Evaluation of Implicit and Explicit Methods of Uncertainty Analysis on a Hydrological Modeling

Arpana Rani Datta

University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation

https://scholar.uwindsor.ca/etd/5397

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
Evaluation of Implicit and Explicit Methods of Uncertainty Analysis on a Hydrological Modeling

by

Arpana Rani Datta

A Dissertation Submitted to the Faculty of Graduate Studies through Civil Engineering in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy at the University of Windsor

Windsor, Ontario, Canada

2011

© 2011 Arpana Rani Datta
Evaluation of Implicit and Explicit Methods of Uncertainty Analysis on a Hydrological Modeling

by

Arpana Rani Datta

APPROVED BY:

Dr. Raghavan Srinivasan, External Examiner
Texas A&M University

Dr. Ram Balachandar
Department of Civil and Environmental Engineering

Dr. Abdul-Fattah Asfour
Department of Civil and Environmental Engineering

Dr. Ronald Barron
Department of Mathematics and Statistics

Dr. Tirupati Bolisetti, Advisor
Department of Civil and Environmental Engineering

Dr. Henry Hu, Chair of Defense
Department of Mechanical, Automotive and Materials Engineering

May 24, 2011
DECLARATION OF CO-AUTHORSHIP/ PREVIOUS PUBLICATION

I hereby declare that this thesis incorporates the outcome of a research carried out by the author under the supervision of Dr. Tirupati Bolisetti. The key ideas, experimental designs, data analysis and interpretation, were performed by the author, and the contributions of the co-authors was primarily through the provision of supervision.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

Based on the research work, the following two papers have been submitted for publication in peer reviewed journals:

<table>
<thead>
<tr>
<th>Thesis Chapter</th>
<th>Title of paper</th>
<th>Publication status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 4</td>
<td>Estimation of model parameters and assessment of uncertainty in a distributed hydrological model using Bayesian approach, submitted to <em>Journal of Hydrology</em></td>
<td>Under Review</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>Quantifying parameter and predictive uncertainties in distributed hydrological modeling considering input uncertainty, submitted to <em>Journal of Hydrology</em></td>
<td>Under Review</td>
</tr>
</tbody>
</table>
I certify that, to the best of my knowledge, my thesis does not infringe upon anyone’s copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.
ABSTRACT

Uncertainty in any hydrological modeling can be quantified either implicitly by lumping all sources of errors or explicitly by addressing different sources of errors individually. This dissertation has evaluated some implicit and explicit methods of uncertainty analysis for a physically based distributed hydrological model called Soil and Water Assessment Tool (SWAT). A multiplicative input error model has been developed considering season-dependent precipitation multipliers for quantifying precipitation uncertainty explicitly in the distributed hydrological modeling. The high-dimensional and computational problems of the existing explicit methods have lead to the development of the seasonal input error model. The model is implemented in the calibration process of SWAT for simulating streamflow in two watersheds of Southwestern Ontario, Canada. The calibration method is based on the Bayesian approach and the Markov Chain Monte Carlo (MCMC) simulations are performed by the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm to analyze the posterior probability distribution of model parameters. By keeping the number of precipitation multipliers equal to the number of distinct seasons, the seasonal input error model has reduced the number of latent variables in the Bayesian modeling and has reduced the dimension of posterior probability distribution.

The study reveals that streamflow prediction uncertainty due to parameter uncertainty is reduced when the autoregressive models are used in the implicit methods to represent the residual errors. However, the model parameters are biased when the Box-Cox transformation of data is used in the calibration process for addressing non-homogeneity and non-normality of the residual errors. The parameter and prediction
uncertainties estimated by the seasonal input error model based calibration method are consistent with that of implicit methods. Model structural uncertainty is observed to be dominating over the input and parameter uncertainties in modeling the study area with SWAT. Hence, the autoregressive models as well as the input error models could not provide global optimum values in the parameter space. The seasonal input error model quantifies that the true precipitation is lower than the measured precipitation and the precipitation uncertainty estimated by the model is comparable to that of existing input error models. The effects of seasonal precipitation multipliers on parameter estimation and model prediction are explained by the correlation of estimated model parameters and by the reliability of model prediction uncertainty.
DEDICATION

To my parents
ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my advisor Dr. T. Bolisetti for his continuous inspiration, guidance and support during my PhD study at the University of Windsor.

I sincerely acknowledge my gratitude to the committee members Dr. R. Balachandar, Dr. A. Asfour and Dr. R. Barron for their valuable time and suggestions for the study.

I am indebted to Dr. J. A. Vrugt for the source codes of SCEM-UA algorithm for Markov Chain Monte Carlo simulation. I would also like to thank Dr. S. Paul for reviewing the mathematical form of the likelihood function considering the second order autoregressive model.

I am grateful to P. Seguin for providing the necessary technical support for running thousands of computer simulation. I would like to thank the Essex Region Conservation authority for providing the GIS data.

I acknowledge the financial support provided by the Ontario Graduate Scholarship, Sustainable Engineering Scholarship and University of Windsor Tuition Scholarship for carrying out my research.

I would like to acknowledge the encouragement and support that I have received from my family during my study. I would also like to thank my colleagues, with whom I have shared an office and carried out my research in a friendly environment.
TABLE OF CONTENTS

DECLARATION OF CO-AUTHORSHIP/ PREVIOUS PUBLICATION .................. iii
ABSTRACT .................................................................................................................. v
DEDICATION ............................................................................................................. vii
ACKNOWLEDGEMENTS ......................................................................................... viii
LIST OF TABLES ....................................................................................................... xii
LIST OF FIGURES ................................................................................................. xiv

CHAPTER

I. INTRODUCTION

1.1 General .............................................................................................................. 1
1.2 Uncertainty analysis in hydrological modeling ................................................. 2
1.3 Input uncertainty in distributed hydrological modeling................................. 5
1.4 Objectives of the research .............................................................................. 8
1.5 Scope of the research ....................................................................................... 9
1.6 Significance of the research ............................................................................ 9
1.7 Organization of the dissertation ..................................................................... 10

II. LITERATURE REVIEW

2.1 Introduction ....................................................................................................... 12
2.2 Bayesian theory ............................................................................................. 13
2.3 Hierarchical Bayesian modeling .................................................................... 14
2.4 Methods of uncertainty analysis ..................................................................... 14
2.5 Uncertainty analysis of SWAT model ............................................................. 21
2.6 Current state of knowledge ........................................................................... 23
2.7 Summary ......................................................................................................... 24

III. METHODOLOGY

3.1 Introduction ...................................................................................................... 25
3.2 SWAT model .................................................................................................... 28
3.3 Selection of calibration parameters ................................................................ 29
3.4 Computational framework ............................................................................. 32
3.5 Experimental design ....................................................................................... 35
3.6 Posterior diagnostics ...................................................................................... 36
IV. IMPLICIT METHODS OF UNCERTAINTY ANALYSIS

4.1 Introduction ...........................................................................................................39
4.2 Uncertainty analysis using AR models ...............................................................42
  4.2.1 Likelihood function for white noise model ..............................................42
  4.2.2 Box-Cox transformation of data ...............................................................43
  4.2.3 Likelihood function for AR(1) model .......................................................44
  4.2.4 Likelihood function for continuous AR model ........................................45
4.3 Formulation of likelihood function with AR(2) model .......................................47
4.4 Evaluation of implicit methods ..........................................................................50
  4.4.1 Study area, model and data ......................................................................50
  4.4.2 Methodology .............................................................................................53
  4.4.3 Estimation of parameter uncertainty .......................................................55
  4.4.4 Estimation of prediction uncertainty .......................................................65
  4.4.5 Test of residual errors ..............................................................................76
4.5 Summary .............................................................................................................91
4.6 Conclusions .........................................................................................................93

V. EXPLICIT METHODS OF UNCERTAINTY ANALYSIS

5.1 Introduction .........................................................................................................94
5.2 Uncertainty analysis using multiplicative input error model .........................96
  5.2.1 Posterior pdf for storm input error model ............................................96
  5.2.2 Posterior pdf for daily input error model .............................................99
  5.2.3 Posterior pdf of Standard calibration method ..................................100
5.3 Development of seasonal input error model ..................................................101
  5.3.1 The conceptual basis of the seasonal input error model ..................101
  5.3.2 Posterior pdf for the seasonal input error model .............................102
5.4 Evaluation of seasonal input error model .......................................................105
  5.4.1 Methodology .........................................................................................105
  5.4.2 Identifying the seasonal precipitation multipliers .............................108
  5.4.3 Estimation of parameter uncertainty and input uncertainty ..............109
  5.4.4 Estimation of prediction uncertainty ...................................................117
  5.4.5 Test of residual errors ..........................................................................122
5.5 Uncertainty analysis by storm input error model based calibration
  method ..................................................................................................................127
  5.5.1 Identifying the precipitation events .....................................................127
  5.5.2 Convergence of Markov Chains ..........................................................127
  5.5.3 Estimation of precipitation uncertainty .................................................129
  5.5.4 Comparison with Standard calibration method ................................129
5.5.5 Comparison with seasonal input error model and daily input error model based calibration methods ..................................................135
5.6 Summary .........................................................................................137
5.7 Conclusions ......................................................................................141

VI. APPLICATION OF SEASONAL INPUT ERROR MODEL

6.1 Introduction .........................................................................................143
6.2 Evaluation of seasonal input error model for the Ruscom River watershed ..............................................................................143
   6.2.1 Study area, model and data ............................................................143
6.2.2 Methodology ..................................................................................147
6.2.3 Estimation of parameter uncertainty and input uncertainty ..................................................................................148
6.2.4 Estimation of prediction uncertainty ..................................................155
6.2.5 Test of residual errors ......................................................................161
6.3 Comparison with the results of the Canard River watershed .......165
6.4 Comparison of seasonal input error model and AR(1) model based calibration methods .................................................................170
   6.4.1 Methodology .................................................................................170
6.4.2 Uncertainty analysis of the Ruscom River watershed modeling ..................................................................................172
6.4.3 Uncertainty analysis of the Canard River watershed modeling ..................................................................................184
6.5 Summary .........................................................................................196
6.6 Conclusions ......................................................................................198

VII. CONCLUSIONS AND FUTURE WORK

7.1 Conclusions .........................................................................................199
7.2 Future work .......................................................................................202

REFERENCES ..........................................................................................204

VITA AUCTORIS .....................................................................................216
LIST OF TABLES

Table 3.1: List of methods used for SWAT model calibration ........................................... 38
Table 4.1: Summary of calibration of SWAT model considering input uncertainty
indirectly in the calibration process .............................................................................. 54
Table 4.2: The prior ranges of parameters ....................................................................... 55
Table 4.3: Mean (standard deviation) of SWAT model parameters in different calibration
methods ..................................................................................................................... 61
Table 4.4: Mean (standard deviation) of AR model parameters and Box-Cox
transformation parameters in different calibration methods ...................................... 61
Table 4.5: The prior ranges of low sensitive model parameters ...................................... 63
Table 4.6: Efficiency of SWAT model parameters obtained at the maximum posterior
density during calibration period .............................................................................. 66
Table 4.7: NS values at different seasons for streamflow simulation during calibration
and validation periods .............................................................................................. 66
Table 4.8: Effects of NS values at different seasons for streamflow simulation
considering high sensitive and high and low sensitive parameters ......................... 68
Table 4.9: Average annual evapotranspiration and streamflow using the model
parameters at the maximum posterior density ......................................................... 68
Table 4.10: Efficiency of optimum values of model parameters in streamflow prediction
during validation period ......................................................................................... 68
Table 4.11: Percentage of observed streamflow data covered by 95% prediction
uncertainty due to parameter uncertainty during model calibration and validation
periods ..................................................................................................................... 70
Table 4.12: Percentage of observed streamflow data covered by 95% prediction
uncertainty during model calibration and validation periods ..................................... 70
Table 5.1: Summary of the explicit methods used for SWAT model calibration .......... 107
Table 5.2: The prior ranges of parameters ....................................................................... 110
Table 5.3: Comparison of mean (standard deviation) of SWAT model parameters ..... 132
Table 5.4: Correlation between estimated model parameters ........................................ 132
Table 5.5: Efficiency of SWAT model parameters obtained at the maximum posterior
density ..................................................................................................................... 132
Table 5.6: Percentage of observed streamflow data covered by 95% prediction
uncertainty .............................................................................................................. 134
Table 5.7: Quality of data covered by 95% prediction uncertainty ................................. 134
Table 5.8: Comparison of three calibration methods based on multiplicative input error
model ...................................................................................................................... 138
Table 6.1: The prior ranges of parameters ................................................................. 148
Table 6.2: Correlation between estimated SWAT model parameters ......................... 150
Table 6.3: Streamflow and precipitation characteristics of the watersheds for the
period of 1981-2000 .................................................................................................. 167
Table 6.4: Comparison of results with seasonal input error model for two watersheds 168
Table 6.5: The prior ranges of parameters ................................................................. 171
Table 6.6: Optimum values of SWAT model parameters with 95% confidence limits
for the Ruscom River watershed ............................................................................. 174
Table 6.7: Correlation of SWAT model parameters for the Ruscom River watershed .. 175
Table 6.8: Efficiency of optimum parameter values for streamflow simulation in the Ruscom River watershed ................................................................. 175
Table 6.9: Percentage of observed streamflow data covered by 95% prediction uncertainty in the Ruscom River watershed .................................................. 177
Table 6.10: Reliability of streamflow prediction in the Ruscom River watershed .... 179
Table 6.11: Correlation of SWAT model parameters for the Canard River watershed.. 186
Table 6.12: Optimum values of SWAT model parameters with 95% confidence limits for the Canard River watershed ............................................................... 186
Table 6.13: Efficiency of optimum parameter values for streamflow simulation in the Canard River watershed ................................................................. 187
Table 6.14: Percentage of observed streamflow data covered by 95% prediction uncertainty in the Canard River watershed .................................................. 189
Table 6.15: Reliability of streamflow prediction in the Canard River watershed ....... 192
LIST OF FIGURES

Figure 3.1: The flow chart for calibration of SWAT model under any uncertainty framework............................................................................................................. 27
Figure 3.2: Movement of water simulated by SWAT at the HRU level for the study area (Adapted from Neitsch et al., 2005)........................................................................................................ 30
Figure 3.3: The computational framework of SWAT model calibration considering the input error model.................................................................................................................. 34
Figure 4.1: Location of the Canard River watershed................................................................................................................................. 51
Figure 4.2: Delineation of the Canard River watershed into sub-basins .................... 52
Figure 4.3: Marginal posterior pdfs of model parameters in white noise and AR model based calibration methods................................................. 56
Figure 4.4: Marginal posterior pdfs of model parameters in data transformation based calibration methods.......................................................... 57
Figure 4.5: Marginal posterior pdf of AR model parameters in AR model based calibration methods.................................................................................. 58
Figure 4.6: Marginal posterior pdf of AR model parameters and transformation parameters in data transformation based calibration methods................................................................. 59
Figure 4.7: Correlation between the estimated model parameters.............................. 62
Figure 4.8: Verification of parameter non-identifiability in the calibration process ...... 64
Figure 4.9: Streamflow prediction uncertainty due to parameter uncertainty in calibration period in Standard and AR(1) model based calibration methods ........................................... 72
Figure 4.10: Streamflow prediction uncertainty due to parameter uncertainty in calibration period in AR(2) model and continuous AR model based calibration methods .................................................................................................................. 73
Figure 4.11: Streamflow prediction uncertainty due to total uncertainty in calibration period in Standard and AR(1) model based calibration methods ................................................................................................................................. 74
Figure 4.12: Streamflow prediction uncertainty due to total uncertainty in calibration period in AR(2) model and continuous AR model based calibration methods ......... 75
Figure 4.13: Streamflow prediction uncertainty due to parameter uncertainty in validation period for Standard and AR(1) model based calibration methods................................. 77
Figure 4.14: Streamflow prediction uncertainty due to parameter uncertainty in validation period for AR(2) model and continuous AR model based calibration methods ...... 78
Figure 4.15: Streamflow prediction uncertainty due to total uncertainty in validation period for Standard and AR(1) model based calibration methods......................................................... 79
Figure 4.16: Streamflow prediction uncertainty due to total uncertainty in validation period in AR(2) model and continuous AR model based calibration methods .................................................................................................................. 80
Figure 4.17: Test of homoscedasticity of standardized residuals .................................. 81
Figure 4.18: ACF and PACF plot of residuals with 95% limits in Standard calibration method................................................................................................................................. 83
Figure 4.19: ACF plot of residuals with 95% limits in AR model based and data transformation based calibration methods ........................................................................................................ 85
Figure 4.20: Normality plot of standardized residuals in Standard and AR model based calibration methods ................................................................................................................................. 87
Figure 4.21: Normality plot of standardized residuals in data transformation based calibration methods................................................................................................................................. 88
Figure 5.1: Marginal posterior pdfs of SWAT model parameters in Standard, seasonal input error model and daily input error model based calibration methods
Figure 5.2: Box plots of marginal posterior probability distribution of seasonal input error model parameters
Figure 5.3: Deviation of estimated precipitation by seasonal input error model against the measured precipitation
Figure 5.4: Marginal posterior probability distribution of daily input error model parameters
Figure 5.5: Comparison of observed and estimated precipitation and observed and simulated streamflow in seasonal input error model and daily input error model based calibration methods
Figure 5.6: Probability distribution function of DRMSE in different calibration methods.
Figure 5.7: Streamflow prediction uncertainty due to total uncertainty and parameter uncertainty in the calibration period
Figure 5.8: Streamflow prediction uncertainty due to total uncertainty and parameter uncertainty in the validation period
Figure 5.9: Predictive QQ plot in calibration and validation periods
Figure 5.10: (a) QQ plot of standardized residuals and (b) ACF of residuals with 95% probability limits during calibration
Figure 5.11: Test of homoscedasticity of standardized residuals during calibration
Figure 5.12: Identification of precipitation events in the study area from February, 1992 to January, 1993
Figure 5.13: Marginal posterior pdf of precipitation multipliers
Figure 5.14: Marginal posterior pdf of SWAT model parameters
Figure 5.15: Streamflow prediction using the parameter values at the maximum posterior density
Figure 5.16: QQ plot of standardized residuals in Standard and Storm_input_error methods
Figure 5.17: ACF plot of residuals in Standard and Storm_input_error methods
Figure 6.1: Location of the Ruscom River watershed
Figure 6.2: Delineation of the Ruscom River watershed into sub-basins
Figure 6.3: Marginal posterior pdfs of SWAT model parameters in Standard and seasonal input error model based calibration methods
Figure 6.4: Box plots of marginal posterior probability distribution of seasonal input error model parameters
Figure 6.5: Deviation of estimated precipitation by seasonal input error method against the measured precipitation
Figure 6.6: Comparison of observed and estimated precipitation and observed and simulated streamflow in seasonal input error model based calibration method and Standard calibration method
Figure 6.7: Probability distribution function of DRMSE
Figure 6.8: Streamflow prediction uncertainty due to total uncertainty and parameter uncertainty in the calibration period. ................................................................. 158
Figure 6.9: Streamflow prediction uncertainty due to total uncertainty and parameter uncertainty in the validation period. ................................................................. 159
Figure 6.10: Predictive QQ plot in calibration and validation periods. ...................... 160
Figure 6.11: QQ plot of standardized residuals during calibration.......................... 162
Figure 6.12: ACF plot of residuals with 95% probability limits during calibration ...... 163
Figure 6.13: Test of homoscedasticity of standardized residuals during calibration ...... 164
Figure 6.14: Location of the Canard River and Ruscom River watersheds.................. 166
Figure 6.15: Marginal posterior pdfs of SWAT model parameters for the Ruscom River watershed ................................................................. 173
Figure 6.16: Marginal posterior pdf of AR(1) model parameters for the Ruscom River watershed .............................................................................................................. 174
Figure 6.17: Posterior pdfs of standard deviation of errors in AR(1) model based calibration method and DRMSE in seasonal input error model based calibration method in the Ruscom River watershed ................................................................. 176
Figure 6.18: Predictive QQ plots in calibration and validation periods in the Ruscom River watershed. ............................................................................................................. 178
Figure 6.19: QQ plot of standardized residuals in the Ruscom River watershed........... 180
Figure 6.20: ACF of residuals with 95% probability limits in the Ruscom River watershed. .......................................................................................................................... 181
Figure 6.21: Test of homoscedasticity of standardized residuals in the Ruscom River watershed. .......................................................................................................................... 182
Figure 6.22: Cumulative periodogram of residuals with 95% limits in the Ruscom River watershed ............................................................................................................. 183
Figure 6.23: Marginal posterior pdfs of SWAT model parameters for the Canard River watershed .................................................................................................................. 185
Figure 6.24: Marginal posterior pdf of AR(1) model parameters for the Canard River watershed .................................................................................................................. 187
Figure 6.25: Posterior pdfs of standard deviation of errors in AR(1) model based calibration method and DRMSE in seasonal input error model based calibration method in the Canard River watershed ................................................................. 188
Figure 6.26: Predictive QQ plots in calibration and validation periods in the Canard River watershed. ............................................................................................................. 191
Figure 6.27: QQ plot of standardized residuals in the Canard River watershed............ 192
Figure 6.28: ACF of residuals with 95% probability limits in the Canard River watershed .......................................................................................................................... 193
Figure 6.29: Test of homoscedasticity of standardized residuals in the Canard River watershed. .......................................................................................................................... 194
Figure 6.30: Cumulative periodogram of residuals with 95% limits in the Canard River watershed ................................................................. 195
CHAPTER I
INTRODUCTION

1.1 General

The hydrological models are used for generating information on different components of the hydrologic system. The model needs to be calibrated against the observed data over a historical period of time before using the modeling results. During the process of calibration, the model parameters are estimated such that the modeling results are close to the observations of the real world system. There is uncertainty in the results of any modeling that arises from different sources (Kay et al., 2009). The uncertainties in the hydrological modeling are due to the uncertainty in model inputs, parameters, structure and outputs (Thyer et al., 2009; Yang et al., 2008; Ajami et al., 2007; Huard and Mailhot, 2006; Kavetski et al., 2006a; Vrugt, 2004). The uncertainty in model inputs, such as precipitation, temperature, evapotranspiration etc., can result from their measurement errors. The uncertainty in model parameters may arise from the non-identifiable model parameters and non-uniqueness of identifiable model parameters. The non-identifiable problem arises when model parameters are not identified as the parameters that are required to be estimated through the calibration process. The problem of non-uniqueness or equifinality arises when different sets of model parameters produce similar observed responses for the hydrologic system. The uncertainty in model structure is due to the simplification of the complex hydrological system and inadequate representation of the system (Abbaspour, 2008). The uncertainty in model outputs is from the measurement errors of the observed data. All sources of uncertainty can propagate through the water resources system management and affect the performance of the
system (Ajami et al., 2008). Thus, proper quantification of uncertainty in model inputs, parameters and predictions is vital for different water resources management problems, such as watershed management, flood control and flood management, aquifer management, reservoir management etc. Considering the importance of uncertainty estimation, Pappenberger and Beven (2006) recommended to develop a 'Code of Practice' for making the uncertainty analysis as an integral part of the hydrological modeling process.

