. of Hydrology, 389*(1-2), 137-145.

Loucks, D. P., Van Beek, E., Stedinger, J. R., Dijkman, J. P. M., and Villars, M. (2005). "Model sensitivity and uncertainty analysis". *Water resources systems planning and management*, 255-290.

Ma, L., Liu, D., Huang, Q., Guo, F., Zheng, X., Zhao, J., Luan, J., Fan, J., and Ming, G. (2023). "Identification of a function to fit the flow duration curve and parameterization of a semi-arid region in North China". *Atmosphere 14*(1), 116.

Mann, H. B. (1945). "Nonparametric tests against trend". *Econometrica*,*13*(3), 245–259.

Mathevet, T., Michel, C., Andreassian, V., and Perrin, C. (2006). "A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins". *IAHS Publication*, *307*, 211.

Masih, I., Uhlenbrook, S., Maskey, S., and Ahmad, M. (2010). "Regionalization of a conceptual rainfall–runoff model based on similarity of the flow duration curve: A case study from the semi-arid Karkheh basin, Iran". *J. of Hydrology, 391*(1-2), 188-201.

Mccuen, R. H. (2005). "Accuracy Assessment of Peak Discharge Models". *J. of Hydrologic Engineering, 10*(1), 16-22.

McCuen, R. H. (1973). "The role of sensitivity analysis in hydrologic modelling". *J. of Hydrology*, *18*(1), 37-53.

Medina, Y., and Muñoz, E. (2020). "Analysis of the Relative Importance of Model Parameters in Watersheds with Different Hydrological Regimes". *Water 12*(9): 2376.

Merz, R., and Blöschl, G. (2004). "Regionalisation of catchment model parameters". *J. of hydrology*, *287*(1-4), 95-123.

Meher, J. K., and Das, L. (2019). "Gridded data as a source of missing data replacement in Station Records". *J. of Earth System Science*, *128, 58*.

Mitra, A. K., Momin, I. M., Rajagopal, E. N., Basu, S., Rajeevan, M. N., and Krishnamurti, T. N. (2013). "Gridded Daily Indian monsoon rainfall for 14 seasons: Merged TRMM and IMD gauge analyzed values". *J. of Earth System Science*, *122*(5), 1173–1182.

Mohamoud, Y.M. (2008). "Prediction of daily flow duration curves and streamflow for ungauged catchments using regional flow duration curves". *Hydrological Sciences J. 53*(4), 706–724.

Montanari, L., Sivapalan, M., and Montanari, A. (2006). "Investigation of dominant hydrological processes in a tropical catchment in a monsoonal climate via the downward approach". *Hydrology and Earth System Sciences Discussions*, *10*(5), 769-782.

Moradkhani, H., and Sorooshian, S. (2009). "General review of rainfall-runoff modeling: model calibration, data assimilation, and uncertainty analysis". In S. Sorooshian et al. (eds.), *Hydrological modelling and the water cycle*. Springer, Berlin, *1-24*.

Mosley, M.P. (1981). "Delimitation of New Zealand hydrologic regions". *J. of Hydrology, 49*(1–2), 173–192.

Moriasi, D. N., Gitau, M. W., Pai, N., and Daggupati, P. (2015). "Hydrologic and water quality models: Performance measures and evaluation criteria". *Transactions of the ASABE*, *58*(6), 1763–1785.

Muleta, M. K. (2011). "Improving model performance using season-based evaluation". *J. of Hydrologic Engineering*, *17*(1), 191-200.

Müftüoğlu, R. (1984). "New models for nonlinear catchment analysis". *J. of Hydrology, 73*(3-4), 335-357.

Mulaomerovic-Seta, A., Blagojevic, B., Mihailovic, V., and Petroselli, A. (2023). "A silhouette-width-induced hierarchical clustering for defining flood estimation regions". *Hydrology, 10*(6), 126.

Nathan, R. J., and McMahon, T. A. (1990). "Identification of homogeneous regions for the purposes of regionalization". J. of Hydrology, *121*(1–4), 217–238.

Nandagiri, Lakshman. (2010). "Regionalization of Flow Duration Quantiles – Comparison of Multiple Least-Squares Regression and Artificial Neural Networks Approaches". *Proceedings: National Conference on 'Sustainable Water Resources Management and Impact of Climate Change*, BITS-Hyderabad, India, March 2010.

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

Nobert, J., Ndayizeye, J., and Mkhandi, S. (2011). "Regional Flow Duration Curve Estimation and its Application in Assessing Low Flow Characteristics for Ungauged Catchment. A Case Study of Rwegura Catchment-Burundi". *Nile Basin Water Science and Engineering J.*, *4*(1),13-23.

Niranjan, S., and Nandagiri, L. (2021). "Effect of local calibration on the performance of the Hargreaves reference crop evapotranspiration equation". *J. of Water and Climate Change, 12*(6): 2654–2673.

Nruthya, K., and Srinivas, V. (2015). "Evaluating Methods to Predict Streamflow at Ungauged Sites Using Regional Flow Duration Curves: A Case Study". *Aquatic Procedia, 4*, 641-648.

Pai, D. S., Latha Sridhar., Rajeevan M., Sreejith, O. P., Satbhai, N. S., and Mukhopadhyay, B. (2014). "Development of new high resolution (0.25^0x0.25^0) long period (1901-2010) gridded rainfall data set over India and its comparison with existing data sets over the region". *Mausam*, *65*(1), 1-18.

Panda, D. K., Kumar, A., and Mohanty, S. (2011). "Recent trends in sediment load of the tropical (Peninsular) river basins of India". *Global and Planetary Change, 75*(3-4), 108-118.

Panthi, J., Talchabhadel, R., Ghimire, G. R., Sharma, S., Dahal, P., Baniya, R., Boving, T., Pradhanang, S. M., and Parajuli, B. (2021). "Hydrologic regionalization under data scarcity: Implications for streamflow prediction". *J. of Hydrologic Engineering, 26*(9), 05021022.

Paul, P. K., Zhang, Y., Ma, N., Mishra, A., Panigrahy, N., and Singh, R. (2021). "Selecting Hydrological Models for Developing Countries: Perspective of Global, Continental, and Country Scale Models over Catchment Scale Models". *J. of Hydrology*. 600. 126561.

Parajka, J., Merz, R. and Blöschl, G. (2005). "A comparison of regionalisation methods for catchment model parameters". *Hydrology and Earth System Sciences, 9*(3), 157–171.

Pechlivanidis, I. G., Jackson, B. M., McIntyre, N. R., and Wheater, H. S. (2011). "Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications". *Global NEST J.*, *13*(3), 193-214.

Perrin, C., Michel, C., and Andréassian, V. (2001). "Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments". *J. of Hydrology*, *242*(3-4), 275-301.

Petrakis, R. E., Norman, L. M., Vaughn, K., Pritzlaff, R., Weaver, C., Rader, A., and Pulliam, H. R. (2021). "Hierarchical clustering for paired watershed experiments: Case study in south eastern Arizona, U.S.A". *Water13*(21), 2955.

Podger, G., (2004). "User Guide for RRL Library". *CRC for Catchment Hydrology, eWaters, Australia.*

Post, D. A., and Jakeman, A. J. (1999). "Predicting the daily streamflow of ungauged catchments in SE Australia by regionalising the parameters of a lumped conceptual rainfall-runoff model". *Ecological Modelling*, *123*(2-3), 91-104.

Qamar, M.U., Ganora, D., Claps, P., Azmat, M., Shahid, M.A., and Khushnood, R.A. (2016). "Flow duration curve regionalization with enhanced selection of donor basins". *J. of Applied Water Engineering and Research, 6*(1), 70–84.

Quimpo, R.G., Alejandrino, A.A. and McNally, T.A. (1983). "Regionalized flow duration for Philippines". *J. of Water Resources Planning and Management, 109*(4), 320–330.

Rao, A. R., and Srinivas, V.V. (2006). "Regionalization by hybrid cluster analysis". *J. of Hydrology, 318*(1), 37–56.

Refsgaard, J. C., and Knudsen, J. (1996). "Operational validation and intercomparison of different types of hydrological models". *Water Resources Research*, *32*(7), 2189-2202.

Riswandi, H., Sukiyah, E. and Tania, D. (2022). "Hydrogeological cluster analysis with average linkage and ward method in the southern slope of Merapi Mountain". In: Proceedings of the 5th International Conference of Geological Engineering Faculty, Universitas Padjadjaran, Indonesia. *TIIKM Publishing. 1*(2), 47–61.