1.2 Uncertainty analysis in hydrological modeling

In the last two decades, many uncertainty analysis techniques were developed to account for different sources of uncertainty explicitly or implicitly in the hydrological modeling. The traditional uncertainty analysis techniques assume that all sources of uncertainties in the hydrological modeling can be accounted for by the parameter uncertainty. Some examples of these techniques are Sequential Uncertainty Fitting (SUFI-2) (Abbaspour et al., 2007, 2004), Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992), etc. In SUFI-2, the model parameters are calibrated so that most of the observed data fall within the 95% prediction uncertainty bound (Abbaspour et al., 2007). The GLUE methodology is an informal Bayesian approach (Vrugt et al., 2009) and based on the concept of equifinality of model structure and/or parameter sets in providing 'behavioral' fits to observational data (Zheng and Keller, 2007). The GLUE methodology has subjectivity in defining the likelihood function and the behavioral criterion of the model (Blasone et al., 2008). For avoiding the subjectivity in the likelihood function, some direct methods were introduced to account for different sources of uncertainty in hydrological modeling. In the direct methods, the uncertainties
in model inputs, structure and outputs are accounted for explicitly by introducing appropriate error models to the calibration framework. The Bayesian Total Error Analysis (BATEA) framework developed by Kavetski et al. (2006a) and the Bayesian framework developed by Huard and Mailhot (2008) fall under the explicit methods. In BATEA (Kavetski et al., 2006a), the input uncertainty is accounted for by assuming a multiplicative error model, the structural uncertainty is represented by varying some model parameters stochastically and the output uncertainty is accounted for by an additive error model. In BATEA, the input and structural error parameters are treated as latent variables to the hierarchical Bayesian modeling. However, Huard and Mailhot (2008) represented different sources of errors by additive error models and the model input and output time series were treated as latent variables to the Bayesian system. In the Bayesian approach, the model parameters are considered as probabilistic variables and the posterior probability density function of parameters are estimated by conditioning on the observed data (Vrugt, 2004; Engeland and Gottschalk, 2002). The parameter inferences are often made by the Markov Chain Monte Carlo (MCMC) methods for estimating the posterior probability density function of model parameters. The posterior probability density function is proportional to the product of the likelihood function and the prior probability density function of parameters. The dimension of the posterior probability distribution increases with the increase of number of variables needed to be inferred and the numerical solution of posterior distribution becomes computationally intensive. Hence, the high-dimensional problem of posterior probability distribution as well as the extensive computational problem arise when the frameworks developed by
Kavetski et al. (2006a) and Huard and Mailhot (2008) are applied for the long calibration period.

Ajami et al. (2007) introduced the Integrated Bayesian Uncertainty Estimator (IBUNE) framework to account for model input, parameter and structural uncertainties. They used the multiplicative error model to account for input uncertainty and the multi-model combination technique to account for the structural uncertainty. Reichert and Mieleitner (2009) corrected the bias in model input and structure explicitly by introducing the stochastic, time-dependent model parameters. In this approach, the time-dependent model parameter is considered as the multiplicative factor of, or additive term to, the model input (Reichert and Mieleitner, 2009) and thus, it is conceptually similar to the multiplicative error model of BATEA. Data assimilation techniques (Salaman and Feyen, 2009; Moradkhani et al., 2005; Vrugt et al., 2005) are often used to account for different sources of uncertainty in hydrological modeling. In the method, the state variables are estimated at each time step of model simulation and thus, it has a computational burden (Yang et al., 2007a).

In the indirect methods of uncertainty analysis, the errors in model inputs, parameters, structure and outputs are lumped together and expressed implicitly as an additive error model. Some examples of this category are the works of Schoups and Vrugt (2010), Laloy et al. (2010), Schaefli et al. (2007), Yang et al. (2007a,b), Engeland et al. (2005), Bates and Campbell (2001), Duan et al. (1988), Kuczera (1983), and Sorooshian and Dracup (1980). The additive error model in the likelihood function aims to make the residuals to be independent and normally distributed with zero mean and constant variance. In most of the cases, the residuals are correlated, non-normal and have
non-constant variance (heteroscedastic). The autoregressive (AR) models are commonly adopted in the implicit methods to account for the correlated errors of residuals (Laloy et al., 2010; Vrugt et al., 2009; Schaeffli et al., 2007; Yang et al., 2007a,b; Bates and Campbell, 2001; Duan et al., 1988; Kuczera, 1983; Sorooshian and Dracup, 1980) and the Box-Cox transformation (Box and Cox, 1964) of data is used to reduce the heteroscedasticity and non-normality of the errors (McLeod et al., 1977; Box and Tiao, 1973).

The explicit methods of uncertainty analysis have some advantages over the implicit methods. In the explicit methods, the effects of different sources of errors on the uncertainty in model prediction can be quantified separately (Renard et al., 2010; Thyer et al., 2009; Huard and Mailhot, 2008). However, the explicit methods are very challenging when applied to the distributed hydrological modeling. A large number of model parameters are used in a distributed hydrological model to describe the hydrologic system and it becomes very difficult to identify the effects of the error models on parameter estimation and model prediction (Abbaspour, 2008). In addition, the computational burden is a constraint for using the explicit methods in uncertainty analysis of the distributed hydrological modeling. Due to the challenges of the explicit methods, the implicit methods are often used for quantifying uncertainty in the distributed hydrological modeling.

1.3 Input uncertainty in distributed hydrological modeling

The uncertainty in precipitation data is the major source of input uncertainty in any hydrological modeling. This type of uncertainty may result from the errors in precipitation measurement and the errors due to its imperfect representation in the
hydrological modeling (Huard and Mailhot, 2006). The precipitation measurement errors may occur at a station due to the effects of wind and evaporation during its measurement and/or instrument error (Salamon and Feyen, 2009). Even though the precipitation measurement is exact, there might be differences between the gauge readings and the model inputs due to the spatial scale difference (Huard and Mailhot, 2006). This difference can be treated as the errors due to imperfect representation of precipitation. Hwang (2005) identified the interpolation techniques of precipitation as a source of input uncertainty in hydrological modeling. The uncertainty in model inputs propagates through the calibration process and causes biasedness in parameter estimation. This results in an increase in model prediction uncertainty. Therefore, the input uncertainty needs to be taken into account during the model calibration process. In BATEA (Kavetski et al., 2006a), the systematic measurement errors of precipitation data are corrected during the calibration process directly by the rainfall multipliers, which are the latent variables to the hierarchical Bayesian system. The temporal scale of the multipliers is either daily or storm-event basis. Thus the dimension of the posterior probability distribution is very high. The input uncertainty represented by the additive input error model (Huard and Mailhot, 2008) has the dimensional problem as well, when the resolution of temporal scale is finer than a month. The sequential data assimilation techniques used to account for input uncertainty are also computationally intensive.

Due to the dimensional and computational constraints of the existing multiplicative input error model, additive input error model and sequential data assimilation method, input uncertainty is commonly corrected implicitly in aggregation with other sources of uncertainty in distributed hydrological modeling (Zhang et al.,
Zhang et al. (2009) used the combined method of Genetic Algorithms (GA) and Bayesian Model Averaging (BMA) for calibration and uncertainty analysis of SWAT model. Li et al. (2009) used the Metropolis algorithm based MCMC approach for uncertainty analysis of SWAT model. Yang et al. (2007a,b) used the continuous time AR model to account for different sources of uncertainty in SWAT model prediction.

The GLUE methodology (Beven and Binley, 1992) has often been used for uncertainty analysis in the distributed hydrological modeling. Some examples are the research of Younger et al. (2009), Yang et al. (2008), Blasone et al. (2008) and Arabi et al. (2007). Younger et al. (2009) applied the GLUE methodology to study the effects of spatial variability of rainfall on TOPMODEL (Beven et al., 1995). Yang et al. (2008) applied GLUE methodology (Beven and Binley, 1992) for analyzing uncertainty of SWAT model (Arnold et al., 1998). Blasone et al. (2008) used the GLUE methodology for assessing all sources of uncertainty of MIKE-SHE model (Graham and Butts, 2006) during multi-response and multi-site calibration. Arabi et al. (2007) used the GLUE methodology for analyzing uncertainty of water quality estimates of SWAT model (Arnold et al., 1998) for the best management practices. Salamon and Feyen (2009) used the sequential data assimilation technique with the particle filter and assessed the uncertainties in model parameter, precipitation and model prediction associated with LISFLOOD model (De Roo et al., 2000).

The recent research direction in any hydrological modeling is to quantify the effects of different sources of errors on model prediction. The effects of different sources of errors on parameter uncertainty and prediction uncertainty can be quantified by using
the explicit methods of uncertainty analysis. Hence, more studies are needed to reduce the dimensional problem of posterior probability distribution so that the explicit methods can be considered as a robust method of uncertainty analysis and can be practiced to quantify uncertainty in the distributed hydrological modeling.

1.4 Objectives of the research

This dissertation addresses the existing limitations of the explicit methods of uncertainty analysis and aims to develop a methodology under the Bayesian approach to account for precipitation uncertainty explicitly in the calibration process of a distributed hydrological model. The study is carried out with a widely-used distributed hydrological model called Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998). The specific objectives of the dissertation are described as follows:

i) Quantifying the uncertainty in parameter estimation and model prediction by the implicit methods of uncertainty analysis. The purpose of this objective is to identify the merits and limitations of the implicit methods for the distributed hydrological models.

ii) Quantifying the parameter uncertainty and prediction uncertainty by implementing the existing multiplicative input error models. This objective is carried out to illustrate the need for development of a new method for the treatment of precipitation uncertainty.

iii) Development of a new seasonal input error model to account for precipitation uncertainty in the distributed hydrological modeling. The method is developed by introducing the season-dependent parameters to the multiplicative input error model.

iv) Evaluation of the seasonal input error model by quantifying input uncertainty, parameter uncertainty and streamflow prediction uncertainty. The purpose of this
objective is to identify the robustness of the seasonal input error model in parameter estimation and model prediction.

v) Application of the seasonal input error model to another watershed having similar hydrologic and climatic conditions. The purpose of this objective is to investigate the performance of the seasonal input error model for analyzing uncertainty of watershed modeling.

1.5 Scope of the research

The dissertation carried out with the above objectives is expected to strengthen the explicit methods of uncertainty analysis. The newly developed method is expected to reduce the uncertainty in parameter estimation during calibration process and to reduce the uncertainty in model prediction. The reduction of biasedness in parameter estimation is important for parameter regionalization, while the improvement in model prediction is useful for managing the extreme hydrological events. In addition, the newly developed input error model expects to reduce the existing high-dimensional problem of the multiplicative input error model and to identify the effects of input error model on parameter estimation and model prediction in the distributed hydrological modeling.

1.6 Significance of the research

This research quantifies uncertainty in hydrological modeling arisen from inputs and model parameters. The methodology of this research can be extended to the uncertainty analysis of other water resources modeling studies, such as, hydraulic modeling, water quality modeling and climate change impact studies. This research is also significant for the studies related to transferring model parameters to the ungauged basins. In this dissertation, an error model is developed for quantifying input uncertainty
explicitly in hydrological modeling. This research is probably the first attempt to extend the explicit method of uncertainty analysis to distributed hydrological modeling. The methodology developed in this dissertation will contribute to reducing dimensional problem and computational cost of solving the posterior probability distributions. Furthermore, this research develops a calibration method by representing the residual errors with the second order autoregressive model. This is probably the first attempt to implement the second order autoregressive model in the calibration process of any hydrological modeling. Uncertainty estimation is usually communicated to the decision makers for understanding the risk associated with uncertainty in modeling results. Therefore, this research will contribute to managing water resources system.

1.7 Organization of the dissertation

The dissertation is organized into seven chapters. The literature related to the existing uncertainty analysis methods is summarized in Chapter II. The methodology for carrying out the objectives of the dissertation is presented in Chapter III. The first objective is addressed in Chapter IV while the second, third and fourth objectives are addressed in Chapter V. In Chapter IV, precipitation uncertainty is accounted for implicitly along with other sources of uncertainties and the results are presented. In Chapter V, precipitation uncertainty is taken care of explicitly and the seasonal input error model is developed. The performance of the seasonal input error model is evaluated in comparison with other existing multiplicative input error models. In Chapter VI, the performance of the seasonal input error model for another watershed is evaluated. A comparison is also made between the results obtained from the implicit method and seasonal input error model based explicit method of uncertainty analysis for two case
studies. Finally, in Chapter VII, the findings of the dissertation are presented in the 'conclusions' section and recommendations for future research are described in the 'future work' section.
CHAPTER II
LITERATURE REVIEW

2.1 Introduction

Most of the uncertainty analysis methods used in hydrological modeling are based on the Bayesian approach. These methods can be classified into three major categories based on how different sources of uncertainty are considered in the methods. Yang et al. (2008) described these methods as: i) all uncertainties represented by parameter uncertainty [Sequential Uncertainty Fitting (SUFI-2) (Abbaspour et al., 2007, 2004); Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992)]; ii) the input and model structural uncertainty considered implicitly by introducing an additive error model [Schoups and Vrugt (2010); Laloy et al. (2010); Schaeffli et al. (2007); Yang et al. (2007a,b); Bates and Campbell (2001); Duan et al. (1988); Kuczera (1983); Sorooshian and Dracup (1980)]; and iii) the input and/or model structural uncertainty considered explicitly by using the stochastic time-dependent parameters (Reichert and Mieleitner, 2009); additive input error model (Huard and Mailhot, 2008); multiplicative input error model (Ajami et al., 2007, Kavetski et al., 2006a; Kuczera et al., 2006); Sequential Data Assimilation (SDA) method (Moradkhani et al., 2005; Vrugt et al., 2005. These methods of uncertainty analysis (UA) are described in this dissertation as UA method-type 1, UA method-type 2 and UA method-type 3. This chapter briefly discusses the Bayesian theory and hierarchical Bayesian modeling, the concepts, application and limitations of three types of UA methods in hydrological modeling and presents some examples of uncertainty analyzing techniques of SWAT model (Arnold et
al., 1998). Moreover, the current state of knowledge in the field of 'uncertainty analysis' and the research gaps are presented.

2.2 Bayesian theory

In the Bayesian approach, the model parameters are considered as probabilistic variables having a joint posterior probability density function, which captures the probabilistic beliefs about the parameters conditioned on the observed data (Vrugt, 2004). According to the Bayesian theory, the posterior probability distribution of model parameters, \( p(\theta | y) \) is expressed as follows (Gelman et al., 2004):

\[
p(\theta | y) = \frac{p(\theta, y)}{\int p(\theta) p(y|\theta) d\theta} = \frac{p(\theta)p(y|\theta)}{\int p(\theta)p(y|\theta)d\theta}
\]  

(2.1)

where \( p(\theta) \) is the prior distribution of parameters, \( p(y|\theta) \) is the sampling distribution\( \int p(\theta)p(y|\theta)d\theta \) is known as the normalizing constant. For a fixed \( y \), this equation can be written as:

\[
p(\theta | y) \propto p(\theta)p(y|\theta)
\]  

(2.2)

The data \( y \) affects the posterior inference through the function \( p(y|\theta) \). When the data \( y \) are given, \( p(y|\theta) \) can be considered as a function of \( \theta \) which is known as the likelihood function of \( \theta \) given \( y \) (Gelman et al., 2004) and can be expresses as \( l(\theta | y) \). Thus eqn. (2.2) can be written as:

\[
p(\theta | y) \propto p(\theta)l(\theta | y)
\]  

(2.3)
2.3 Hierarchical Bayesian modeling

According to Bayesian theory, for the parameters $\theta$ and given data, $y$, the posterior probability distribution of $\theta$ is written as (eqn. (2.2))

$$p(\theta|y) \propto p(y|\theta)p(\theta) \tag{2.4}$$

where $p(\theta)$ is the prior distribution of $\theta$. If the prior distribution of $\theta$ depends on some other parameters, $\tau$; according to hierarchical Bayesian modeling the posterior probability distribution can be written as follows (Gelman et al., 2004):

$$p(\theta, \tau|y) \propto p(y|\theta, \tau)p(\theta, \tau) \tag{2.5}$$

The prior $p(\theta, \tau)$ can be replaced by a prior $p(\theta|\tau)$, and a prior of $\tau$, $p(\tau)$, and

$$p(\theta, \tau|y) \propto p(y|\theta, \tau)p(\theta|\tau)p(\tau) \tag{2.6}$$

The parameters $\tau$ are known as the hyperparameters in the hierarchical Bayesian modeling. These variables are introduced in the system to modify the posterior distribution of model parameters.

2.4 Methods of uncertainty analysis

UA method-type 1 assumes all sources of uncertainties in hydrological modeling can be accounted for by parameter uncertainty. Hence, the methods aim to find the most likely solutions of model parameters using a likelihood function and provide the uncertainty in parameter estimation. Examples are SUFI-2 (Abbaspour et al., 2007, 2004) and GLUE (Beven and Binley, 1992). In SUFI-2, the parameter uncertainty is described by a multivariate uniform distribution in a parameter hypercube and the model parameters are calibrated to bracket most of the measured data within the 95% prediction uncertainty band (Abbaspour et al., 2007). The output uncertainty is quantified by the
95% prediction uncertainty (95PPU) calculated at the 2.5% and 97.5% levels of the cumulative distribution of an output variable obtained through Latin Hypercube Sampling (Abbaspour et al., 2007). Two indices, the P-factor and the R-factor, are used to quantify the goodness of calibration/uncertainty performance. The P-factor is the percentage of data bracketed by the 95PPU band (maximum value 100%), and the R-factor is the average width of the band divided by the standard deviation of the corresponding measured variable. In ideal condition when the uncertainty model is perfect, P-factor will be 1 and the R-factor will be 0.