Robinson, S. (2006). "Conceptual modeling for simulation: issues and research requirements". In *Simulation Conference, 2006. WSC 06. Proceedings of the Winter* (pp. 792-800). IEEE.

Seibert, J. (1999). "Regionalisation of parameters for a conceptual rainfall-runoff model". *Agricultural and forest meteorology*, *98*, 279-293.

Sen, P.K. (1968). "Estimates of the regression coefficient based on Kendall's tau". *J. of the American Statistical Association, 63*, 1379–1389.

Setti, S., Maheswaran, R., Sridhar, V., Barik, K. K., Merz, B., and Agarwal, A. (2020). "Inter-comparison of gauge-based gridded data, reanalysis and satellite precipitation product with an emphasis on hydrological modelling". *Atmosphere*, *11*(11), 1252.

Shaban, M., Urban, B., El Saadi, A., and Faisal, M. (2010). "Detection and mapping of water pollution variation in the Nile delta using multivariate clustering and GIS techniques". *J. of Environmental Management*, *91*(8), 1785–1793.

Shao, J., and Tu, D. (1995). "The Jackknife and Bootstrap. Springer Series in Statistics". *Springer Inc.*, New York.

Sharma, N., Zakaullah, M., Tiwari, H., and Kumar, D. (2015). "Runoff and sediment yield modeling using ANN and support vector machines: A case study from Nepal watershed". *Modeling Earth Systems and Environment, 1*(3).

Shen, H., Tolson, B. A., and Mai, J. (2022). "Time to update the split-sample approach in hydrological model calibration". *Water Resources Research*, *58*, e2021WR031523.

Shu, C., and Ouarda, T. B. (2012). "Improved methods for daily streamflow estimates at ungauged sites". *Water Resources Research*, *48*(2), W02523.

Silva, R., Blanco, C. J., and Pessoa, F. C. (2019). "Alternative for the regionalization of flow duration curves". *J. of Applied Water Engineering and Research*, *7*(3), 198–206.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E. (2003a). "IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences". *Hydrological sciences J.*, *48*(6), 857-880.

Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R. (2003b). "Downward approach to hydrological prediction". *Hydrological processes*, *17*(11), 2101-2111.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., and Xu, C. (2015). "Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications". *J. of hydrology*, *523*, 739-757.

Song, Z., Xia, J., Wang, G., She, D., Hu, C. and Hong, S. (2022). "Regionalization of hydrological model parameters using gradient boosting machine". *Hydrology and Earth System Sciences*, *26*(2), 505–524.

Sorooshian, S., and Gupta, V. K. (1985). "The Analysis of Structural Identifiability: Theory and Application to Conceptual Rainfall-Runoff Models". *Water Resources Research*, *21*(4), 487-495.

Srivastava, K., Rajeevan, M., and Kshirsagar S. R. (2009). "Development of high resolution daily gridded temperature data set (1969-2005) for the Indian region". *Atmospheric Science Letters. 10*(4), 249-254.

Subramanya, K. (2008). "Engineering Hydrology". 3rd Edition, McGraw Hill Co. Ltd. New Delhi.

Sugawara, M. I., Watanabe, I., Ozaki, E., and Katsuyame, Y. (1983). "Reference manual for the TANK model". *National Research Center for Disaster Preview, Tokyo.*

Sugiyama, H., Vudhivanich, V., Whitaker, A. C., and Lorsirirat, K. (2003). "Stochastic Flow Duration Curves for Evaluation of Flow Regimes in Rivers". *J. of the American Water Resources Association, 39*(1), 47-58.

Sukristiyanti, S., Maria, R., and Lestiana, H. (2018). "Watershed-based Morphometric Analysis: A Review". *IOP Conference Series: Earth and Environmental Science*, *118*, 012028.

Suryoputro, N., Suhardjono, Soetopo, W., and Suhartanto, E. (2017). "Calibration of infiltration parameters on hydrological tank model using runoff coefficient of rational method". In *AIP Conference Proceedings*, *AIP Publishing*, *1887*(1), 020056.

Swain, J. B., and Patra, K. C. (2017). "Streamflow estimation in ungauged catchments using regional flow duration curve: Comparative study". *J. of Hydrologic Engineering, 22*(7), 04017010.

Tang, Y., Reed, P., Wagener, T., and van Werkhoven, K. (2007). "Comparing sensitivity analysis methods to advance lumped Watershed Model Identification and Evaluation". *Hydrology and Earth System Sciences*, *11*(2), 793–817.

Tasker, G.D. (1982). "Comparing methods of hydrologic regionalization". *J. of the American Water Resources Association*, *18*(6), 965–970.

Tayyab, M., Zhou, J., Zeng, X., Chen, L., and Ye, L. (2015). "Optimal application of conceptual rainfall–runoff hydrological models in the Jinshajiang River basin, China". *Proceedings of the International Association of Hydrological Sciences*, *368*, 227-232.

Todini, E. (1996). "The ARNO rainfall–runoff model". *J. of Hydrology.*, *175*:339–382.

Uhlenbrook, S., Seibert, J. A. N., Leibundgut, C., and Rodhe, A. (1999). "Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure". *Hydrological Sciences J.*, *44*(5), 779-797.

Vallam, P., Qin, X. S., and Yu, J. J. (2014). "Uncertainty quantification of hydrologic model". *APCBEE procedia*, *10*, 219-223.

Vaze, J., Chiew, F.H., Perraud, J. M., Viney, N., Post, D., Teng, J., Wang, B., Lerat, J., and Goswami, M. (2011). "Rainfall-runoff modelling across southeast Australia: datasets, models and results". *Australas J Water Resource, 14*(2):101–116.

Venkatesan, A. (2014). "Geoinformatics in fluvial geomorphological study of thoppaiyar sub basin Tamil Nadu India", *Doctoral Dissertation Report*, Periyar University, Tamil Nadu.

Vogel, R. M., and Fennessey, N. M. (1994). "Flow-Duration Curves I: New Interpretation and Confidence Intervals". *J. of Water Resources Planning and Management, 120*(4), 485-504.

Vogel, R. M., and Fennessey, N. M. (1995). "Flow-Duration Curves II: A review of Applications in Water Resources Planning". *J. of the American Water Resources Association, 31*(6), 1029-1039.

Vogel, R. M., Wilson, I., and Daly, C. (1999). "Regional Regression Models of Annual Streamflow for the United States". *J. of Irrigation and Drainage Engineering, 125*(3), 148-157.

Vogel, R., Chapra, S., and Limbrunner, J. (2005). "A Parsimonious Watershed Model". *Watershed Models,* 549-567.

Wagener, T., and Wheater, H. S. (2006). "Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty". *J. of Hydrology, 320*(1-2), 132-154.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., and Gupta, H. V. (2004). "Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis". *Hydrological Processes*, *18*(13), 2699-2722.

Wood, E. F., Sivapalan, M., Beven, K., and Band, L. (1988). "Effects of spatial variability and scale with implications to hydrologic modeling". *J. of Hydrology*, *102*(1–4), 29–47.

Xu, C.Y., and Vandewiele, G.L. (1995). "Parsimonious monthly rainfall runoff models for humid basins with different input requirements". *Advances in Water Resources*,*18*, 39-48.

Xu, C. Y., and Singh, V. P. (1998). "A review on monthly water balance models for water resources investigations". *Water Resources Management*, *12*(1), 20-50.

Xu, C. Y. (1999). "Estimation of parameters of a conceptual water balance model for ungauged catchments". *Water Resources Management*, *13*(5), 353-368.

Xu, C. Y. (2002). "Textbooks of Hydrologic models". Uppsala University, Department of Earth Sciences Hydrology Sweden.

Yang, X., Magnusson, J., Rizzi, J., and Xu, C.Y. (2017). "Runoff prediction in ungauged catchments in Norway: Comparison of regionalization approaches". *Hydrology Research, 49*(2), 487–505.

Yang, J., Reichert, P., Abbaspour, K. C., Xia, J., and Yang, H. (2008). "Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China". *J. of Hydrology*, *358*(1-2), 1-23.

Yang, S., Gao, M., Liu, J., Wu, P., and Yang, Y. (2023). "Regionalizing streamflow regime function through integrations of geographical controls in mountainous basins". *Water, 15*(2), 280.

Yates, D., and Strzepek, K.M. (1994). "Potential Evapotranspiration Methods and their Impact on the Assessment of River Basin Runoff Under Climate Change". *IIASA Working Paper. IIASA, Laxenburg, Austria*: WP-94-046.