The GLUE approach is widely used for analyzing uncertainty in distributed hydrological modeling. The GLUE approach is known as an informal Bayesian method since it can be used with a statistically informal likelihood function (Vrugt et al., 2009). In GLUE, the parameter sets are randomly sampled from the prior distribution of parameters. All parameter sets meeting the predefined behavioral criterion are selected as behavioral parameter sets and a 'likelihood weight' is given to each behavioral parameter set. The prediction uncertainty is calculated by the percentiles of cumulative distribution realized from the weighted behavioral sets. The major drawback of the GLUE approach is its subjectivity of defining the likelihood function and the behavioral criterion (Blasone et al., 2008).

UA method-type 2 considers the model residuals as a combination of errors due to model inputs, parameters, model structure and outputs. In this approach, the residual errors are represented by an additive error model to the model outputs. The residual errors are described by a statistical model and the likelihood function is developed based on the assumptions of the statistical error model. In most of the cases, the assumptions of
statistical error models are violated resulting in biasedness in parameter estimation. This subsequently affects the parameter uncertainty and prediction uncertainty (Schoups and Vrugt, 2010; Thyer et al., 2009). In classical calibration method, the model residuals are described by the normal distribution with zero mean and constant variance and are assumed to be uncorrelated. When the residuals are correlated (Schoups and Vrugt, 2010; Laloy et al., 2010; Vrugt et al., 2009; Schaefli et al., 2007; Yang et al., 2007a,b; Bates and Campbell, 2001; Duan et al., 1988; Kuczera, 1983; Sorooshian and Dracup, 1980), the autoregressive (AR) models are used to remove the correlation of errors. When the errors are heteroscedastic (Laloy et al., 2010; Vrugt et al., 2009; Schaefli et al., 2007; Yang et al., 2007a,b; Bates and Campbell, 2001; Duan et al., 1988; Kuczera, 1983; Sorooshian and Dracup, 1980) and non-normal (McLeod et al., 1977; Box and Tiao, 1973), the Box-Cox transformation (Box and Cox, 1964) of data is used. Recently Schoups and Vrugt (2010) used an explicit statistical model to account for heteroscedasticity and non-normality of residuals instead of using the Box-Cox transformation of data. In their approach, the standard deviation of errors is modeled as a linear function of simulated response to account for the heteroscedasticity and the error distribution is described by considering the kurtosis and skewness of model residuals to account for the non-normality of residuals (Schoups and Vrugt, 2010). Sometimes Normal Quantile Transformation (NQT) is used to account for the non-normality of residuals along with the AR model (Engeland et al., 2010).

Sorooshian and Dracup (1980) used the first order autoregressive [AR(1)] scheme and the weighting approach with power transformation to reduce the correlation error and the heteroscedasticity of residuals, respectively, while Kuczera (1983) applied the
autoregressive moving average (ARMA) model and the power transformation for the similar problem. Duan et al. (1988) used the continuous time autoregressive error model during the hydrologic model calibration for the autocorrelation error of data recorded at unequal time interval. Bates and Campbell (2001) used the data transformation and the higher order autoregressive model to remove the problems of non-constant variance and autocorrelation of errors. They reported that the likelihood function based on data transformation might not lead to independent, normally distributed residuals with zero mean and constant variance. Schaeffli et al (2007) used the AR(1) model to account for the correlated errors and used a mixture of normal distribution error model with two mixture components of high flow and low flow to account for the non-normality of errors. Yang et al. (2007a) used the data transformation and continuous time autoregressive error model to account for the heteroscedasticity and correlation of errors. They used the seasonal variation of the statistical error model parameters such as variance and characteristic correlation time to reduce the model structural uncertainty. Yang et al. (2007b) used the t distribution to describe the residual errors to account for the non-normality of residuals and continuous time autoregressive error model to account for the correlation of errors. Laloy et al. (2010) used the data transformation and AR(1) model to account for the non-constant variance and correlation of model residuals.

The major limitation of UA method-type 1 and UA method-type 2 is that the effect of different sources of errors on hydrologic model prediction cannot be separated using these methods. The effects of different sources of errors on model prediction cannot be separated unless each source of errors is considered explicitly. In the UA method-type 3, the input, model structure and output uncertainty are explicitly accounted
for and are represented by separate statistical error models (Renard et al., 2010; Huard and Mailhot, 2008). Renard et al. (2010) used the BATEA framework (Kavestki et al., 2006a) to identify the model input and structural errors. In BATEA, a multiplicative Gaussian input error model which is independent for each storm is introduced and the rainfall multipliers are considered as latent variables in the hierarchical Bayesian modeling. The input error model of BATEA corrects the systematic error of rainfall measurement on storm-event basis. Hence, the number of latent variables increases as the length of calibration period increases and BATEA becomes computationally intensive. BATEA has been used for quantifying prediction uncertainty in conceptual hydrological models (Renard et al., 2010; Thyer et al., 2009; Kavetski et al., 2006b).

Huard and Mailhot (2008) developed a Bayesian uncertainty analysis framework considering three errors of input, structural and output errors separately. There are two common ways to relate errors and data, additively and multiplicatively. To allow the use of Gaussian distributions, Huard and Mailhot (2008) applied additive error model. In this method, the model input and output time series are the latent variables to the Bayesian system and are inferred along with other model parameters. The approach is similar to the approach implicitly used by Vrugt et al. (2005) in simultaneous optimization and data assimilation (SODA) based on ensemble Kalman filters. The uncertainty framework is developed for monthly time series data and the dimensional problem of posterior probability distribution would arise if it is extended for daily time series.

Ajami et al. (2007) developed the Integrated Bayesian Uncertainty Estimator (IBUNE) framework to account for input, output and structural uncertainties in hydrologic rainfall-runoff predictions. For input uncertainty, the rain multiplier concept
(Kavetski et al., 2006a) has been implemented in IBUNE by using the mean and variance of the rainfall multipliers as latent variables in the Bayesian modeling. In IBUNE, the model structural uncertainty is accounted for by the Bayesian model combination approach. The IBUNE input error model reduces the high-dimensional problem of BATEA input error model. The IBUNE multipliers are not storm dependent and can be used in real-time forecasting to account for input error uncertainty (Ajami et al., 2009). Even though the IBUNE input error model reduces the dimension of the posterior distribution, Renard et al. (2009) reported some difficulties in implementing the IBUNE input error model under the Bayesian framework. According to Renard et al. (2009), "the likelihood and the posterior of IBUNE become random function of their arguments, which violates the fundamental requirement for probability density functions." The IBUNE input error model is also limited to be applicable for a relatively small variance of rainfall multipliers (Ajami et al., 2009). Vrugt et al. (2008) applied the storm multiplier concept (Kavetski et al., 2006a) for analyzing forcing data error explicitly in hydrologic model calibration. While applying the storm multiplier concept, Vrugt et al. (2008) assumed noninformative prior distribution of rainfall multipliers rather than informative prior. Kavetski et al. (2006a) recommended using the informative prior to avoid the ill-posedness in parameter inferences. Due to the high dimension of posterior distribution, Kavetski et al. (2006a) suggested using the Newton-type optimization methods and Hessian-based covariance analysis for solving the optimum parameter values. Vrugt et al. (2008) developed the Differential Evolution Adaptive Metropolis (DREAM) algorithm for solving the high-dimensional posterior probability distribution. The DREAM algorithm is an adaptation of the Shuffled Complex Evolution Metropolis (SCEM-UA)
(Vrugt et al., 2003), a global optimization algorithm. The DREAM algorithm maintains detailed balance and ergodicity and is more applicable for complex, highly nonlinear and multimodal target distributions (Vrugt et al., 2008). Vrugt et al. (2009) combined the AR(1) model with the rainfall multiplier model to account for the structural, input and parameter uncertainty.

Reichert and Mieleitner (2009) corrected the bias in hydrologic model or input data explicitly by considering stochastic, time-dependent parameters rather than considering bias in model outputs with AR error model. In this approach, the time-dependent parameters are used to correct the rainfall time series which is similar to the rainfall multipliers techniques (Kavetski et al., 2006a). This approach removes the heteroscedasticity of the residuals by applying data transformation and is applicable to nonlinear, dynamic models. Moradkhani et al. (2005) used the sequential data assimilation approach for estimating model parameters and state variables using Bayesian particle filters and observed improved uncertainty estimates of hydrological model parameters. Sequential data assimilation is a process where the system state is recursively estimated/corrected each time an observation becomes available (Moradkhani et al., 2005). Vrugt et al. (2005) introduced a simultaneous parameter optimization and data assimilation (SODA) method to assess the input, output, parameter and model structural uncertainties in hydrologic modeling. They combined the strengths of the parameter search efficiency and explorative capabilities of the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm (Vrugt et al., 2003) with the power and computational efficiency of the ensemble Kalman filter (Evensen, 1994). The main characteristic of SODA is to make the deterministic hydrologic model stochastic and combine parameter
with state estimation. In SODA, different sources of errors are accounted for in terms of state variables and the theoretical issues related to input errors are not focused (Huard and Mailhot, 2006). The difficulty of SODA is that it involves state estimation and increases the computational burden (Yang et al., 2007b).

2.5 Uncertainty analysis of SWAT model

SWAT is a commonly used distributed hydrological model for studying the effects of land use change, climate change and management practices on water resources system. Due to the extensive application of SWAT model, different techniques have been developed for its uncertainty analysis. A suite of tools called SWAT calibration and uncertainty programs (SWAT-CUP2) (Abbaspour, 2008) was developed for sensitivity analysis, calibration, validation and uncertainty analysis of SWAT model. In SWAT-CUP2, four uncertainty analyzing techniques are used for automated calibration and uncertainty analysis of SWAT model. These are SUFI-2 (Abbaspour et al., 2004, 2007), GLUE (Beven and Binley, 1992), Parameter Solution (Parasol) (Van Griensven and Meixner, 2006) and Metropolis-Hastings based MCMC method. The method of automated calibration and uncertainty analysis of highly parameterized SWAT model has become convenient to apply due to the development of 'aggregate parameter concept' (Yang et al., 2005). Some examples of applying the 'aggregate parameter concept' to SWAT model are the research works of Li et al. (2010) and Yang et al. (2008; 2007a,b).

Setegn et al. (2010) used SUFI-2, Parasol and GLUE for estimating prediction uncertainty of SWAT model for the Lake Tana Basin, Ethiopia. Li et al. (2009) used the Bayesian MCMC approach for parameter estimation and uncertainty analysis of SWAT model for the upper reaches of the Heihe River Basin in China and observed relatively
small contributions of parameter uncertainty on model simulation uncertainty. Li et al. (2010) also used the bootstrap method (Stine, 1985) for analyzing parameter uncertainty of SWAT model in Yingluoxia watershed in northwest China. Ghaffari et al. (2010) applied SWAT model for studying the impacts of land-use changes on hydrology of Zanjanrood Basin, northwest Iran and used SUFI-2 for analyzing the uncertainty of SWAT model prediction. Faramarzi et al. (2009) and Schuol et al. (2008) also used SUFI-2 for analyzing uncertainty of blue and green water resources availability in Iran and Africa, respectively using SWAT model. Xie and Zhang (2010) applied the sequential data assimilation technique, the ensemble Kalman filter (EnKF) for combined state-parameter estimation of SWAT model. Zhang et al. (2009) used the combined method of Genetic Algorithms (GA) and Bayesian Model Averaging (BMA) for calibration and uncertainty analysis of SWAT model. Yang et al. (2008) applied five different uncertainty analysis techniques to the SWAT model; GLUE, Parasol, SUFI-2, and a Bayesian framework implemented using MCMC and Importance Sampling (IS) techniques. Yang et al. (2007a,b) used the continuous time AR models with Box-Cox transformation of data for uncertainty analysis of SWAT model.

Since SWAT is a distributed model, the uncertainty in model prediction may arise from the methods of distribution of rainfall inputs as well as from the spatial scale of sub-watershed delineation. Cho et al. (2009) studied the effects of spatial distribution of rainfall and the effects of sub-watershed delineation on the temporal and spatial uncertainties of streamflow prediction and water quality results generated by SWAT model. Kumar and Merwade (2009) studied the effects of sub-watershed delineation and soil data resolution on calibration and parameter uncertainty of SWAT model.
The literature shows that the physically based distributed model SWAT has been extensively used for watershed management in different climatic and hydrologic conditions. Different uncertainty analyzing methods are adopted for quantifying parameter uncertainty and prediction uncertainty of SWAT model. Most of the methods fall under category UA method-type 1. Explicit methods of analyzing uncertainty such as BATEA (Kavetski et al., 2006a), IBUNE (Ajami et al., 2007), uncertainty framework of Huard and Mailhot (2008) etc., have not yet been adopted for analyzing uncertainty of the widely used SWAT model. The explicit methods assume specific error model for a particular source of errors. For example, BATEA uses the multiplicative input error model to account for rainfall uncertainty in hydrological modeling. The input error model of BATEA assumes inputs as a random variable. Application of such input error models for assessing uncertainty in SWAT model prediction is a challenging task due to the use of a large number of variables to describe the hydrologic system (Abbaspour, 2008).

2.6 Current state of knowledge

To improve model parameter estimation and reducing parameter and prediction uncertainties, many powerful numerical simulation and optimization tools, such as Shuffled Complex Evolutionary Metropolis algorithm (SCEM-UA) (Vrugt et al., 2003), Differential Evolution Adaptive Metropolis (DREAM) (Vrugt et al., 2008), etc., have been developed. These tools are being efficiently used for MCMC simulation and uncertainty analysis. The recent research studies have been carried out for improving the efficiency of optimization tools (Chu et al., 2010) and increasing the efficiency of Markov Chain Monte Carlo (MCMC) sampler (Kuczera at el., 2010). Moreover, for reducing the computational cost of uncertainty-based calibration method, Razavi et al.
(2010) developed the 'model preemption' concept where a simulation model is terminated early if the current model parameter set does not benefit the parameter searching scheme by looking at the intermediate simulation model results.

2.7 Summary

Quantification of parameter and prediction uncertainty has been practiced for the last three decades in hydrological modeling. Despite an extensive improvement in the area of uncertainty in hydrological modeling, there are some research gaps. In distributed hydrological modeling, the effects of different sources of errors on parameter estimation and model prediction have not yet been quantified separately. Due to the difficulties in implementing explicit methods, they are not commonly used for uncertainty analysis of the distributed models. Moreover, the applicability of the multiplicative input error models has not yet been explored for quantifying the input uncertainty in the distributed hydrological modeling. The existing input error models have some dimensional and computational problems when they are applied to the highly parameterized distributed model for a long calibration period. The present study aims to develop a new uncertainty analysis method suitable for a distributed hydrologic model. The SWAT model has been selected as a tool for evaluating the performance of the developed methodology. The study expects to reduce the existing research gaps of the uncertainty analysis by introducing a season-dependent input error model for quantifying precipitation uncertainty in distributed hydrological modeling.
CHAPTER III
METHODOLOGY

3.1 Introduction

To address the objectives of this dissertation, the SWAT model (Arnold et al., 1998) is calibrated by an automated calibration procedure using different uncertainty analyzing frameworks. The uncertainty frameworks include the traditional method, implicit methods and explicit methods of uncertainty analysis. In the traditional method, the uncertainty in model parameters is considered only during the calibration process. In the implicit methods, the appropriate AR models are used in the likelihood function to account for the model input uncertainty in aggregation with other sources of uncertainty. In the explicit methods, input uncertainty is accounted for by some input error models in the calibration process while the uncertainty in model structure and observed outputs is not considered explicitly.

The automated calibration of SWAT model under any uncertainty framework is based on Bayesian approach and the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm (Vrugt et al., 2003), a MCMC based calibration and optimization tool is used for solving the posterior probability distribution. The SCEM-UA algorithm is based on the Metropolis-Hastings algorithm (Hastings, 1970; Metropolis et al., 1953) and complex shuffling procedure for sampling the model parameters (Vrugt et al., 2003) and finds the target posterior probability density function and the global optimum values of model parameters. In general, the MCMC is an approach to sample parameters from an approximate distribution and then correct the samples to better approximate the target posterior probability density function (Gelman et al., 2004). A Markov Chain is a
sequence of random samples for which at any iteration, \( t \), the distribution of parameters given all previous samples depend on the most recent value of parameters. A Markov Chain is generated by sampling \( \theta_{(t+1)} \approx z(\theta_{(t)}|\theta_{(t)}) \), where \( z \) is called the proposal distribution of the Markov Chain. The main feature of the Markov Chain is to create a Markov process whose stationary distribution is the target posterior probability distribution of parameters. Therefore, the simulation is run long enough so that the Markov Chain converges to the stationary posterior probability distribution (Gelman et al., 2004). The convergence depends on the shape and size of the proposal distribution \( z(\cdot) \) (Laloy et al., 2010).

The general methodology of analyzing uncertainty of SWAT model parameters and model prediction is presented by a flowchart in Figure 3.1. To make parameter inferences using the MCMC sampler, the likelihood function is developed for each uncertainty analyzing frameworks considering the respective error models. In the calibration process, the closeness of fit between the model predictions and the observed data are described by the likelihood function, which represents the modeling errors via a stochastic model (Engeland et al., 2005). The likelihood function mainly controls the estimation of model parameters (Engeland et al., 2005; Boyle et al., 2001). After the convergence of the Markov Chains, the posterior probability density function of model parameters is analyzed and the uncertainty in model parameters is quantified. The model parameters obtained at the maximum of the posterior density, known as the 'optimum parameter', is also recorded to check the closeness of fit between the observed data and simulated values. In the case of input error model, the parameters of input error models are sampled with the SWAT model parameters and the input uncertainty is quantified.
Figure 3.1: The flow chart for calibration of SWAT model under any uncertainty framework

1. Develop the appropriate likelihood function
2. Prepare input data and selection of calibration parameters
3. Make parameter inferences using the likelihood function by the SCEM algorithm
4. Check the convergence of Markov Chains
5. Analyze the posterior probability distribution of model parameters using the samples after the convergence of Markov Chains
6. Estimate parameter uncertainty, prediction uncertainty and input uncertainty in the model calibration period
7. Estimate prediction uncertainty in the validation period
8. Carry out the posterior diagnostics of residual errors and input error models
with the samples after the posterior probability distribution reaches the stationary
distribution. The parameter uncertainty and the input uncertainty quantified in the
calibration period are propagated in SWAT model simulation during the model validation
period and the prediction uncertainty is quantified. Finally, the assumptions of the
residual error models and input error models are tested using the standard tools of
verification. In the subsequent sections of this chapter, the structure of SWAT model, the
calibration parameters, the computational framework and the experimental design for
carrying out the objectives of this dissertation are briefly described. At the end of this
chapter, some statistical tests and graphical plots that have been used in this research for
carrying out posterior diagnostics are briefly described.

3.2 SWAT model

SWAT (Arnold et al., 1998) is a public domain distributed hydrologic model. SWAT has an interface with ArcGIS called ArcSWAT. The ArcGIS is an integrated
collection of Geographic Information System (GIS) software for spatial analysis, data
management and mapping (ESRI Canada website: https://www.esricanada.com). The
ArcSWAT has the capabilities of preprocessing, interface and post-processing of SWAT
data and output. The ArcSWAT divides the watershed into a number of sub-basins and
extracts model input data from the map layers and other databases for each sub-basin.
Overlying the land use and soil maps on the Digital Elevation Model (DEM) map, each
sub-basin is divided into a number of Hydrologic Response Units (HRUs) and the SWAT
model simulates water, sediment, nutrients and pesticides transport at a HRU level on a
daily basis. This dissertation presents the generation of runoff at HRU level and
transportation of water from the HRUs to the watershed outlet. The movement of water at
the HRU level simulated by the SWAT model for the watershed selected in this study is shown in Figure 3.2. Infiltration and surface runoff from daily precipitation are calculated in SWAT by the SCS curve number method (Soil Conservation Services, 1972). The potential evapotranspiration is estimated for the watershed by the Penman-Monteith method (Monteith, 1965). Lateral subsurface flow is computed using the kinematic storage model (Sloan and Moore, 1984) and groundwater flow is computed as return flow to stream from the shallow aquifer (Arnold et al., 1998). For routing the channel water, the Muskingum method (Cunge, 1969) is used.