Yu, B. (2015). "How would peak rainfall intensity affect runoff predictions using conceptual water balance models?". *Proceedings of the International Association of Hydrological Sciences*, *371*, 109–115.

Yu, B., and Zhu, Z., (2015). "A comparative assessment of AWBM and SimHyd for forested watersheds". *Hydrological Sciences J.*, *60*(7-8), 1200-1212.

Yu, P., Yang, T., and Wang, Y. (2002). "Uncertainty Analysis of Regional Flow Duration Curves". *J. of Water Resources Planning and Management, 128*(6), 424-430.

Yu, P. S., and Yang, T.-C. (1996). "Synthetic regional flow duration curve for southern Taiwan". Hydrol. Process., *10*: 373-391.

Zamani, R., Mirabbasi, R., Nazeri, M., Meshram, S. G., and Ahmadi, F. (2018). “Spatio-temporal analysis of daily, seasonal and annual precipitation concentration in Jharkhand State, India”. *Stoch. Environ. Res. Risk Assess., 32*(4):1085–1097.

Zhang, Z., Lu, W., Chu, H., Cheng, W., and Zhao, Y. (2014). “Uncertainty analysis of hydrological model parameters based on the bootstrap method: A case study of the SWAT model applied to the Dongliao River Watershed, Jilin Province, Northeastern China”. *Science China Technological Sciences*, *57*(1), 219-229.

LIST OF PUBLICATIONS

INTERNATIONAL JOURNAL

- Hiremath C. G., and Nandagiri L. (2023). "Regionalization of Flow Duration Curves for Catchments in Southern India using a Hierarchical Cluster Approach". *Journal of Water and Climate Change*, *14*(12), 4875-4898.

  DOI permalink: https://doi.org/10.2166/wcc.2023.467

SCOPUS INDEXED BOOK CHAPTER

- Hiremath C. G., and Nandagiri L. (2020). "Regionalization of Flow Duration Curve for West Flowing Rivers of India". Sustainability Trends and Challenges in Civil Engineering, Part of Lecture Notes in Civil Engineering Book series, *Springer Singapore*, *162,* 657 to 668.

  DOI Permalink: https://doi.org/10.1007/978-981-16-2826-9_42

INTERNATIONAL CONFERENCES

i. Hiremath C. G., and Nandagiri L. (2020). "Regionalization of Flow Duration Curve for West Flowing Rivers of India". International Conference on Civil Engineering Trends and Challenges for Sustainability (CTCS 2020), NMAM Institute of Technology (Nitte), Udupi, India, 22nd to 23rd Dec, 2020.

   — Received Best Paper Presentation Certificate for the above-mentioned paper presented during the international conference.

ii. Hiremath C. G., and Nandagiri L. (2024). “Conceptual Hydrological Modelling Using AWBM”. International Conference on Hazardous Materials Management and Environmental Practices, Visvesvaraya Technological University, Belagavi, India, 12th to 13th Feb 2024.

— Received Best Paper Presentation Certificate (First Place) for the above-mentioned paper presented during the international conference.

BRIEF BIO

Name:	**CHANDRASHEKARAYYA G. HIREMATH**
Permanent Address:	Jyoti Nivas, Rajnagar-2 cross, Gulaganjikoppa, Dharwad. Karnataka - 580008
Education:	MASTER OF TECHNOLOGY (2007-09) *Discipline:* Water Resources Engineering and Management Indian Institute of Technology (IIT), Guwahati, Assam. BACHELOR OF ENGINEERING (2002-06) *Discipline:* Civil Engineering B. V. Bhoomaraddi College of Engineering and Technology, Hubballi, Karnataka. *Affiliated to:* Visvesvaraya Technological University, Belgaum.
Job Experience:	Industry: 3.5 years Teaching: 11.5 years
Contact No.:	+91-9620313624
Email ID:	iitgchandu@gmail.com
ORCID:	https://orcid.org/0000-0001-7407-270X
VIDWAN ID:	https://vidwan.inflibnet.ac.in/profile/111171