3.3 Selection of calibration parameters

SWAT is a highly parameterized distributed model. For the calibration of SWAT model, the model parameters are aggregated to reduce the number of parameters needed to be calibrated. The ‘aggregate parameter’ concept was developed by Yang et al. (2005) and is expressed in the following format (Abbaspour, 2008):

\[ \hat{x} _{parname}_{ext}_{hydrogrp}_{soltext}_{landuse}_{subbsn}_{slope} \]  

(2.1)

where \( \hat{x} \) indicates the type of change to be applied to the parameter (such as, \( v \) means existing parameter will be replaced by the given value, \( a \) means the given value will be added to the existing parameter value, \( r \) means the relative change to the existing parameter value). \( \{parname\} \) is the SWAT parameter name, \( \{ext\} \) indicates the extension of the SWAT input file which contains the parameter needed to be changed, \( \{hydrogrp\} \) represents the soil hydrologic group, \( \{soltext\} \) means the type of soil texture, \( \{landuse\} \) indicates the name of the land use type, \( \{subbsn\} \) indicates the sub-basin number and \( \{slope\} \) indicates slope of the HRU.
Rahman (2007) carried out a sensitivity analysis of SWAT model parameters for simulating streamflow in the Canard River watershed which has been selected as the study area in this dissertation. In the sensitivity analysis, one parameter was changed at a time by ± 10 percent of its initial value and its effects on annual streamflow was quantified. Four model parameters i.e., curve number (CN), available water holding capacity (AWC), the plant uptake compensation factor (EPCO) and soil evaporation compensation factor (ESCO) were observed to be the most sensitive parameters for the

---

Figure 3.2: Movement of water simulated by SWAT at the HRU level for the study area (Adapted from Neitsch et al., 2005)
Canard River watershed modeling. The parameters related to snow hydrology such as, maximum melt rate for snow (SMFMN) and minimum melt rate for snow (SMFMX) showed low sensitivity for simulating annual streamflow. Furthermore, the parameters related to groundwater flow, such as baseflow alpha factor (ALPHA_BF) and groundwater delay time (GW_DELAY) showed low sensitivity. Therefore, four aggregate model parameters, such as, CN, AWC, EPCO and ESCO are primarily selected in this study as the calibration parameters. However, the sensitivity of SWAT model parameters related to snow hydrology and groundwater flow in the context of aggregate model parameters are also analyzed by including these model parameters in automated calibration processes. These parameters are observed to be low sensitive to streamflow simulation. Therefore, four model parameters: CN, AWC, EPCO and ESCO are estimated through the calibration procedure for the study area. The parameter, CN is very sensitive for the estimation of surface runoff, while the parameter AWC is very sensitive for estimating the soil storage and evapotranspiration. The value of CN varies non-linearly with the moisture content of soil. CN drops when the moisture content of soil approaches the wilting point and when soil approaches saturation, CN may increase to near 100 (Neitsch et al., 2005). The parameter EPCO indicates the changes in the depth distribution of soil layers used to meet water uptake demand of plant and the parameter ESCO indicates the changes in the depth distribution of soil layers to meet the soil evaporative demand. The values of EPCO and ESCO can range from 0.01 to 1.0. The value of EPCO near one means that the water uptake demand will be met by the lower layers of soil. The value of EPCO close to zero means that the model allows less variation from the original depth distribution to take place (Neitsch et al., 2005). As the
value of ESCO reduces, the model will extract more of the evaporative demand from lower levels (Neitsch et al., 2005). Evapotranspiration is the largest component of the water balance in the study area. Moreover, the seasonal occurrence of evapotranspiration is highly variable in the selected watersheds. Therefore, evapotranspiration related model parameters, such as, EPCO and ESCO are observed to be highly sensitive to streamflow simulation. In the current research, the aggregate parameters are expressed as a-CN2.mgt, a-SOL_AWC (.sol, a-EPCO.bsn and a-ESCO.bsn, respectively. The variable ‘a’ indicates that the value is added to the existing values of those parameters. The symbol ’( )’ indicates that the value will be changed to all of the layers of soil. The mgt, sol and bsn indicate the extension of data files that contain the parameters needed to be calibrated. The calibration process is sensitive to the type of changes applied to the model parameters. After some sensitivity analyses, addition to the existing values is selected as the type of parameter changes for automated calibration of SWAT model so that the values of estimated model parameters would represent the hydrologic system of the study area. This type of changes to the existing values was employed in the research works of Li et al. (2010) and Yang et al. (2007a).

3.4 Computational framework

To analyze the posterior probability distribution, the SWAT model is simulated at the predefined number of iterations and the posterior distribution are analyzed using the samples after the convergence of the Markov Chains. The computational time increases with the increase in the dimension of posterior probability distribution. Five parallel Markov Chains are used for Monte Carlo simulations and the convergence of the chain is checked using the convergence diagnostic called the scale reduction factor (\( \hat{R} \)) (Gelman
and Rubin, 1992). A value of \( \hat{R} \) close to 1 for each of the parameters indicates convergence of the chain. Since it is difficult to achieve the value of unity, Gelman and Rubin (1992) recommended using the value of \( \hat{R} \) less than 1.2 as an indicator of convergence of the Markov Chain to a stationary distribution. The equation of \( \hat{R} \) is given as (Gelman and Rubin, 1992):

\[
\hat{R} = \sqrt{\frac{g - 1}{g} + \frac{q + 1}{g} \frac{B}{W}}
\]

(2.2)

where \( g \) is the number of iterations within each sequence, \( B \) is the variance between the \( q \) sequence means, and \( W \) is the average of the \( q \) within-sequence variances for the parameter under consideration. The product of \( q \) and \( g \) is identical to the total number of iterations.

The MCMC simulations are performed under the GNU OCTAVE environment. The GNU OCTAVE is a publicly available high-level language and mostly compatible with MATLAB. In each iteration, the model parameters are sampled by the SCEM-UA algorithm (Vrugt et al., 2003) from the pre-specified distribution and the SWAT model is simulated with the sampled parameters and the text format input data files for each of the HRU. Once the data files are extracted in text format by the ArcSWAT interface, it can be used outside the GIS environment for SWAT model simulation by running the SWAT executable file. The text format input data files are kept in the 'Backup' directory and the SCEM-UA algorithm (Vrugt et al., 2003) is used for parameter inferences. The calibration framework using the SCEM-UA algorithm (Vrugt et al., 2003) is shown in Figure 3.3. The computational flowchart except the inclusion of the input error model is similar to the SWAT calibration and uncertainty programs (SWAT-CUP) developed by
Figure 3.3: The computational framework of SWAT model calibration considering the input error model
(The framework outside the dotted line is similar to Abbaspour, 2008)
Abbaspour (2008). The executable file SWAT_edit uses the input data files kept in the 'Backup' directory and makes changes to the model parameters as directed in the 'model.in' file. The changed data files are kept in the current working directory and the model outputs are generated by using the changed data files with the SWAT2005 executable program. The executable file SWAT_extract extracts data from SWAT output files and saves in the 'model.out' file in the current working directory. After extracting the model outputs, the density function value is estimated by the likelihood function. To incorporate the input error models directly in SWAT model calibration, a computer program called 'Multiplier' has been developed to make necessary changes to the precipitation input data file. The executable files SWAT_edit, SWAT2005, SWAT_extract and Multiplier are called from OCTAVE for the model simulation. The codes of SCEM-UA algorithm (Vrugt et al., 2003) are written in MATLAB and also compatible with OCTAVE.

3.5 Experimental design

To carry out the research objectives of this study, the SWAT model is calibrated by the traditional method of calibration, six implicit methods and three explicit methods of uncertainty analysis. The traditional method of calibration is described here as the Standard method. Three implicit methods are carried out by implementing the first order autoregressive model [AR(1) model], second order autoregressive model [AR(2) model] and continuous time autoregressive model (ARcont model) separately in the likelihood function. The remaining three implicit methods include the Box-Cox transformation (Box and Cox, 1964) of data along with the AR models to account for the non-homogeneity and non-normality of model residuals. The justification of using the AR models are
described in the following chapter. The likelihood function considering the AR(2) model is developed in this dissertation to account for the lumping errors from model inputs, parameters and structure. The explicit methods of uncertainty analysis are carried out by implementing three input error models in the calibration process. These are the seasonal input error model, the daily input error model developed by Ajami et al. (2007) and the original storm-event basis input error model developed by Kavetski et al. (2006a). The seasonal input error model is developed and the likelihood function considering the seasonal input error model is formulated in this dissertation. The details of the likelihood functions used for different calibration methods are described in the following chapters. A summary of the experimental design is presented in Table 3.1.

3.6 Posterior diagnostics

In this dissertation, Kruskal-Wallis test or H-test is used to verify the homoscedasticity of residuals errors. The Kruskal-Wallis test or H-test verifies the null hypothesis that \( k \) independent random samples are from the identical populations. The form of \( H \)- statistic is given by :

\[
H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)
\]  

(2.3)

where \( R_i \) is the sum of the ranks of \( n_i \) observations of the \( i \)th sample and \( n_1 + n_2 + n_3 + \ldots + n_k = n \). The null hypothesis is rejected at a significance level \( \alpha \) if the computed \( H \) is larger than \( \chi^2_{1-\alpha,k-1} \).

In this study, the series of residuals are divided into four groups (i.e \( k = 4 \)) based on the long term average of computed discharge. The groups are defined as follows:

Group (a): if the computed discharge is less than 50% of long term average discharge
Group (b): if the computed discharge is between 50% to 75% of long term average
Group (c): if the computed discharge is between 75% to 125% of long term average
Group (d): if the computed discharge is greater than 125% of long term average

Further, for quantifying the reliability of model prediction, the predictive Quantile-Quantile (QQ) plot are used in this dissertation. The predictive QQ plot (Laio and Tamea, 2007) is a useful tool to verify the probabilistic forecasts of hydrological variables. The details of the construction of predictive QQ plot are described in Laio and Tamea (2007) and Thyer et al. (2009). The predictive QQ plot helps to quantify the reliability of the streamflow prediction. If the predictive distribution of \( x_i \) is correct, the probability density function of \( x_i \) coincides with the true distribution of \( x_i \). If \( z_i \) represents the value from the cumulative distribution function of the predictions corresponding to the observed value of \( x_i \), the distribution of \( z_i \) is uniform, \( U[0,1] \) (Laio and Tamea, 2007). If the \( z \)-value curve is close to the bisector (the 1:1 line), the predictive distribution of \( x_i \) seems to be reliable, otherwise it indicates the biasness in prediction (Laio and Tamea, 2007). The deviation from the bisector can be quantified using the reliability index which is related to the area between the \( z \)-value curve and bisector line (Renard et al., 2010). If the area between the \( z \)-value curve and bisector line is close to zero, the value of the reliability index will be close to one. The value of reliability index close to 1 shows perfect reliability and the value of the index close to zero shows worst reliability of prediction (Renard et al., 2010).
<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Type of uncertainty analysis</th>
<th>Consideration of input error in model calibration</th>
<th>Notations used for the method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Traditional</td>
<td>No input error is considered</td>
<td>Standard</td>
</tr>
<tr>
<td>2</td>
<td>Implicit</td>
<td>Input errors are lumped with other errors and first order autoregressive model is used for describing the modeling errors</td>
<td>AR(1)_model</td>
</tr>
<tr>
<td>3</td>
<td>Implicit</td>
<td>Input errors are lumped with other errors and second order autoregressive model is used for describing the modeling errors</td>
<td>AR(2)_model</td>
</tr>
<tr>
<td>4</td>
<td>Implicit</td>
<td>Input errors are lumped with other errors and continuous time autoregressive model is used for describing the modeling errors</td>
<td>ARcont_model</td>
</tr>
<tr>
<td>5</td>
<td>Implicit</td>
<td>First order autoregressive model is used with Box-Cox transformation of data to describe the modeling errors</td>
<td>t_AR(1)_model</td>
</tr>
<tr>
<td>6</td>
<td>Implicit</td>
<td>Second order autoregressive model is used with Box-Cox transformation of data to describe the modeling errors</td>
<td>t_AR(2)_model</td>
</tr>
<tr>
<td>7</td>
<td>Implicit</td>
<td>Continuous time autoregressive model is used with Box-Cox transformation of data to describe the modeling errors</td>
<td>t_ARcont_model</td>
</tr>
<tr>
<td>8</td>
<td>Explicit</td>
<td>Seasonal input error model is used</td>
<td>Seasonal_input_error</td>
</tr>
<tr>
<td>9</td>
<td>Explicit</td>
<td>Daily input error model is used</td>
<td>Daily_input_error</td>
</tr>
<tr>
<td>10</td>
<td>Explicit</td>
<td>Storm-event basis input error model is used</td>
<td>Storm_input_error</td>
</tr>
</tbody>
</table>
CHAPTER IV
IMPLICIT METHODS OF UNCERTAINTY ANALYSIS

4.1 Introduction

In the implicit methods of uncertainty analysis, input uncertainty is lumped together with other sources of uncertainty in the hydrological model calibration and the errors are expressed as an additive error model to the outputs. In the Bayesian approach of uncertainty analysis, the estimation of the posterior distribution of model parameters is dominated by the likelihood function (Smith et al., 2010; Box et al., 2008). Therefore, the appropriate form of likelihood function in the Bayesian inferences is equally important as in the frequentist inferences. While formulating the likelihood function in the implicit methods, a white noise model is generally applied assuming that the errors are uncorrelated and normally distributed with zero mean and a constant variance (homoscedastic). The assumptions of white noise are not often satisfied in model calibration and parameter optimization processes due to the presence of different sources of uncertainty in hydrological modeling. Since the Bayesian theory is not limited to the assumption of normal distribution of the errors, little attention is paid to verify the distributional assumption in Bayesian approach based calibration methods. However, without reasonable description of the modeling errors, the form of likelihood function can be inadequate for searching model parameter values. This may result in inefficient parameter estimates (Kuczera, 1983) and erroneous assessment of parameter uncertainty (Yang et al., 2007a; Kuczera, 1983). The unrealistic assessment of parameter uncertainty may cause unreliable model prediction uncertainty (Li et al., 2010; Yang et al., 2007a; Kuczera and Parent, 1998). Therefore, it is of utmost importance to select the appropriate
statistical error models to formulate the likelihood function and to test the adequacy of
the model.

The problem of autocorrelation of residuals may be caused by the errors in model
input and model structure (Reichert and Mieleitner, 2009). The problem of
heteroscedasticity may arise when the hydrologic data used for calibration have non-
stationary properties (Sorooshian and Dracup, 1980). The AR models are adopted with
appropriate order to describe the correlated errors (Schoups and Vrugt, 2010; Laloy et al.,
2010; Vrugt et al., 2009; Schaefl et al., 2007; Yang et al., 2007a,b; Bates and Campbell,
2001; Duan et al., 1988; Kuczera, 1983; Sorooshian and Dracup, 1980). The order of AR
models is often identified by plotting the autocorrelation function (ACF) and partial
autocorrelation function (PACF) of errors. The AR models may be discrete (Laloy et al.,
2010; Vrugt et al., 2009; Schaefl et al., 2007; Bates and Campbell, 2001; Kuczera, 1983;
Sorooshian and Dracup, 1980) or continuous (Yang et al., 2007a,b; Duan et al., 1988)
depending on the characteristics of modeling errors. The AR models with first order
[AR(1) model] is most commonly used to account for the correlated errors in the
hydrological modeling (Laloy et al., 2010; Vrugt et al., 2009; Schaefl et al., 2007; Bates
and Campbell, 2001; Kuczera, 1983; Sorooshian and Dracup, 1980). Therefore, the
mathematical formulation of the likelihood function based on the AR(1) model is
available in the literature. The AR models with second order [AR(2) model] are seldom
used in the likelihood function for parameter inferences. However, it is often used in time
series modeling of hydrological data (McLeod et al., 1977; Delleur et al., 1976). Hence,
the functional form of the likelihood considering AR(2) process is not readily available in
the literature. In this dissertation, the likelihood function considering AR(2) model is
developed and used for parameter inferences. If the errors are heteroscedastic, the Box-Cox transformation of data is used to make the variance of the errors to be constant (Laloy et al., 2010; Vrugt et al., 2009; Schaefli et al., 2007; Yang et al., 2007a,b; Bates and Campbell, 2001; Duan et al., 1988; Kuczera, 1983; Sorooshian and Dracup, 1980). The Box-Cox transformation (Box and Cox, 1964) of data is often used to reduce the non-normality of the errors (McLeod et al., 1977; Box and Tiao, 1973).

In this dissertation, the implicit methods are used in the uncertainty analysis of SWAT model for the Canard River watershed located in Southwestern Ontario, Canada. Identifying that the modeling errors are correlated, the AR(1), AR(2) and continuous AR models are adopted in the likelihood function for parameter inferences. To reduce the heterogeneity and non-normality of errors, the Box-Cox transformation of data is applied in three calibration methods along with three different AR models.

In the subsequent section, the mathematical formulation of likelihood function with the white noise model assumption, AR(1) model and continuous time AR model are described briefly. The mathematical formulation of the likelihood function with the AR(2) model developed in this research is presented in section 4.3. These likelihood functions are used for parameter inferences and calibration of SWAT model. The results based on different likelihood functions are presented in section 4.4. The posterior diagnostic checks of residuals are carried out to verify the assumptions of the stochastic model and to conform the residual errors to the assumed stochastic error model (Thyer et al., 2009). The findings of this chapter are summarized and conclusions are drawn at the end of this chapter.
4.2 Uncertainty analysis using AR models

4.2.1 Likelihood function for white noise model

Any hydrological model can be represented mathematically by the following equation:

\[ q_t = g_t([X], \theta) \quad t = 1, 2, \ldots, n \] (4.1)

where \( g_t(\cdot) \) is the response function in the hydrologic system, \( q_t \) is true response, such as streamflow, groundwater level etc. of the hydrologic system at time step \( t \), \( n \) is the number of time steps, \( [X] = \{x_1, x_2, x_3, \ldots, x_n\} \) is a vector of input data, such as precipitation, temperature etc. to the hydrologic system and \( \{\theta\} = \{\theta_1, \theta_2, \theta_3, \ldots, \theta_s\} \) is a vector containing the \( s \) model parameters that need to be estimated through the calibration process.

Considering errors in the modeling results, eqn. (4.1) can be modified as follows:

\[ q_t = f_t([X], \theta) + e_t \quad t = 1, 2, \ldots, n \] (4.2)

where, \( f_t(\cdot) \) is the selected hydrologic model for the watershed response and \( e_t \) represents the modeling errors that may arise from measurement errors in the calibration data, model inputs and model structural errors (Duan et al., 1988; Kuczera, 1983). The computed response vector of the model for a given parameter vector \( \{\theta\} \) can be represented as \( \{Q\}_{com} = \{\hat{q}_1, \hat{q}_2, \hat{q}_3, \ldots, \hat{q}_n\}^T \) while the observed response of the system and the error vector can be represented as \( \{Q\}_{obs} = \{q_1, q_2, q_3, \ldots, q_n\}^T \) and \( \{E\} = \{e_1, e_2, e_3, \ldots, e_n\}^T \), respectively. The superscript \( T \) represents the transpose of the vector.
Using the Bayesian theory, the posterior probability density function \((pdf)\) of model parameters conditioned on observed streamflow data \(\{Q\}_{obs}\) can be written as:

\[
p(\theta | \{Q\}_{obs}) \propto p(\theta) p(\{Q\}_{obs} | \theta)
\]

where \(p(\theta)\) is the prior pdf of model parameters, which represents the prior knowledge about the parameters \(\{\theta\}\), and \(p(\theta | \{Q\}_{obs})\) is the posterior pdf of parameters, \(\{\theta\}\). The observed data \(\{Q\}_{obs}\) affects the posterior inference through the function \(p(\{Q\}_{obs} | \theta)\), and for given data \(\{Q\}_{obs}\), the function is the likelihood function of \(\{\theta\}\) (Gelman et al., 2004).

Since the white noise model assumes that the errors are uncorrelated and the probability distribution of errors are normal with zero mean and constant variance, the likelihood function of \(\{\theta\}\) can be written as (Laloy et al., 2010; Vrugt et al., 2009; Feyen et al., 2007):

\[
I(\{\theta\}, \{Q\}_{obs}) = (2\pi)^{-\frac{n}{2}} \left(\sigma_e^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \left(\sigma_e^2 \sum_{t=1}^{n} (e_t)^2\right)\right]
\]

(4.4)

where the errors are defined as the difference between observed and computed streamflow and can be expressed as \(\{e_t = (q_t - \hat{q}_t)\} \quad t = 1,2,\ldots,n\). The variance of errors are expressed as \(\sigma_e^2\).

4.2.2 Box-Cox transformation of data

To account for the non-homogeneity and non-normality of modeling errors, the following Box-Cox transformation of data is applied to the simulated and the measured data:
\[ q_* = \begin{cases} \frac{q^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln(q) & \text{if } \lambda = 0 \end{cases} \]  

and

\[ q_* = \begin{cases} \frac{q^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln(q) & \text{if } \lambda = 0 \end{cases} \]  

where \( \hat{q}_* \) is transformed observed data and \( q_* \) is transformed simulated data. The computed response vector of the model and the observed response of the system in the transformed space can be represented as:

\[ \{Q_*\}_{\text{com}} = \{\hat{q}_{s1}, \hat{q}_{s2}, \ldots, \hat{q}_{sn}\}^T \] and \[ \{Q_*\}_{\text{obs}} = \{q_{s1}, q_{s2}, q_{s3}, \ldots, q_{sn}\}^T \]

respectively.

The errors in the transformed space then become:

\[ q_* = \hat{q}_* + e_* \quad t = 1, 2, \ldots, n \]  

4.2.3 Likelihood function for AR(1) model

In the AR(1) process without any data transformation, the errors are expressed as:

\[ e_t = \phi_1 e_{t-1} + \nu_t \quad t = 1, 2, \ldots, n \]  

where \( \phi_1 \) is the autoregressive model parameter for the errors and \( \{\nu\} = \{\nu_1, \nu_2, \ldots, \nu_n\}^T \) is a vector of random components. The errors \( \nu_t \) represent the unexplained errors of the stochastic model. Assuming \( e_0 \) at \( t = 0 \) as zero and that the errors, \( \nu_t \), are normally distributed with zero mean and constant variance \( \sigma^2 \) for all \( t > 1 \), the likelihood function for estimating the model parameters \( \{\theta\} \) and autoregressive
parameter \( \phi_1 \) can be constructed as follows (Vrugt et al., 2009; Sorooshian and Dracup, 1980):

\[
l(\theta, \phi_1, \{Q\}_{obs}) = (2\pi)^{-\frac{n}{2}} \left( \frac{\sigma_v^2}{1-\phi_1^2} \right)^{\frac{n}{2}} \exp \left[ -\frac{1}{2} \sigma_v^{-2} \left( 1 - \phi_1^2 \right) e_1^2 + \left( \sum_{t=2}^{n} \left( v_t \right)^2 \right) \right]
\]

(4.9)

The AR(1) model for the transformed errors can be expressed as:

\[
e_{st} = \phi_{st} e_{s(t-1)} + v_{st} \quad t = 1, 2, \ldots, n
\]

(4.10)

where \( \phi_{st} \) is the autoregressive model parameter for the errors and \( \{V_s\} = \{v_{s1}, v_{s2}, v_{s3}, \ldots, v_{sn}\}^T \) is a vector of random components.

Assuming \( e_{s0} \) as zero and the errors \( v_{st} \) are normally distributed with zero mean and constant variance \( \sigma_v^2 \) for all \( t > 1 \), the likelihood function for estimating the model parameters \( \{\theta\} \), the transformation parameter \( \lambda \) and the autoregressive parameter \( \phi_{s1} \) can be constructed as follows:

\[
l(\theta, \lambda, \phi_{s1}, \{Q\}_{obs}) = (2\pi)^{-\frac{n}{2}} \left( \frac{\sigma_v^2}{1-\phi_{s1}^2} \right)^{\frac{n}{2}} \exp \left[ -\frac{1}{2} \sigma_v^{-2} \left( 1 - \phi_{s1}^2 \right) e_{s1}^2 + \left( \sum_{t=2}^{n} \left( v_{st} \right)^2 \right) \right]
\]

(4.11)

4.2.4 Likelihood function for continuous AR model

The continuous time series models are often used to account for the autocorrelation of residuals. Duan et al. (1988) modeled the errors with unequal time interval using the continuous time AR process where the correlation between the errors are increased with the closely spaced data and decreased with widely spaced data. Yang et al. (2007a,b) accounted for the correlated errors using the continuous time AR model considering the seasonally variable stochastic model parameters, such as characteristic correlation time.
and variance of errors. The continuous time AR model is closely related to the depletion of reservoir storage with time (Duan et al., 1988) and it can be anticipated that the continuous time AR model may represent the correlation errors better than the discrete time AR models.

The continuous time AR model can be represented as (Yang et al., 2007a):

\[ e_t = e_{t-1} \exp \left\{ -\frac{U_t - U_{t-1}}{\tau} \right\} + a_t \quad t = 1, 2, \ldots, n \quad (4.12) \]

where \( (U_t - U_{t-1}) \) is the difference between successive time steps, \( \tau \) is called the characteristic correlation time and \( \{A\} = \{a_1, a_2, a_3, \ldots, a_n\}^T \) is a vector of random components. Assuming \( e_0 \) as zero and the errors \( a_t \) are normally distributed with zero mean and constant variance \( \sigma_a^2 \), the likelihood function for estimating the model parameters \( \{\theta\} \) and characteristic correlation time \( \tau \) can be constructed as follows (Duan et al., 1988):

\[
l(\theta, \tau, \{Q\}_{obs}) = (2\pi)^{-\frac{n}{2}} \left( \sigma_a^2 \right)^{-\frac{n}{2}} \exp \left[ -\frac{1}{2} \sigma_a^2 \sum_{t=1}^{n} \left( e_t - e_{t-1} \exp \left( -\frac{U_t - U_{t-1}}{\tau} \right) \right)^2 \right] \quad (4.13) \]

In the transformed space, the continuous time AR model can be written as:

\[ e_{s_t} = e_{s(t-1)} \exp \left\{ -\frac{U_{s_t} - U_{s_{t-1}}}{\tau_s} \right\} + a_{s_t} \quad t = 1, 2, \ldots, n \quad (4.14) \]

where \( \tau_s \) is the characteristic correlation time and \( \{A_s\} = \{a_{s_1}, a_{s_2}, a_{s_3}, \ldots, a_{s_n}\}^T \) is a vector of random components in transformed space. Assuming \( e_{s0} \) as zero and the errors \( a_{s_t} \) are normally distributed with zero mean and constant variance, \( \sigma_{s_a}^2 \), the likelihood...
function for estimating the model parameters $\{\theta\}$ and characteristic correlation time $\tau_*$ can be constructed as follows:

$$l(\theta, \tau_*, \{Q_*\}_{obs}) = (2\pi)^{-\frac{n}{2}} (\sigma_{*a}^2)^{\frac{n}{2}} \exp \left[ -\frac{1}{2} \sigma_{*a}^{-2} \sum_{t=1}^{n} \left( e_{*t} - e_{*(t-1)} \exp \left( -\frac{T_t - T_{t-1}}{\tau_*} \right) \right)^2 \right]$$  \hspace{1cm} (4.15)

### 4.3 Formulation of likelihood function with AR(2) model

In this dissertation, the mathematical formulation of the likelihood function is developed to account for the correlation of the modeling errors. In the AR(2) process, the errors can be expressed as follows:

$$e_i = \phi_1 e_{i-1} + \phi_2 e_{i-2} + v_i \hspace{1cm} t = 1, 2, \ldots, n$$  \hspace{1cm} (4.16)

where $\phi_1$ and $\phi_2$ are the autoregressive parameters for the errors and $\{V\} = \{v_1, v_2, v_3, \ldots, v_n\}^T$ is a vector of random components. The errors are assumed to be independent and the distribution of the random error vector $\{V\}$ is assumed to be multivariate normal distribution with zero mean and covariance matrix $\Phi$. The covariance matrix, $\Phi$, it can be written as,

$$\Phi = E[VV^T] = \sigma_v^2 I$$  \hspace{1cm} (4.17)

where $\sigma_v^2$ is the variance of the random components for all $t>2$. It is assumed that $e_0$ at $t=0$ and $e_{-1}$ at $t=-1$ as zero.

If the errors are correlated, the covariance matrix of the error vector $\{E\}$ can be presented as follows (Siddiqui, 1958; Box et al., 2008):
\[
\Omega_n = \sigma_e^2 \begin{bmatrix}
1 & \rho_1 & \rho_2 & \ldots & \rho_{n-1} \\
\rho_1 & 1 & \rho_1 & \ldots & \rho_{n-2} \\
& \ddots & \ddots & \ddots & \ddots \\
\rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \ldots & 1
\end{bmatrix}
\] (4.18)

where \( \sigma_e^2 \) is the variance of the errors \( e_t \) for \( t = 1, 2, \ldots, j, \ldots, n \) and \( \rho_t \) is the autocorrelation between \( e_t \) and \( e_j \), which is defined as

\[
\rho_t = \frac{E e_t e_j}{\sigma_e^2}
\] (4.19)

Thus, in terms of errors, the covariance matrix \( \Omega_n \) can be written as follows (Judge et al., 1982):

\[
\Omega_n = E[EE^T] = E \begin{bmatrix}
e_1^2 & e_1e_2 & \ldots & e_1e_n \\
e_2e_1 & e_2^2 & \ldots & e_2e_n \\
& \ddots & \ddots & \ddots \\
e_ne_1 & e_ne_2 & \ldots & e_n^2
\end{bmatrix}
\] (4.20)

Now, the likelihood function of model \( \{\theta\} \) can be constructed as follows:

\[
l(\{\theta\}) = (2\pi)^{-n/2} |\Omega_n|^{-1/2} \exp \left[ -\frac{1}{2} (\{Q\}_\text{obs} - \{Q\}_\text{com})^T \Omega_n^{-1} (\{Q\}_\text{obs} - \{Q\}_\text{com}) \right]
\] (4.21)

In terms of error vector, eqn. (4.21) can be written as:

\[
l(\{\theta\}) = (2\pi)^{-n/2} |\Omega_n|^{-1/2} \exp \left[ -\frac{1}{2} (E\Omega_n^{-1}E)^T \right]
\] (4.22)

According to Siddiqui (1958), for the AR(2) process, the likelihood function [eqn. 4.22)] of the parameters \( \{\theta\} \) and autoregressive parameters \( \phi_1 \) and \( \phi_2 \) can be written as follows:
\[ l(\theta, \phi_1, \phi_2, \{Q\}_{\mathrm{obs}}) = (2\pi)^{-\frac{n}{2}}|\Omega_2|^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} \left( \frac{1}{\Sigma_\varepsilon^2} \left( \sum_{i=3}^{n} (v_i)^2 \right) \right) \right] \] (4.23)

According to Box et al. (2008), the inverse of the covariance matrix and the determinant of the inverse can be expressed as follows:

\[ \Omega_2^{-1} = \sigma_v^{-2} \begin{bmatrix} 1 - \phi_2^2 & -\phi_1 (1 + \phi_2) \\ -\phi_1 (1 + \phi_2) & 1 - \phi_2^2 \end{bmatrix} \] (4.24)

and

\[ |\Omega_2^{-1}| = \sigma_v^{-2} \left( (1 - \phi_2^2)^2 - (1 + \phi_2)^2 \phi_1^2 \right) \] (4.25)

Using eqns. (4.24) and (4.25) in eqn. (4.23), it can be written as:

\[ l(\theta, \phi_1, \phi_2, \{Q\}_{\mathrm{obs}}) = (2\pi)^{-\frac{n}{2}} \sigma_v^{-2} \left( 1 - \phi_2^2 \right)^2 - (1 + \phi_2)^2 \phi_1^2 \]

\[ \times \exp \left[ -\frac{1}{2} \left( \sigma_v^{-2} \left( \sum_{i=3}^{n} (v_i)^2 \right) \right) \right] \] (4.26)

The AR(2) model for the transformed errors can be expressed as:

\[ e_{st} = \phi_1 e_{s(t-1)} + \phi_2 e_{s(t-2)} + v_{st}, \quad t = 1, 2, \ldots, n \] (4.27)

where \( \phi_1 \) and \( \phi_2 \) are the autoregressive parameters for the errors and \( \{v_s\} = \{v_{s1}, v_{s2}, v_{s3}, \ldots, v_{sn}\}^T \) is a vector of random components.

Assuming \( e_{s0} \) and \( e_{s(-1)} \) as zero and the errors \( v_{st} \) are normally distributed with zero mean and constant variance \( \sigma_v^2 \) for all \( t > 2 \), the likelihood function for estimating the model parameters \( \theta \) and autoregressive parameters \( \phi_1 \) and \( \phi_2 \) can be constructed as follows:

\[ l(\theta, \phi_1, \phi_2, \{Q\}_{\mathrm{obs}}) = (2\pi)^{-\frac{n}{2}} \sigma_v^{-2} \left( 1 - \phi_2^2 \right)^2 - (1 + \phi_2)^2 \phi_1^2 \]

\[ \times \exp \left[ -\frac{1}{2} \left( \sigma_v^{-2} \left( \sum_{i=3}^{n} (v_i)^2 \right) \right) \right] \] (4.28)
4.4. Evaluation of implicit methods

4.4.1 Study area, model and data

To evaluate the implicit methods of uncertainty analysis, SWAT model is calibrated against the observed streamflow data of the Canard River watershed (Figure 4.1) located in the Essex region, Southwestern Ontario, Canada. The area of the watershed is 348 km$^2$ and consists of relatively flat clay plane. The major land use of the watershed is agriculture that occupies 85% of the area. The land elevation of the watershed ranges from 175 m to 197 m and is treated as mild slope. The subsurface formation of the study area consists of a series of aquifers like overburden aquifers, contact aquifers and bedrock aquifers (ERCA, 2007). The overburden aquifers include confined and unconfined aquifers. The water table in the shallow aquifer is seldom deeper than 5 meters. Due to the nature of the clay soil and the aquifer characteristics, the occurrence of groundwater recharge is very low in the area. To facilitate the root zone aeration and agricultural operations, the tile drains are extensively used in the watershed for the removal of excess water from the fine-textured clay soil (Tan et al., 2002). The major components of water budget of the area are precipitation, evapotranspiration, surface runoff, tile drain and groundwater flow (Rahman, 2007).

For SWAT model simulation, the necessary Geographic Information System (GIS) data, such as watershed boundary, Digital Elevation Model (DEM), land use and soil have been obtained from the Essex Region Conservation Authority (ERCA), Ontario, Canada. Using the ArcSWAT, the watershed is delineated into 32 sub-basins and model input data are extracted for each sub-basin. The delineation of the watershed into 32 sub-basins is shown in Figure 4.2. Based on the information of elevation, land use and soil,
each sub-basin is divided into a number HRUs and the SWAT model simulates water balance at the HRU level. For the study area, the sub-basins are divided into 170 HRUs. The climate data for the watershed, such as daily precipitation, temperature, humidity and wind speed were obtained from the Environment Canada website (http://climate.weatheroffice.gc.ca/climateData/canada_e.html) for the Windsor Airport climatic station (Figure 4.2). The climatic record of Windsor Airport shows that the annual average precipitation in the study area is 920 mm for the period of 1971 to 2000
Figure 4.2: Delineation of the Canard River watershed into sub-basins

with an average annual rainfall of 805 mm. Most of the snowfall occurs during the winter months of December - February. The winter temperature usually falls below 0°C while the average summer (June - August) temperature is around 20°C.

There is one streamflow measuring station (Figure 4.2) in the Canard River Watershed. For the calibration of SWAT model under different implicit methods of uncertainty analysis, the daily streamflow data of this gauging station obtained from the
Environment Canada website are used. The daily streamflow data for the period from 1990 to 1993 are used for SWAT model calibration. Each of the implicit methods of uncertainty analysis is evaluated by validating the SWAT model for a second set of daily streamflow data for the period from 2000 to 2003. In each case, one year is considered as a warm-up period to stabilize the initial state variables of the SWAT model.

4.4.2 Methodology

Seven likelihood functions formulated in the previous sections are used as objective functions in the calibration process for parameter inferences of the SWAT model. The summary of the implicit methods used for SWAT model calibration are described in Table 4.1. The parameter inferences are made using the SCEM-UA algorithm. For performing the MCMC analysis, the SWAT model is simulated 30,000 times in GNU OCTAVE environment using the text format input data files generated by the ArcSWAT interface for each HRU. The computational flowchart (Figure 3.3) is similar to that developed by Abbaspour (2008) for SWAT calibration and uncertainty programs (SWAT-CUP). The transformation parameters and the autoregressive model parameters are sampled together with the model parameters. The prior probability distribution of model parameters, the transformation parameters and the autoregressive model parameters are assumed to be uniform. The prior ranges of model parameters are selected by performing sensitivity analysis. The prior range of one model parameter is changed at a time by keeping the prior ranges of other model parameters unchanged and the sensitivity of prior ranges are evaluated by the posterior distribution of model parameters and the efficiency of optimum parameter values in streamflow simulation.
Table 4.1: Summary of calibration of SWAT model considering input uncertainty indirectly in the calibration process

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Error model</th>
<th>Likelihood function</th>
<th>Parameters inferred by MC sampler</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>White noise model</td>
<td>Eqn. (4.4)</td>
<td>a__CN2.mgt, a__SOL_AWC ().sol, a__EPCO.bsn, a__ESCO.bsn</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>AR(1) model without applying data transformation</td>
<td>Eqn. (4.9)</td>
<td>a__CN2.mgt, a__SOL_AWC ().sol, a__EPCO.bsn, a__ESCO.bsn, $\phi_1$</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>AR(2) model without applying data transformation</td>
<td>Eqn. (4.26)</td>
<td>a__CN2.mgt, a__SOL_AWC ().sol, a__EPCO.bsn, a__ESCO.bsn, $\phi_1, \phi_2$</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>Continuous AR model without applying data transformation</td>
<td>Eqn. (4.13)</td>
<td>a__CN2.mgt, a__SOL_AWC ().sol, a__EPCO.bsn, a__ESCO.bsn, $\tau$</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>AR(1) model with applying data transformation</td>
<td>Eqn. (4.11)</td>
<td>a__CN2.mgt, a__SOL_AWC ().sol, a__EPCO.bsn, a__ESCO.bsn, $\lambda, \phi_1$</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>AR(2) model with applying data transformation</td>
<td>Eqn. (4.28)</td>
<td>a__CN2.mgt, a__SOL_AWC ().sol, a__EPCO.bsn, a__ESCO.bsn, $\lambda, \phi_1, \phi_2$</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>Continuous AR model with applying data transformation</td>
<td>Eqn. (4.15)</td>
<td>a__CN2.mgt, a__SOL_AWC ().sol, a__EPCO.bsn, a__ESCO.bsn, $\lambda, \tau$</td>
</tr>
</tbody>
</table>
The prior ranges of these parameters are presented in Table 4.2. In each case of the Markov Chain simulations, five parallel Markov Chains are used for sampling and the optimum parameter values are estimated by the SCEM-UA algorithm. The posterior probability distribution of parameters are analyzed using the samples after the chain has reached the stationary distribution.