APPENDIX-I

Catchment Characteristics of Delineated Basins in West Flowing Rivers

Sl No.	Sub Basin Name	MAXe (km)	MINe (km)	ΔH	ΔH/P	S	A (km^2)	P (km)	L (km)	W (km)	L_p (km)	FF	SF	R_c	R_L	D_d
1	Santeguli	0.797	0.012	0.785	0.003	0.023	988	248	34	29	80	0.84	1.19	0.20	0.45	1.15
2	Avershe	1.152	0.014	1.138	0.007	0.029	276	168	39	7	49	0.182	5.49	0.12	0.38	1.63
3	Yennehole	1.225	0.020	1.205	0.009	0.045	334	127	27	12	36	0.465	2.15	0.26	0.58	1.70
4	Addoor	1.886	0.000	1.886	0.009	0.041	690	221	46	15	64	0.330	3.03	0.18	0.46	1.70
5	Bantwal	1.716	0.001	1.715	0.003	0.020	3204	491	85	37	128	0.438	2.28	0.17	0.50	0.65
6	Erinjipuzha	1.471	0.017	1.454	0.006	0.022	852	235	67	13	90	0.191	5.24	0.19	0.37	1.27
7	Kidangoor	1.188	0.002	1.186	0.006	0.029	593	193	40	15	56	0.364	2.75	0.20	0.49	1.58
8	Kalloopara	1.404	0.001	1.403	0.006	0.030	700	249	46	15	75	0.327	3.06	0.14	0.40	1.59
9	Thumpaman	1.921	0.008	1.913	0.007	0.028	796	282	69	11	91	0.165	6.04	0.13	0.35	1.59
10	Ayilam	1.705	0.006	1.699	0.008	0.037	533	208	46	12	58	0.257	3.89	0.16	0.45	1.67
11	Kuniyil	2.621	-0.006	2.627	0.007	0.040	1998	403	65	31	97	0.466	2.15	0.15	0.52	1.67
12	Karathodu	1.344	0.006	1.338	0.006	0.028	770	226	47	16	79	0.345	2.90	0.19	0.40	1.62
13	Kalampur	1.176	0.010	1.166	0.007	0.030	358	157	39	9	53	0.238	4.21	0.18	0.40	1.61
14	Mahuwa	1.388	0.005	1.383	0.003	0.014	1701	420	101	17	148	0.166	6.01	0.12	0.31	2.38
15	Haladi	0.965	0.004	0.961	0.006	0.032	524	174	30	17	56	0.583	1.72	0.22	0.46	0.76
16	Nanipalasan	1.031	0.094	0.937	0.004	0.020	719	215	46	16	78	0.335	2.99	0.20	0.39	0.72
17	Ozerkheda	1.195	0.078	1.117	0.006	0.035	659	177	31	21	60	0.664	1.51	0.26	0.48	0.77
18	Pulamanthole	2.373	0.007	2.366	0.008	0.038	904	288	63	14	76	0.227	4.40	0.14	0.44	0.74

Catchment Characteristics of Delineated Basins in Krishna River Basin

Sl No.	Sub Basin Name	MAXe (km)	MINe (km)	ΔH	ΔH/P	S	A (km^2)	P (km)	L (km)	W (km)	L_p (km)	FF	SF	R_c	R_L	D_d
1	Kellodu	1.921	0.652	1.269	0.002	0.014	4271	554	94	45	114	0.485	2.06	0.18	0.65	0.29
2	Talikot	0.753	0.501	0.252	0.000	0.002	2411	570	145	17	182	0.115	8.67	0.09	0.30	1.90
3	Navalgund	0.846	0.566	0.280	0.000	0.003	3080	600	91	34	96	0.373	2.68	0.11	0.65	1.89
4	Balehonnur	1.885	0.700	1.185	0.005	0.022	810	253	53	15	74	0.287	3.49	0.16	0.43	0.70
5	Khanapur	1.039	0.639	0.400	0.002	0.010	518	214	39	13	51	0.336	2.98	0.14	0.51	0.73
6	Marol	0.842	0.511	0.331	0.001	0.002	5142	622	133	39	205	0.291	3.44	0.17	0.39	0.74
7	Halia	0.704	0.128	0.576	0.001	0.006	3250	457	94	35	121	0.366	2.73	0.20	0.53	0.81
8	Naguleru	0.587	0.065	0.522	0.002	0.011	561	212	47	12	65	0.250	4.00	0.16	0.41	0.77