Table 4.2: The prior ranges of parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Upper bound</th>
<th>Lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SWAT model parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a__CN2.mgt</td>
<td>5.00</td>
<td>-5.00</td>
</tr>
<tr>
<td>a__SOL_AWC ().sol</td>
<td>0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>a__EPCO.bsn and</td>
<td>0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td><strong>Statistical model parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$\phi_3$</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$\phi_4$</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.0001</td>
<td>10.00</td>
</tr>
<tr>
<td>$\tau_*$</td>
<td>0.0001</td>
<td>10.00</td>
</tr>
</tbody>
</table>

4.4.3 Estimation of parameter uncertainty

To analyze the posterior pdf of model parameters, a total of 5,000 samples are used after the Markov Chain is converged in each calibration method. The marginal posterior pdf of SWAT model parameters, AR model parameters and transformation parameters are presented in Figures 4.3-4.6. These figures show that the marginal posterior pdf of SWAT model parameters are not normal. This represents the inadequacy of the likelihood functions to search for the global optimum values in the parameter
Figure 4.3: Marginal posterior pdfs of model parameters in white noise and AR model based calibration methods
Figure 4.4: Marginal posterior pdfs of model parameters in data transformation based calibration methods
Figure 4.5: Marginal posterior pdf of AR model parameters in AR model based calibration methods.
Figure 4.6: Marginal posterior pdf of AR model parameters and transformation parameters in data transformation based calibration methods
space. The mean and standard deviation of estimated SWAT model parameters are presented in Table 4.3. The mean and variance of AR model parameters and Box-Cox transformation parameters are shown in Table 4.4. Table 4.3 shows that when the data transformation based likelihood function is used, the variance of estimated model parameters is reduced, but the mean values of a__CN2.mgt parameter are changed significantly. The change in the values of parameter a__CN2.mgt has large implications on the estimation of direct runoff, tile drain and groundwater flow. The marginal posterior pdfs of a__SOL_AWC (.sol and a__ESCO.bsn are changed significantly in AR model based simulations from that of the Standard method. Even though the posterior distributions of parameters a__CN2.mgt and a__EPCO.bsn are not unimodal in AR model based simulations, all of the estimated SWAT model parameters are observed to be independent. The correlation between the estimated model parameters in different calibration methods is presented in Figure 4.7. When data transformation is used along with AR models in parameter inference processes, the marginal posterior pdf of a__EPCO.bsn remains multi-modal, but the marginal posterior pdf of a__ESCO.bsn is changed significantly. The marginal posterior distributions of AR model parameters (Figure 4.5) shows that the mode of the first order autoregressive parameter is 0.6 and the mode of the second order autoregressive parameter is 0.02, when data transformation is not used. This clearly indicates the presence of correlated errors in the Canard River watershed modeling. The mode of characteristic correlation time is observed to be 2 days. Therefore, the temporal dependence of errors can be considered short for the present watershed modeling. The modes of AR model parameters exhibit higher values (Figure 4.5) when data transformation is incorporated in the likelihood functions. However, the
Table 4.3: Mean (standard deviation) of SWAT model parameters in different calibration methods

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>a_CN2.mgt</th>
<th>a__SOL_AWC.sol</th>
<th>a__EPCO.bsn</th>
<th>a__ESCO.bsn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0.42 (2.61)</td>
<td>0.032 (0.009)</td>
<td>0.001 (0.027)</td>
<td>-0.046 (0.003)</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>0.50 (2.66)</td>
<td>0.014 (0.014)</td>
<td>0.004 (0.025)</td>
<td>-0.038 (0.011)</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>0.90 (2.4)</td>
<td>0.013 (0.014)</td>
<td>0.001 (0.028)</td>
<td>-0.039 (0.009)</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>0.38 (2.56)</td>
<td>0.014 (0.015)</td>
<td>0.004 (0.026)</td>
<td>-0.039 (0.011)</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>-4.32 (0.65)</td>
<td>-0.03 (0.003)</td>
<td>-0.009 (0.029)</td>
<td>0.047 (0.001)</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>-4.30 (0.67)</td>
<td>-0.03 (0.003)</td>
<td>0.006 (0.024)</td>
<td>0.047 (0.001)</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>-4.15 (0.6)</td>
<td>0.047 (0.002)</td>
<td>0.006 (0.023)</td>
<td>0.047 (0.001)</td>
</tr>
</tbody>
</table>

Table 4.4: Mean (standard deviation) of AR model parameters and Box-Cox transformation parameters in different calibration methods

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Trans_par</th>
<th>AR(1)_par</th>
<th>AR(2)_par</th>
<th>ARcont_par</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>-</td>
<td>0.58 (0.026)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>-</td>
<td>0.57 (0.025)</td>
<td>0.005 (0.005)</td>
<td>-</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.83 (0.15)</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>0.21 (0.008)</td>
<td>0.92 (0.011)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>0.22 (0.07)</td>
<td>0.92 (0.011)</td>
<td>0.004 (0.004)</td>
<td>-</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>0.24 (0.007)</td>
<td>-</td>
<td>-</td>
<td>5.78 (0.16)</td>
</tr>
</tbody>
</table>

time dependency increases in data transformation based calibration processes most likely due to increase in model structural uncertainty. Model structural uncertainty may arise if some parameters remain non-identifiable in the calibration process. Since four model parameters are considered in model calibration process, there might exist non-identifiable problem of model parameters for simulating streamflow. To verify this, the model parameters which were identified as low sensitive parameters in the research works of Rahman (2007) are included in the calibration processes considering the AR(1) model. These parameters are inferred along with the four highly-sensitive parameters. These parameters include snow pack temperature lag factor (TIMP), snow melt base
Figure 4.7: Correlation between the estimated model parameters

(In figure, AR_par_1 indicates $\phi_1$ or $\phi_{11}$, AR_par_2 indicates $\phi_2$ or $\phi_{22}$, AR_cont_par indicates $\tau$ or $\tau$, and Trans_par indicates $\lambda$.)
temperature (SMTMP), minimum melt rate for snow (SMFMX), maximum melt rate for snow (SMFMN), baseflow alpha factor (ALPHA_BF) and groundwater delay time (GW_DELAY). The parameters TIMP, SMTMP, SMFMX and SMFMN are related to snow hydrology while ALPHA_BF and GW_DELAY are related to groundwater flow. In terms of aggregate parameter concept, these model parameters are expressed as $v_{\text{TIMP+bsn}}$, $v_{\text{SMTMP+bsn}}$, $v_{\text{SMFMX+bsn}}$, $v_{\text{SMFMN+bsn}}$, $v_{\text{ALPHA_BF+gw}}$ and $v_{\text{GW_DELAY+gw}}$, respectively. The variable ‘$v$’ refers replacement to the initial parameter value. The prior ranges of these model parameters are presented in Table 4.5.

Figure 4.8 shows that the marginal posterior pdfs of previous model parameters remain unchanged when low sensitive model parameters are included in the calibration process. Therefore, it can be concluded that structural uncertainty may arise from other sources that are not accounted for by the AR processes.

### Table 4.5: The prior ranges of low sensitive model parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Initial value</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{\text{TIMP+bsn}}$</td>
<td>1.0</td>
<td>U (0.01, 1.0)</td>
</tr>
<tr>
<td>$v_{\text{SFTMP+bsn}}$</td>
<td>1.0</td>
<td>U (-5.0, 5.0)</td>
</tr>
<tr>
<td>$v_{\text{SMTMP+bsn}}$</td>
<td>0.5</td>
<td>U (-5.0, 5.0)</td>
</tr>
<tr>
<td>$v_{\text{SMFMX+bsn}}$</td>
<td>4.5</td>
<td>U (0.0, 10.0)</td>
</tr>
<tr>
<td>$v_{\text{SMFMN+bsn}}$</td>
<td>4.5</td>
<td>U (0.0, 10.0)</td>
</tr>
<tr>
<td>$v_{\text{ALPHA_BF+gw}}$</td>
<td>0.43</td>
<td>U (0.0, 1.0)</td>
</tr>
<tr>
<td>$v_{\text{GW_DELAY+gw}}$</td>
<td>31</td>
<td>U 0, 300)</td>
</tr>
</tbody>
</table>
Figure 4.8: Verification of parameter non-identifiability in the calibration process
4.4.4 Estimation of prediction uncertainty

The efficiency of the model parameter values obtained at the maximum posterior probability density in simulating streamflow is evaluated in terms of the Nash-Sutcliffe coefficient of efficiency (NS) criteria (Nash and Sutcliffe, 1970) during the calibration period (Table 4.6). The efficiency is compared on both daily and monthly timescales. The peak streamflows are usually underestimated by the SWAT model in all of the calibration methods. The model has a general tendency to overestimate low flows. The discrepancy of daily simulated flow with daily observed data for the study area indicates the presence of high model structural uncertainty in the calibration process. To identify the efficiency of SWAT model for simulating daily streamflow in different seasons, the NS values are estimated for the time period from January to April, May to June, July to October and November to December. The NS values for simulating streamflow in different seasons during model calibration and validation periods are presented in Table 4.7. This table reveals that the efficiency of streamflow simulation is lower in May to June and July to October. There is consistency in NS values during January to April and November to December in different calibration methods. However, the efficiency is lower than 0.6 during these seasons. This concludes that there exists model structural uncertainty in simulating high flow as well as low flow for the study area.

The NS values for different seasons obtained from high sensitive parameters are compared to that of high and low sensitive model parameters. This comparison is made to verify whether the efficiency of streamflow simulation is increased by including the parameters related to snow hydrology and groundwater in the calibration processes. The calibration method is carried out by considering the AR(1) model and the NS values are
Table 4.6: Efficiency of SWAT model parameters obtained at the maximum posterior density during calibration period

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>NS value on Daily timescale</th>
<th>NS value on Monthly timescale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0.47</td>
<td>0.87</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>0.47</td>
<td>0.87</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>0.47</td>
<td>0.87</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>0.47</td>
<td>0.90</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>0.36</td>
<td>0.60</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>0.36</td>
<td>0.60</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>0.42</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Table 4.7: NS values at different seasons for streamflow simulation during calibration and validation periods

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration Jan-Apr</th>
<th>Calibration May-Jun</th>
<th>Calibration Jul-Oct</th>
<th>Calibration Nov-Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0.42</td>
<td>0.47</td>
<td>0.30</td>
<td>0.54</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>0.42</td>
<td>0.40</td>
<td>0.28</td>
<td>0.54</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>0.42</td>
<td>0.40</td>
<td>0.28</td>
<td>0.54</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>0.41</td>
<td>0.48</td>
<td>0.29</td>
<td>0.53</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>0.42</td>
<td>-0.19</td>
<td>-0.30</td>
<td>0.48</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>0.42</td>
<td>-0.19</td>
<td>-0.30</td>
<td>0.49</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>0.42</td>
<td>0.47</td>
<td>0.30</td>
<td>0.54</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Validation</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0.41</td>
<td>0.16</td>
<td>0.38</td>
<td>0.33</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>0.41</td>
<td>0.24</td>
<td>0.55</td>
<td>0.36</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>0.41</td>
<td>0.24</td>
<td>0.55</td>
<td>0.36</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>0.41</td>
<td>0.24</td>
<td>0.55</td>
<td>0.36</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>0.42</td>
<td>0.30</td>
<td>0.59</td>
<td>0.28</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>0.42</td>
<td>0.30</td>
<td>0.59</td>
<td>0.28</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>0.41</td>
<td>0.16</td>
<td>0.38</td>
<td>0.33</td>
</tr>
</tbody>
</table>
presented in Table 4.8. This table reveals that the efficiency of streamflow simulation by including parameters related to snow hydrology and groundwater is marginally improved during model validation period while it is decreased during calibration period in May to June and July to October. These results also indicate the presence of model structural uncertainty in high and low flow simulation. The optimum parameter values have performed better in the calibration period when data transformation is not used in parameter inferences. This reveals that the efficiency of model parameters for simulating streamflow is reduced when the parameters are estimated by the data transformation based calibration processes. The effects of estimated model parameters on two major elements of the hydrologic cycle (evapotranspiration and streamflow) in the study area are presented in Table 4.9. The streamflow is considered here as the total contribution of direct runoff, tile flow and groundwater flow. The observed average annual streamflow during the calibration period is 36% of observed average annual precipitation. The table shows that SWAT model predicts significantly higher streamflow than the observed value using the parameters obtained from data transformation based inferences. The computed tile flow and the groundwater flow are high in data transformation based calibration method and this has contributed to increase in the simulated streamflow. On the other hand, the predicted streamflow is close to the observed value when the parameters are inferred by AR model based likelihood functions. The difference in streamflow prediction in data transformation based calibration process is most likely due to the significant changes in the value of parameter a__CN2.mgt. The efficiency of optimum parameter values for model prediction during the validation period is presented in Table 4.10. It appears that the efficiency of optimum parameter values in streamflow
Table 4.8: Effects of NS values at different seasons for streamflow simulation considering high sensitive and high and low sensitive parameters

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration method</th>
<th>Jan-Apr</th>
<th>May-Jun</th>
<th>Jul-Oct</th>
<th>Nov-Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>High sensitive parameters</td>
<td></td>
<td>0.56</td>
<td>0.37</td>
<td>0.24</td>
<td>0.53</td>
</tr>
<tr>
<td>High and low sensitive parameters</td>
<td></td>
<td>0.42</td>
<td>0.40</td>
<td>0.28</td>
<td>0.54</td>
</tr>
</tbody>
</table>

| Validation | | Jan-Apr | May-Jun | Jul-Oct | Nov-Dec |
| High sensitive parameters | | 0.38 | 0.29 | 0.59 | 0.38 |
| High and low sensitive parameters | | 0.41 | 0.24 | 0.55 | 0.36 |

Table 4.9: Average annual evapotranspiration and streamflow using the model parameters at the maximum posterior density

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Average annual evapotraspiration (% of annual precipitation)</th>
<th>Average annual computed streamflow (% of annual precipitation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>62.0</td>
<td>38.0</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>62.0</td>
<td>38.0</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>62.0</td>
<td>38.0</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>66.0</td>
<td>34.0</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>51.0</td>
<td>49.0</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>51.0</td>
<td>49.0</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>58.0</td>
<td>42.0</td>
</tr>
</tbody>
</table>

Table 4.10: Efficiency of optimum values of model parameters in streamflow prediction during validation period

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>NS value on Daily timescale</th>
<th>NS value on Monthly timescale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0.39</td>
<td>0.76</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>0.42</td>
<td>0.80</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>0.42</td>
<td>0.80</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>0.42</td>
<td>0.80</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>0.43</td>
<td>0.80</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>0.43</td>
<td>0.80</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>0.43</td>
<td>0.84</td>
</tr>
</tbody>
</table>
prediction is almost same in all AR model based methods during the validation period. However, 95% streamflow prediction uncertainty due to parameter uncertainty during calibration period is reduced when AR models are used without any data transformation. The observed streamflow data covered by prediction uncertainty due to parameter uncertainty in different calibration methods are presented in Table 4.11. The 95% prediction uncertainty bounds due to parameter uncertainty is constructed by running SWAT model using 5,000 parameter sets obtained from the posterior parameter distribution and then by calculating the 2.5th and 97.5th percentiles of streamflow from 5,000 simulated flow at each time step.

The observed streamflow data covered by 95% prediction uncertainty bounds in any calibration method are shown in Table 4.12. The standard deviation of the errors (\(\sigma\)) is estimated by the maximum likelihood procedure for each of 5,000 model simulations. Then 95% prediction uncertainty bounds due to total uncertainty are constructed by adding a constant error term \(\pm 1.96 \times \sigma\) to the computed streamflow at each time step. The streamflow obtained by using the optimum parameter set is considered as the computed flow for estimating prediction uncertainty. Total uncertainty is defined as the lumped errors arising from model parameter, input, structural and output errors. In the case of data transformation, the standard deviation of the errors in the transformed space is estimated and the constant error term is added to the transformed computed streamflow at each time step. Then the resulting outputs are back-transformed to the original output space to obtain 95% prediction uncertainty bounds. Tables 4.11 and 4.12 show that the uncertainty in streamflow prediction is increased when the AR models are applied along with data transformation. In the data transformation based parameter inferences, the
Table 4.11: Percentage of observed streamflow data covered by 95% prediction uncertainty due to parameter uncertainty during model calibration and validation periods

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration period</th>
<th>Validation period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>9.3</td>
<td>6.1</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>12.5</td>
<td>11.2</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>12.9</td>
<td>11.8</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>12.5</td>
<td>11.4</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>0.6</td>
<td>0.55</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>0.6</td>
<td>0.54</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>0.3</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Table 4.12: Percentage of observed streamflow data covered by 95% prediction uncertainty during model calibration and validation periods

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration period</th>
<th>Validation period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>95.2</td>
<td>95.4</td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>95.2</td>
<td>95.4</td>
</tr>
<tr>
<td>AR(2)_model</td>
<td>95.2</td>
<td>95.3</td>
</tr>
<tr>
<td>ARcont_model</td>
<td>95.3</td>
<td>95.3</td>
</tr>
<tr>
<td>t_AR(1)_model</td>
<td>18.1</td>
<td>12.9</td>
</tr>
<tr>
<td>t_AR(2)_model</td>
<td>17.5</td>
<td>12.9</td>
</tr>
<tr>
<td>t_ARcont_model</td>
<td>17.1</td>
<td>12.3</td>
</tr>
</tbody>
</table>

model parameters are estimated in the transformed streamflow space assuming that the modeling errors have zero mean. This assumption might not be true for the modeling errors in the retransformed streamflow space (Schaefl et al., 2007). However, the uncertainty in parameter estimation is small in data transformation based calibration methods. This may lead to narrow prediction uncertainty due to parameter uncertainty in these methods. Moreover, an increase in model structural uncertainty in the data transformation based calibration methods may produce unrealistic model prediction uncertainty due to total uncertainty. This reveals that even though the transformation of data for making the residuals normal is statistically interesting for using some nonlinear transformation to make the time series stationary, and to reduce the non-homogeneity and
non-normality, it can sometimes produce misleading results. Schaeffli et al. (2007) also raised the question about the use of data transformation for reducing the non-homogeneity of residuals in the context of parameter inference and model uncertainty estimation. So the accuracy of model parameters inferred by data transformation based calibration method needs to be verified before using them in model prediction. The performance of the three AR models without applying any data transformation in predicting streamflow is observed to be the same. Since, the value of second order autoregressive parameter is close to zero, the performance of the AR(2) model is similar to the AR(1) model. Furthermore, the value of characteristic correlation time in continuous AR model leads to the identical value of first order autoregressive parameter. Thus, insignificant differences are observed in the three AR models when data transformation is not applied. By incorporating the AR models in the likelihood function, the uncertainty in streamflow prediction due to parameter uncertainty has been reduced from that of Standard calibration method. The uncertainty in streamflow prediction due to total uncertainty in AR model based calibration methods is similar to that of Standard method since the variance of errors in the Standard method is equivalent to that of AR model based methods. The 95% streamflow prediction uncertainty due to parameter uncertainty and total uncertainty in calibration period are graphically presented in Figures 4.9-4.12. The streamflow prediction using the optimum parameter values are also shown in the figures as the simulated flow. Since, the data transformation based likelihood functions did not produce reasonable parameter estimation, the streamflow prediction uncertainty produced in data transformation based calibration methods are not presented graphically. The same parameter uncertainty and total uncertainty as estimated in the
Figure 4.9: Streamflow prediction uncertainty due to parameter uncertainty in calibration period in Standard and AR(1) model based calibration methods.
Figure 4.10: Streamflow prediction uncertainty due to parameter uncertainty in calibration period in AR(2) model and continuous AR model based calibration methods
Figure 4.11: Streamflow prediction uncertainty due to total uncertainty in calibration period in Standard and AR(1) model based calibration methods
Figure 4.12: Streamflow prediction uncertainty due to total uncertainty in calibration period in AR(2) model and continuous AR model based calibration methods
calibration process are propagated in model simulation during the validation period and the 95% streamflow prediction uncertainty due to parameter uncertainty and total uncertainty is estimated. The results are presented graphically in Figures 4.13-4.16. The figures reveal that the prediction uncertainty caused by parameter uncertainty is narrower than that of total uncertainty. This indicates the need for improvement in the model structure used to represent the hydrologic system for a watershed (Vrugt et al., 2003). The additive errors are dominant over the errors caused by estimated model parameters and have created the wider prediction uncertainty bounds due to total uncertainty. Thus the parameter uncertainty can be considered as a second level source of uncertainty contributing to model prediction uncertainty. However, 95% streamflow prediction uncertainty due to total uncertainty is similar in any AR model based calibration method. Since the standard deviation of modeling errors are very close in all of the methods, the 95% streamflow prediction uncertainty bounds are not changed for adopting different AR models in describing the residual errors.

4.4.5 Test of residual errors

The assumptions of the statistical models regarding the residual errors need to be verified. In this dissertation, some standard graphical tools are adopted for testing the homoscedasticity, correlation and normality of residuals. The tests are performed using the residuals generated by the parameter sets obtained at the maximum posterior density. The homoscedasticity of residuals is tested by plotting the standardized residuals against simulated streamflow (Figure 4.17). The residuals are calculated as the difference between the observed and simulated streamflow and are standardized by the standard deviation estimated by the different calibration methods. In the Standard method, there is
Figure 4.13: Streamflow prediction uncertainty due to parameter uncertainty in validation period for Standard and AR(1) model based calibration methods.
Figure 4.14: Streamflow prediction uncertainty due to parameter uncertainty in validation period for AR(2) model and continuous AR model based calibration methods.
Figure 4.15: Streamflow prediction uncertainty due to total uncertainty in validation period for Standard and AR(1) model based calibration methods
Figure 4.16: Streamflow prediction uncertainty due to total uncertainty in validation period in AR(2) model and continuous AR model based calibration methods
Figure 4.17: Test of homoscedasticity of standardized residuals
a systematic bias in low streamflow and the variability of residuals increases with the increase of streamflow. This indicates that the variance of the residuals are not constant. However, the variability of residuals with the increase of simulated flow is marginally reduced in the AR process based stochastic error models compared to the Standard method. But the non-homogeneity is reduced significantly by the AR processes with data transformation. The graphical observations are also tested numerically by using the Kruskal-Wallis statistics (Kruskal and Wallis, 1952) at 5% level of significance. To perform the H-test the series of residuals are divided into four groups that were described in Section 3.6. The results of H-test shows that the computed values of $H$ in all of the calibration methods are higher than the value of $\chi^2$ at 5% level of significance for 3 degrees of freedom.

The correlation of residuals are tested by the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) plots of the residuals (Figure 4.18) obtained in the Standard calibration method at the maximum posterior density. The value of correlation coefficient appears to be statistically significant at lag 1, lag 2, lag 3 and lag 6 since the PACF exceeds the 95% limit at those lags. The damping pattern of ACF up to lag 5 indicates the presence of non-seasonality in the residual series. Hence, the residuals are described by the AR processes only. The covariance matrix of the errors, its determinant and inversion becomes complicated to solve mathematically as the order of the AR model increases. Moreover, the dimension of the posterior pdf becomes larger with the increase in order of AR models. Hence, AR models with order 1 and 2 are applied, even though the ACF and PACF of residuals show significant correlation at lag 3 and lag 6. Due to the variability of input forcing of the watershed such as precipitation
Figure 4.18: ACF and PACF plot of residuals with 95% limits in Standard calibration method
and temperature, the variability of the hydrologic responses such as streamflow and evapotranspiration is very high in different seasons of the watershed. Usually the high streamflow occurs in winter, while low streamflow and high evapotranspiration occur in summer. The plot of standardized residuals vs simulated streamflow shows the overestimation of low streamflow. Due to the dynamics of precipitation-runoff process during the high flow occurrence, it is most likely that the time dependence of the residuals occurs in low flow simulation. However, based on the precipitation input and the watershed characteristics, it can be anticipated that the dependency of errors may not extend beyond lag 3. The ACF plots of residuals for other calibration methods are presented in Figure 4.19. This figure shows that the correlation of errors are still significant at 5% level in lag 2 when data transformation is not applied and in lag 1 when data transformation is applied along with the AR models. It is noted that the adopted likelihood functions have partially removed the correlation of residuals. Hence, the values of autoregressive parameters are less than that of the Standard calibration method. The performance of the continuous time AR model in reducing correlation of residuals is similar to that of discrete time AR models. This indicates that the correlation of residuals cannot be explained only by the storage effects of the watershed. It may be explained by the uncertainty in input data and uncertainty in the model itself. The uncertainty in input data and model may exaggerate the storage effects on model predictions (Reichert and Mieleitner, 2009) and subsequently the residuals are correlated. Moreover, the correlation of residuals is not reduced by incorporating the AR(2) model in the likelihood function over the AR(1) model. So it is not obvious whether the AR model of order higher than 2
Figure 4.19: ACF plot of residuals with 95% limits in AR model based and data transformation based calibration methods
will further improve the results. The treatment of input data and model structure may be an alternative to the AR models for accounting the correlation of residuals.

The normality of residuals are tested by the normal probability plots (Figures 4.20-4.21). These figures show that tails of the error distribution are far from the theoretical straight line and thus the normality assumption of the error distribution is not followed in any method. All of the AR model based calibration methods have reduced the non-normality of residuals over the white noise model based calibration method. The calibration methods with both data transformation and the AR processes have reduced the deviations of observed probability lines from the theoretical straight lines and the performance is better than the other methods. However, the assumption of normality of modeling errors is not critical for the estimation of model parameters if the modeling errors are independent and have constant variance (Hipel et al., 1977). In practice, the confidence intervals for the forecasted data are easier to calculate if the normality assumption of the residuals are satisfied (Hipel et al., 1977). The cumulative periodogram plots of residuals (Figures 4.22-4.23) reveal that the deviations of cumulative periodogram from the theoretical straight line joining the points (0,0) and (0.5,1) are beyond the 95% confidence limit lines. Therefore, the modeling errors can be described as non-random in any method. However, the deviation of cumulative periodogram from the 95% limit lines is higher in data transformation based methods compared to only AR model based methods. The inadequacy of the AR process to represent the correlated errors in the transformed space may be responsible for this. This reveals that the AR process with data transformation increases the model structural uncertainty that has resulted in non-randomness of modeling errors.
Figure 4.20: Normality plot of standardized residuals in Standard and AR model based calibration methods
Figure 4.21: Normality plot of standardized residuals in data transformation based calibration methods
Figure 4.22: Cumulative periodogram of residuals with 95% limits in Standard and AR model based calibration methods
Figure 4.23: Cumulative periodogram of residuals with 95% limits in data transformation based calibration methods
4.5 Summary

In this chapter, different sources of uncertainty in distributed hydrological modeling are accounted for implicitly by incorporating the AR models in the likelihood function. The model residuals are correlated significantly at lag 1, lag 2, lag 3 and lag 6. Hence, two discrete time series models: AR(1) and AR(2) and one continuous time AR model are incorporated in the likelihood function to account for the correlation of residuals. The continuous time AR model is used to identify the effects of model structural uncertainty on correlation of residuals. To account for the heteroscedasticity and non-normality of residuals, Box-Cox transformation of data is adopted in the likelihood function. The study reveals that the inclusion of autoregressive models in the likelihood function reduces the correlation of residuals, but cannot completely remove the non-randomness of modeling errors. This reveals the presence of model structural uncertainty in the calibration process. Moreover, similar performance of discrete time series models and continuous time series model indicates that the non-randomness of the errors may be caused by the model structural uncertainty. Due to the presence of high model structural uncertainty, the correlation of errors remains unexplained by the adopted AR models. This results in the non-uniqueness in model parameter estimation in the AR model based calibration methods. Therefore, the likelihood function needs to be further improved by considering data and model structural uncertainty. The data transformation based likelihood functions reduces the non-homogeneity and non-normality of residuals, but changes the parameter inferences significantly, especially the curve number from that of the Standard calibration method. The effects of the changes in parameter estimation on the prediction of streamflow and on the estimation of annual water budget of the Canard
River watershed. The data transformation based parameters estimated at the maximum posterior density predicts annual streamflow almost 10% higher than the observed value. The model prediction uncertainty due to parameter uncertainty is increased in data transformation based calibration methods. Moreover, the prediction uncertainty due to total uncertainty becomes unrealistic in these methods. Therefore, data transformation can be used to make the residuals homoscedastic and normal, but it is essential to check the reliability of the results based on data transformation based calibration processes.

The uncertainty in streamflow prediction due to model parameter uncertainty is reduced in AR model based calibration methods over that of the Standard calibration method. The estimated variance of errors in the AR model based calibration methods are similar to that of the Standard calibration method. Hence, the uncertainty in streamflow prediction due to total uncertainty is similar in the AR model based methods as well as in the Standard method. The uncertainty boundary of model prediction due to total errors is wider than that of parameter uncertainty in both calibration and validation periods due to the dominance of additive errors over the errors caused by the estimated parameters.

The application of data transformation has increased the model structural uncertainty in the calibration process. This results in unrealistic assessment of parameter uncertainty and causes increase in streamflow prediction uncertainty due to parameter uncertainty. Therefore, the estimation of prediction uncertainty due to parameter uncertainty has identified the weakness of data transformation based calibration process, even though it is a second level source of uncertainty contributing to prediction uncertainty.
4.6 Conclusions

Based on the findings of this chapter, the following conclusions can be drawn:

- Model structural uncertainty is high for streamflow simulation for the study area.
- None of the adopted AR models is adequate to describe the correlated errors.
- None of the adopted likelihood functions has provided unique solution in the parameter space.
- Parameter estimation becomes biased when data transformation is adopted in the calibration process.
- Three different AR models based calibration methods show similar performance in terms of streamflow simulation.
- The posterior diagnostics of model residuals verify the adequacy of the stochastic model to represent the modeling errors.
- Contribution of parameter uncertainty to model simulation uncertainty is low compared to that of total uncertainty.
- Major limitation of the implicit method is to identify the appropriate stochastic model to represent the modeling errors.
5.1 Introduction

In the explicit methods of uncertainty analysis, different sources of errors in hydrological modeling are accounted for separately during the model calibration process. In this dissertation, the uncertainty in precipitation input is considered explicitly by using the multiplicative input error model. Kavetski et al. (2006a) introduced the precipitation multiplier concept on storm-event basis to account for the systematic measurement errors of precipitation data. Thyer et al. (2009) applied the precipitation multiplier model on daily time scale, but they excluded the insensitive precipitation multipliers from the Bayesian inferences to reduce the computational cost. The insensitive precipitation multipliers are the multipliers that have little impact on model simulation and are effectively redundant (Thyer et al., 2009). Considering the precipitation multipliers as the latent variables, the posterior probability distribution of model parameters and precipitation multipliers are estimated under the Bayesian approach. Ajami et al. (2007) identified two major limitations of the precipitation multiplier approach. The first limitation is to know the true input forcing in the real world problem and thus to assess the likelihood of the input error model. The second limitation is the increase in the dimension of the posterior probability distribution caused by the precipitation multipliers. To reduce the dimensional problem, Ajami et al. (2007) introduced the mean and variance of precipitation multipliers to the system as latent variables instead of searching for every single multiplier as a latent variable. In this dissertation, the precipitation multiplier model of Kavetski et al. (2006a) is termed as the storm input error model and
the precipitation multiplier model of Ajami et al. (2007) is termed as the daily input error model. The major limitation of the daily input error model is that it performs better for a small range of variance of precipitation multipliers (Ajami et al., 2009). Moreover, Renard et al. (2009) identified that the likelihood function based on the daily input error model became a random function of the arguments when implemented under the Bayesian framework.

To reduce the dimensional and computational problems of the existing explicit methods and to identify the effects of precipitation multiplier model on parameter estimation and model prediction in distributed hydrological modeling, this dissertation has developed a multiplicative input error model by introducing the season-dependent parameters. The newly developed input error model is referred to the seasonal input error model. For evaluating the performance of the seasonal input error model, precipitation uncertainty is estimated by the storm input error and the daily input error models during the calibration of SWAT model for the Canard River watershed. The streamflow prediction obtained from seasonal input error model are compared with that of the daily input error model based calibration method and the Standard calibration method. In the Standard calibration method, the precipitation data are assumed to be known exactly and no input error model is used in the calibration process. The Standard method is selected to examine if there exists any error in observed precipitation data. The posterior probability density functions (pdfs) of model parameters based on the storm input error, the daily input error models and the posterior pdf of the Standard calibration method are described briefly in the subsequent section. The posterior pdf of model parameters and multipliers based on the seasonal input error model are developed and presented in a
separate section of this chapter. The findings of this chapter are summarized and conclusions are drawn at the end of this chapter.

5.2 Uncertainty analysis using multiplicative input error model

5.2.1 Posterior pdf for storm input error model

For the mathematical formulation of the multiplicative input error model, the hydrological model is expressed in the following form:

\[ q_t = g_r([X], \{\theta\}) \quad t = 1, 2, \ldots, n \]

(5.1)

where \( g_r(\cdot) \) is the response function in the hydrologic system, \( q_t \) is true response, such as streamflow, groundwater level etc. of the hydrologic system at time step \( t \), \( n \) is the number of time steps, \( [X] = \{x_1, x_2, x_3, \ldots, x_n\} \) is a vector of true input data, such as precipitation, temperature etc. to the hydrologic system and \( \{\theta\} = \{\theta_1, \theta_2, \theta_3, \ldots, \theta_s\} \) is a vector containing the \( s \) model parameters that need to be estimated through the calibration process.

Assuming that the precipitation data are corrupted on storm-event basis, the 'storm input error' model can be represented by the following equation (Kavetski et al., 2006a):

\[ x_k = \phi_k \tilde{x}_k, \quad k = 1, 2, \ldots, n_{\phi} \]

(5.2)

where \( \phi_k \) is the multiplicative error for the observed precipitation \( \tilde{x}_k \) at the \( k \)th storm, \( \phi_k \) is known as the precipitation multiplier and \( x_k \) is the true precipitation at the \( k \)th storm. \( n_{\phi} \) is the number of precipitation multipliers during the calibration period.

Now, incorporating the storm input error model, the hydrological model can be expressed as follows:
\[ q_t = g_t(\{X_t\}, \{\varphi\}, \{\theta\}) \quad t = 1, 2, \ldots, n \quad (5.3) \]

where \( \{\varphi\} \) is a vector of \( n_\varphi \) precipitation multipliers.

Considering errors in the modeling that are not captured by the 'storm input error' model, the eqn. (5.3) can be modified as follows:

\[ q_t = h_t(\{X_t\}, \{\theta\}, \{\varphi\}) + e_t \quad t = 1, 2, \ldots, n \quad (5.4) \]

where \( h_t(.) \) is the selected hydrologic model for the watershed response and \( e_t \) represents the errors due to the measurement errors in observed streamflow data and model structure.

The computed response vector of the model for a given hydrological model parameter vector \( \{\theta\} \) and precipitation multiplier vector \( \{\varphi\} \) can be represented as:

\[ \{Q\}_{\text{com}} = \{h_1, h_2, h_3, \ldots, h_n\}^T. \]

The observed response of the system and the error vector can be represented as \( \{Q\}_{\text{obs}} = \{q_1, q_2, q_3, \ldots, q_n\}^T \) and \( \{E\} = \{e_1, e_2, e_3, \ldots, e_n\}^T \), respectively. The superscript \( T \) represents the transpose of the vector.

Using the Bayesian theory, the posterior probability density function of model parameters and multipliers conditioned on observed precipitation data \( \{\tilde{X}\} \), and observed streamflow data \( \{Q\}_{\text{obs}} \), can be written as:

\[ p(\{\theta\}, \{\varphi\} | \{\tilde{X}\}, \{Q\}_{\text{obs}}) \propto p(\{\theta\}, \{\varphi\})p(\{Q\}_{\text{obs}} | \{\theta\}, \{\varphi\}, \{\tilde{X}\}) \quad (5.5) \]

Assuming that the multipliers have Gaussian distribution with mean \( \mu_\varphi \) and variance \( \sigma_\varphi^2 \), and applying the hierarchical Bayesian modeling, eqn. (5.5) can be written as:

\[ p(\{\theta\}, \{\varphi\} | \{\tilde{X}\}, \{Q\}_{\text{obs}}) \propto p(\{\theta\})p(\{\varphi\} | \mu_\varphi, \sigma_\varphi^2)p(\{Q\}_{\text{obs}} | \{\theta\}, \{\varphi\}, \{\tilde{X}\}) \quad (5.6) \]
Describing the variance of multipliers $\sigma_{\phi}^2$ by an inverse gamma prior (Kavetski et al., 2006a), the eqn. (5.6) can be expressed as follows:

$$p(\{\theta\}, \{\phi\}|\{X\}, \{Q\}_{obs}, \mu_{\phi}, \nu_0, s_0^2) \propto \left[ SS_{\phi}(\phi) + \nu_0 s_0^2 \right]^{\frac{n_{\nu_0}-1}{2}} \left( SS(\theta, X, \{Q\}_{obs}) \right)^{-\frac{n}{2}}$$

(5.7)

where $\nu_0$ and $s_0$ are the shape parameter and scale parameter of the inverse gamma distribution, respectively. Note that, in eqn. (5.7)

$$SS_{\phi}(\phi) = \sum_{k=1}^{n_{\nu_0}} (\phi_k - \mu_{\phi})^2$$

(5.8)

$$SS(\theta, X, \{Q\}_{obs}) = \sum_{t=1}^{n_{\nu_0}} \left( (Q)_{obs, t} - h_t(\theta, \phi, X) \right)^2$$

(5.9)

Equation (5.7) is the objective function used for parameter and multiplier inferences in storm input error model based calibration method. The details of eqn. (5.7) is described in Kavetski et al. (2006a). The variance of precipitation multipliers $\sigma_{\phi}^2$ has significant effects on hydrological model parameter estimation. If $\sigma_{\phi}^2 \rightarrow 0$, the precipitation data can be assumed to be known exactly, and if $\sigma_{\phi}^2 \rightarrow \infty$, the precipitation multipliers will follow a uniform prior distribution. For addressing these two issues, the inverse gamma distribution prior on $\sigma_{\phi}^2$ with $\nu_0 > 0$ and $s_0 > 0$ is introduced in the Bayesian system. The values of $\nu_0$ and $s_0$ can be fixed by the sensitivity analysis. Kavetski et al. (2006) recommended to consider $\mu_{\phi} = 1$ as a first approximation.

As mentioned earlier, the true input data are difficult to know (Ajami et al., 2007) and thus it is problematic to find out the parameters of storm input error model. Moreover, the identification of the storm events is not straightforward, especially in the climatic regions where the variation of precipitation is very low.
5.2.2 Posterior pdf for daily input error model

Assuming that the precipitation data are corrupted daily, the daily input error model can be represented by the following equation (Ajami et al., 2007):

\[ x_i = M_t \tilde{x}_i, \quad t = 1,2, \ldots, n \text{ and } M_t \approx N(\mu_m, \sigma_m^2) \quad (5.9) \]

\( \tilde{x}_i, \ x_i \) are the observed and true precipitations at \( t \)th time step and \( M_t \) represents a random rainfall multiplier at \( t \)th time step. The rainfall multipliers are assumed to be normally distributed with mean equal to \( \mu_m \) and variance equal to \( \sigma_m^2 \).

As mentioned earlier, the mean and variance of precipitation multipliers are introduced as latent variables to the Bayesian system in the daily input error model rather than introducing the individual precipitation multipliers as latent variables. By incorporating the daily input error model in eqn. (5.1), the eqn. (5.4) takes the following form:

\[ q_i = h_i(\{X^1\}, \{\theta\}) + e_i, \quad t = 1,2, \ldots, n \quad (5.10) \]

Using the Bayesian theory, the posterior pdf of model parameters and mean and variance of precipitation multipliers conditioned on observed precipitation data \( \{\tilde{X}^1\} \), and observed streamflow data \( \{Q^1\}_{obs} \), can be written as:

\[ p(\theta, \mu_M, \sigma_M^2 | \tilde{X}^1, \{Q^1\}_{obs}) \propto SS(\theta, \{X^1\}, \{Q^1\}_{obs})^{-\frac{n}{2}} \quad (5.11) \]

where

\[ SS(\theta, \{X^1\}, \{Q^1\}_{obs}) = \sum_{i=1}^{n} (q_i - h_i(\theta, X))^2 \quad (5.12) \]

The details of the above posterior pdf are available in Ajami et al. (2007). The daily input error model has reduced the dimensional problem of the storm input error model. The
dimension of the posterior distribution is equal to the number of hydrological model parameters and number of precipitation multipliers in the storm input error model, while it is equal to the number of hydrological model parameters and two latent variables in the daily input error model.

To solve eqn. (5.11), the MCMC simulation is commonly adopted. In each iteration, the mean and variance of precipitation multipliers are sampled along with the hydrological model parameters. The precipitation multipliers are then generated at each time step randomly from a normal distribution with the sampled mean and variance of multipliers. Thus, the likelihood function becomes a random function of the arguments (Renard et al., 2009).