Catchment Characteristics of Delineated Basins in Cauvery River Basin

Sl No.	Sub Basin Name	MAXe (km)	MINe (km)	ΔH	ΔH/P	S	A (km^2)	P (km)	L (km)	W (km)	L_p (km)	FF	SF	R_c	R_L	D_d
1	Sakleshpura	1.423	0.889	0.534	0.002	0.011	601	214	47	13	53	0.277	3.61	0.16	0.53	1.65
2	KMVadi	1.595	0.769	0.826	0.002	0.013	1454	343	63	23	97	0.369	2.71	0.16	0.44	0.82
3	E_Mangalam	1.972	0.135	1.837	0.003	0.012	3465	576	152	23	172	0.150	6.65	0.13	0.39	1.82
4	Bendrahalli	1.815	0.630	1.185	0.003	0.017	1838	353	70	26	98	0.370	2.70	0.18	0.49	0.81
5	Hogenakkal	1.387	0.257	1.130	0.003	0.013	1562	373	87	18	126	0.208	4.81	0.14	0.35	0.78
6	Kudlur	1.669	0.443	1.226	0.006	0.034	720	195	36	20	52	0.560	1.79	0.24	0.59	0.73
7	Thevur	1.637	0.174	1.463	0.005	0.021	1190	291	69	17	93	0.247	4.05	0.18	0.42	0.81
8	Thoppur	1.637	0.331	1.306	0.008	0.040	332	172	33	10	48	0.310	3.22	0.14	0.43	0.77
9	Nellithurai	2.634	0.307	2.327	0.008	0.041	1482	301	57	26	89	0.458	2.18	0.21	0.49	0.77
10	Thengumarahada	2.632	0.343	2.289	0.008	0.041	1357	298	56	24	96	0.440	2.28	0.19	0.43	0.75
11	T. Bekuppe	1.457	0.611	0.846	0.002	0.007	3336	553	118	28	146	0.240	4.17	0.14	0.45	0.74

Catchment Characteristics of Delineated Basins in East flowing River Basin

Sl No.	Sub Basin Name	MAXe (km)	MINe (km)	ΔH	ΔH/P	S	A (km^2)	P (km)	L (km)	W (km)	L_p (km)	FF	SF	R_c	R_L	D_d
1	Kashipatnam	1.483	0.110	1.373	0.013	0.050	171	106	28	6	37	0.225	4.45	0.19	0.40	0.74
2	Seedhi	1.495	0.038	1.457	0.006	0.030	1116	249	49	23	79	0.459	2.18	0.23	0.48	0.78
3	Ambasamudram	1.902	0.300	1.602	0.007	0.033	703	219	49	14	71	0.296	3.38	0.18	0.42	0.79
4	Salur	1.619	0.140	1.479	0.010	0.035	308	151	42	7	52	0.174	5.75	0.17	0.38	0.80
5	Gunupur	1.505	0.075	1.430	0.002	0.012	6930	801	117	59	151	0.503	1.99	0.14	0.62	0.77

Catchment Characteristics of Delineated Basins in Cauvery River Basin

Sl No.	Sub Basin Name	MAXe (km)	MINe (km)	ΔH	ΔH/P	S	A (km^2)	P (km)	L (km)	W (km)	L_p (km)	FF	SF	R_c	R_L	D_d
1	Pedagedadda	0.929	0.186	0.743	0.007	0.025	181	110	30	6	35	0.204	4.90	0.19	0.43	0.73
2	Ramakona	1.170	0.337	0.833	0.002	0.008	2493	503	103	24	150	0.233	4.28	0.12	0.38	0.79
3	Wairagarh	0.660	0.197	0.463	0.001	0.006	1832	409	76	24	95	0.321	3.11	0.14	0.51	0.80
4	Amabal	0.787	0.530	0.257	0.001	0.003	1943	439	99	20	136	0.197	5.08	0.13	0.37	0.76
5	Tumnar	1.268	0.331	0.937	0.003	0.017	1732	368	56	31	87	0.556	1.80	0.16	0.54	0.79
6	Cherribeda	0.852	0.576	0.276	0.001	0.005	870	224	52	17	76	0.322	3.11	0.22	0.44	0.74
7	Gandlapet	0.669	0.312	0.357	0.001	0.006	1396	311	61	23	78	0.377	2.65	0.18	0.54	0.82
8	Sonarpal	0.910	0.532	0.378	0.001	0.007	1483	308	57	26	71	0.461	2.17	0.20	0.61	0.74