5.2.3 Posterior pdf of Standard calibration method

The Standard calibration method assumes that there is no error in observed precipitation data. So, no input error model is used in the Bayesian framework for formulating the posterior pdf of hydrological model parameters and the mathematical form of the posterior distribution is similar to that of the standard least square regression method as shown below:

\[ p(\theta | \{\tilde{X}_{f}, \{Q\}_{obs}\}) \propto SS(\theta, \{\tilde{X}\}, \{Q\}_{obs})^{-\frac{n}{2}} \]  \hspace{1cm} (5.13)

where

\[ SS(\theta, \{\tilde{X}\}, \{Q\}_{obs}) = \sum_{i=1}^{n} (q_i - h_i(\theta, \tilde{X}))^2 \]  \hspace{1cm} (5.14)

The details of the above posterior pdf are available in Kavetski et al. (2006a).
5.3 Development of seasonal input error model

5.3.1 The conceptual basis of the seasonal input error model

The concept of time-dependent error model parameters is not new in hydrological modeling. To account for inputs, model structure and outputs uncertainty, the state variables are randomly perturbed at every time step in the sequential data assimilation approach with Bayesian filter (Vrugt et al., 2005; Salamon and Feyen, 2009). Reichert and Meileitner (2009) used the time-dependent stochastic parameters to account for model input and structural uncertainties. Kuczera et al. (2006) represented the modeling errors by the storm-dependent hydrological model parameters. The basic idea behind the timescale of storm-event for varying model parameters stochastically is that the rainfall during a storm-event is the forcing to the catchment water balance and the persistence of errors are likely over the storm-event timescale (Kuczera et al., 2006).

The subject of representation of input, structural and output errors by the season-dependent stochastic model parameters has drawn the attention of recent researches in hydrological modeling. For example, Yang et al. (2007a) used the variance and characteristic correlation time of the stochastic error model for dry season and wet season separately to represent the model input and structural errors. Schaeffli et al. (2007) used the mixture of two normal distributions to account for the modeling errors. The mixture components were used to represent the low flow and high flow discharge regimes.

In this study, a timescale coarser than the storm-event is assumed for perturbing the input data. If there is any error in precipitation measurement, it will collectively affect the observed seasonal precipitation. If the errors in observed seasonal precipitation can be accounted for, the errors can be distributed over the observed daily precipitation within a particular season. It is assumed that the input errors are multiplicative errors to the
observed seasonal precipitation and, for simplicity, the errors in seasonal precipitation are taken as equivalent to the errors in the daily observed precipitation within a particular season. The observed precipitation data contain information on both depth of storm and pattern (Kavetski et al., 2006a). It is assumed that the season-dependent input error model represents mainly the pattern errors of the observed daily precipitation. Since the timescale of seasonal input error model is a season, the number of precipitation multipliers to account for the input uncertainty is lower than the number of storm-event based precipitation multipliers. Thus the season-dependent input error model hopes to reduce the dimensional and computational problem of the existing explicit methods of accounting for input uncertainty and to describe the effects of input errors better in the distributed hydrological modeling.

5.3.2 Posterior pdf for the seasonal input error model

Assuming that the seasonal precipitation is corrupted by the pattern errors, the seasonal input error model can be represented by the following equation:

\[
\{X_s\} = f(\{\tilde{X}_s\}, \{m\})
\]

(5.15)

where \(\{m\} = \{m_1, m_2, m_3, \ldots, m_i\}\) is a vector containing the input error model parameters to estimate the true inputs \(\{X_s\}\), given \(\{\tilde{X}_s\} = \{\tilde{x}_{s1}, \tilde{x}_{s2}, \tilde{x}_{s3}, \ldots, \tilde{x}_{si}\}\), which is a vector of observed seasonal precipitation data.

Assuming that the seasonal input errors are multiplicative, the true precipitation in the \(i\)th season can be expressed as follows:

\[x_{si} = m_i \tilde{x}_{si} \quad i = 1, 2, \ldots, S\]

(5.16)
where $m_i$ is the multiplicative error for the precipitation of the $i$th season referred to as seasonal precipitation multiplier for the $i$th season, $S$ is the number of distinct seasons in a year considered during the calibration period and $\tilde{x}_{si}$, $x_{si}$ are the observed and true precipitations in the $i$th season, respectively. Here, we assume that the errors in seasonal precipitation are equivalent to the errors in the daily observed precipitation within a particular season. Hence, the true daily precipitation can be represented as follows:

$$x_j = m_i \tilde{x}_j \quad j = 1, 2, \ldots, n \quad (5.17)$$

where $\tilde{x}_j$, $x_j$ are the observed and true precipitations at $j$th time step and at the $i$th season, respectively.

Now, by incorporating the seasonal input error model in eqn. (5.1), the hydrological model given by eqn. (5.4) can be written as follows:

$$q_t = h_t([X], \theta, [m]) + e_t \quad t = 1, 2, \ldots, n \quad (5.18)$$

Using the Bayesian theory, the posterior pdf conditioned on observed precipitation data $\{\tilde{X}\}$ and observed streamflow data $\{Q\}_{obs}$ can be written as:

$$p([\theta], [m]|\tilde{X}, Q_{obs}) \propto p([\theta], [m])p([Q]_{obs}|[\theta], [m], \tilde{X}) \quad (5.19)$$

where $\{\tilde{X}\} = \{\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \ldots, \tilde{x}_n\}$ is a vector of observed daily precipitation data. The observed data $\{Q\}_{obs}$ affects the posterior inference through the function $p([Q]_{obs}|[\theta], [m], \{\tilde{X}\})$ and $p([\theta], [m])$ is the prior pdf of hydrological model parameters $\{\theta\}$ and seasonal input error model parameters $\{m\}$. The prior pdf represents the prior knowledge about the hydrological model parameters and seasonal input error model
parameters. Assuming the uniform prior distribution of hydrological model parameters and input error model parameters, the eqn. (5.19) can be written as:

$$p\left(\hat{\theta} \mid \bar{X}, \{Q\}_{obs} \right) \propto p\left(\tilde{\theta} \right)p\left(\{Q\}_{obs} \mid \tilde{\theta}, \{\bar{X}\} \right)$$  \hspace{1cm} (5.20)

where $$\{\tilde{\theta}\} = [\{\theta\}, \{m\}].$$

The use of uniform distribution for precipitation multipliers may result in ill-posed parameter inferences (Kavetski et al., 2006a). Hence, the Gaussian distribution of the precipitation multipliers with unknown variance is used as prior by Kavetski et al. (2006a) to correct the precipitation measurement errors. However, Vrugt et al. (2008) used the uniform prior distribution for the precipitation multipliers and made parameter inferences by the Differential Evolution Adaptive Metropolis (DREAM) algorithm (Vrugt et al., 2008). They did not observe any ill-posedness in parameter inferences. The uniform prior distribution presents that the information content in the observed precipitation is limited to pattern only and no useful information on storm-depth (Kavestki et al., 2006a, Vrugt et al., 2008). In this study, it is assumed that the seasonal multipliers correct the precipitation pattern errors in hydrological modeling. Hence the uniform distribution of seasonal precipitation multipliers is assumed as the prior distribution.

Now, assuming the distribution of the output and model errors as Gaussian with zero mean and constant variance $$\sigma^2$$, the posterior pdf can be written as:

$$p\left(\tilde{\theta}, \sigma^2 \mid \bar{X}, \{Q\}_{obs} \right) \propto p\left(\tilde{\theta}, \sigma^2 \right)p\left(\{Q\}_{obs} \mid \tilde{\theta}, \{\bar{X}\} \right)$$  \hspace{1cm} (5.21)

According to the Jeffry's rule, for the noninformative prior (Box and Tiao, 1992),

$$p(\tilde{\theta}, \sigma) \propto \frac{1}{\sigma}$$  \hspace{1cm} (5.22)
Thus the posterior pdf becomes

\[
p(\tilde{\theta}, \sigma^2 | \bar{x}, \{Q\}_{\text{obs}}) \propto \frac{1}{\sigma^2} (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} \exp \left[ -\frac{1}{2} \left\{ \sigma^2 \left( \sum_{t=1}^{n} (q_t - h_t(\tilde{\theta}, X))^2 \right) \right\} \right] \tag{5.23}
\]

Integrating \( \sigma^2 \) out, the posterior pdf becomes

\[
p(\tilde{\theta} | \bar{x}, \{Q\}_{\text{obs}}) \propto \text{SS}(\tilde{\theta}, \{X\}, \{Q\}_{\text{obs}})^{-\frac{n}{2}} \tag{5.24}
\]

where \( \text{SS}(\tilde{\theta}, \{X\}, \{Q\}_{\text{obs}}) = \sum_{t=1}^{n} (q_t - h_t(\tilde{\theta}, X))^2 \) \tag{5.25}

It is difficult to solve the posterior pdf [eqn. (5.24)] analytically. Hence, the Markov Chain Monte Carlo (MCMC) type numerical solution method is needed to solve the posterior pdf. The seasonal input error models cannot quantify the errors in the zero-depth precipitation measurements. This limitation is applied to the storm input error and daily input error model as well.

5.4 Evaluation of seasonal input error model

5.4.1 Methodology

To evaluate the seasonal input error model, the study area (Figure 4.1), model and data as used in the implicit methods are considered and the uncertainty in the precipitation input is quantified for the measured precipitation data of Windsor Airport station (Figure 4.2). The observed daily streamflow data and climatic data for the period of 1990 to 1993 are used for calibration and the observed daily streamflow data and climatic data for the period of 2000 to 2003 are used in the evaluation period.

The seasonal input error model is implemented in the calibration process of SWAT model and the posterior pdf of model parameters and input error model parameters are estimated by the SCEM-UA algorithm. To illustrate the performance of
the seasonal input error model in calibration, the SWAT model parameters are estimated by the Standard calibration method and the daily input error model based calibration method. The SWAT model parameters and input error model parameters estimated by the calibration process are used to predict streamflow in the validation period. The results obtained from three calibration methods are illustrated by three categories. These are the performance of the methods in estimating the parameter uncertainty, input uncertainty and prediction uncertainty. Moreover, the assumptions regarding the residual error models are verified.

The storm input error model is implemented in the calibration process to examine how the distributed hydrological model behaves if the observed precipitation data are perturbed on the storm-event basis. The results obtained from the storm input error model based calibration method are compared with that of the Standard calibration method. Finally, the results obtained from four calibration methods are summarized. The summary of the explicit methods used for SWAT model calibration are presented in Table 5.1.

To incorporate the multiplicative input error models in the calibration process of SWAT model, a separate program is added to the computational flowchart of SWAT-CUP (Abbaspour, 2008) (Figure 3.3). For performing the MCMC analysis, the SWAT model is simulated in GNU OCTAVE environment using the text format input data files generated by the ArcSWAT interface for each of the HRU. The input error model parameters are sampled in together with the SWAT model parameters. The prior probability distribution of SWAT model parameters, the seasonal input error model parameters and the daily input error model parameters are assumed to be uniform. In the
<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Input error model</th>
<th>Posterior pdf used for parameter inferences</th>
<th>Parameters inferred by MC sampler</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>No input error model</td>
<td>Eqn. (5.13)</td>
<td>a__CN2.mgt, a__SOL_AWC (.).sol, a__EPCO.bsn, a__ESCO.bsn</td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>Seasonal input error model</td>
<td>Eqn. (5.24)</td>
<td>a__CN2.mgt, a__SOL_AWC (.).sol, a__EPCO.bsn, a__ESCO.bsn, Season based precipitation multipliers</td>
</tr>
<tr>
<td>Daily_input_error</td>
<td>Daily input error model</td>
<td>Eqn. (5.11)</td>
<td>a__CN2.mgt, a__SOL_AWC (.).sol, a__EPCO.bsn, a__ESCO.bsn, $\mu_M, \sigma_M^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean ($\mu_M$) and variance ($\sigma_M^2$) of daily precipitation multipliers</td>
</tr>
<tr>
<td>Storm_input_error</td>
<td>Storm input error model</td>
<td>Eqn. (5.7)</td>
<td>a__CN2.mgt, a__SOL_AWC (.).sol, a__EPCO.bsn, a__ESCO.bsn, Storm-event based precipitation multipliers</td>
</tr>
</tbody>
</table>
case of storm input error model, the prior distribution of precipitation multipliers is assumed to follow a Gaussian distribution with mean one and the variance of the multipliers is assumed to follow an inverse gamma distribution with the parameters $\nu_0 = 1, s_0 = 0.01$. The prior distribution of SWAT model parameters are assumed to be uniform. In each case of the Markov Chain simulations, five parallel Markov Chains are used for sampling and the optimum parameter values are estimated by the SCEM-UA algorithm. The posterior probability distribution of parameters are analyzed using the samples after the chain has reached the stationary distribution.

5.4.2 Identifying the seasonal precipitation multipliers

Five seasonal multipliers are identified in the Canard River watershed to account for the pattern errors in the measured precipitation data. The observed streamflow is the hydrologic response of the true precipitation input to the watershed. Since the seasonal variation of precipitation is not very high in the area, the seasonal input error model parameters are selected on the basis of observed seasonal variation of streamflow in the watershed, assuming that the measured streamflow data are exact. There is a rise in the streamflow during November-December months and the peak flow occurs during the months of February-March when the temperature is frequently above the freezing temperature. The streamflow starts to recess at the end of April and takes the lowest value in the months of July-August when the occurrence of the evapotranspiration is the highest in the study area. Hence the months from January to April, May to June, July to August, September to October and November to December are selected as the distinct seasons in the watershed to quantify the precipitation pattern errors conditioned on the observed streamflow and observed precipitation data. The five seasonal precipitation multipliers
are cited here as the Jan_Apr_mult, May_Jun_mult, Jul_Aug_mult, Sep_Oct_mult and Nov_Dec_mult indicating the multipliers to correct the measured daily precipitation for the seasons corresponding to January to April, May to June, July to August, September to October and November to December, respectively.

5.4.3 Estimation of parameter uncertainty and input uncertainty

The prior ranges of SWAT model parameters, seasonal input error and daily input error model parameters used in different calibration methods are presented in Table 5.2. The parameters are sampled by the MCMC sampler using the prior ranges and considering the posterior pdf as the objective function. After the convergence of the Markov Chain, 10,000 samples are analyzed to estimate the uncertainty in SWAT model parameters in each calibration method.

The marginal posterior probability distribution of the aggregate SWAT model parameters are presented in Figure 5.1. The posterior probability distribution is changed from that of the Standard method for two model parameters a__SOL_AWC ( ).sol and a__ESCO.bsn when the input uncertainty is accounted for in the calibration procedure. The correlation between the estimated SWAT model parameters is negligible in any calibration method, even though none of the marginal posterior parameter distribution shows exact normal distribution. No significant difference is observed in posterior probability distribution of parameters obtained from the daily input error and seasonal input error methods. The uncertainty bounds of a__CN2.mgt, a__EPCO.bsn and a__ESCO.bsn are almost similar in two methods. Moreover, the distributions of a__CN2.mgt and a__EPCO.bsn remain almost uniform in all calibration methods. This indicates the non-uniqueness of model parameters. The non-uniqueness of model
Table 5.2: The prior ranges of parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Upper bound</th>
<th>Lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>a__CN2.mgt</td>
<td>5.00</td>
<td>-5.00</td>
</tr>
<tr>
<td>a__SOL_AWC ( ).sol</td>
<td>0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>Jan_Apr_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
<tr>
<td>May_Jun_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
<tr>
<td>Jul_Aug_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
<tr>
<td>Sep_Oct_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
<tr>
<td>Nov_Dec_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
<tr>
<td>( \mu_M )</td>
<td>0.80</td>
<td>1.20</td>
</tr>
<tr>
<td>( \sigma_M^2 )</td>
<td>1e-5</td>
<td>1e-2</td>
</tr>
</tbody>
</table>
Figure 5.1: Marginal posterior pdfs of SWAT model parameters in Standard, seasonal input error model and daily input error model based calibration methods.
parameters may arise from the existence of model structural uncertainty and other uncertainties that are not considered in the calibration process.

The input error model parameters are inferred along with other model parameters in both seasonal input error and daily input error model based methods. The marginal posterior probability distribution of input error model parameters are generated by the 10,000 samples after the stationary distribution is achieved. The marginal posterior probability distribution of seasonal input error model parameters is shown in Figure 5.2. This figure shows that the distribution of each seasonal precipitation multiplier is almost normal, but the mean value of any seasonal multiplier is different from the unique value. This clearly indicates the existence of errors in precipitation input to the model. The errors vary from 7% in May-June months to 32% in July-August months. The overall mean of the seasonal precipitation multipliers is 0.97 that indicates underestimation of observed precipitation in the watershed, on average. To check whether the estimated precipitation by the seasonal input error model based method is independent of the measured precipitation, the deviation between the estimated precipitation and measured precipitation are graphically presented in Figure 5.3. This figure shows that the estimated precipitation conditioned on the observed streamflow does not have any correlation with the measured precipitation. The optimal values of the seasonal precipitation multipliers obtained at the maximum posterior density are used for the estimated precipitation shown in Figure 5.3. When the precipitation data is perturbed on a daily basis in the daily input error model based method, on average, the mean of the precipitation multipliers is observed to be 0.96, which is very close to the overall mean of seasonal precipitation multipliers. Thus the daily input error model also quantifies that the measured
precipitation data are higher than the true precipitation value. The marginal posterior distributions of mean and variance of precipitation multipliers are shown in Figure 5.4.

To evaluate the estimated precipitation uncertainty at the optimal input error model parameters, a comparison of observed and estimated precipitation and observed and simulated streamflow is shown in Figure 5.5. The daily streamflow data from February/1992 to January/1993 is selected for the comparison so that most of the streamflow peaks can be covered. Figure 5.5 shows that except for one peak flow in July/1992, the streamflow simulated by the seasonal input error model based method is consistent with that of daily input error model based method. The mean value of Jul_Aug_mult is less than one in the seasonal input error model. So, the method estimates precipitation less than the observed precipitation and generates less streamflow. It is noticeable that the SWAT model usually underestimates the observed streamflow peaks in the watershed in all calibration methods. This indicates the uncertainty in model structure to simulate the high flows. However, during the calibration period, the NS value for daily streamflow simulation is 0.51 in Seasonal_input_error method while it is 0.47 in Daily_input_error method. If the model efficiency is estimated using the monthly streamflow data, the values of NS are observed to be 0.89 in Seasonal_input_error and 0.88 in Daily_input_error methods. The seasonal precipitation multipliers may be responsible for allowing some extra degrees of freedom during the calibration process so that the efficiency of Seasonal_input_error method is slightly higher than that of the Daily_input_error method. To verify this, the predictive Quantile-Quantile (QQ) plot (Laio and Tamea, 2007) in the validation period, as suggested by Thyer et al. (2009), is examined in the subsequent section.
Figure 5.2: Box plots of marginal posterior probability distribution of seasonal input error model parameters
(ends of box represent 25% and 75% quantiles, vertical bars indicate 5.0% and 95.0% quantiles, horizontal bars indicate median values and the circles indicate the mean values of seasonal precipitation multipliers)
Figure 5.3: Deviation of estimated precipitation by seasonal input error model against the measured precipitation.

Figure 5.4: Marginal posterior probability distribution of daily input error model parameters.
Figure 5.5: Comparison of observed and estimated precipitation and observed and simulated streamflow in seasonal input error model and daily input error model based calibration methods.
5.4.4 Estimation of prediction uncertainty

When the input error model is implemented in the calibration process, the modeling errors are changed significantly. In the Standard calibration method, the modeling errors are captured by the parameter uncertainty only and in the Daily_input_error and Seasonal_input_error methods, the modeling errors are represented by both parameter uncertainty and input uncertainty. The probability distribution function of Daily Root Mean Square Error (DRMSE) in three calibration methods is shown in Figure 5.6.

![Figure 5.6: Probability distribution function of DRMSE in different calibration methods.](image)

Figure 5.6 clearly shows that consideration of input uncertainty has resulted in reduction in modeling errors. This finding indicates that the explicit treatment of input uncertainty can compensate for the model structural uncertainty (Thyer et al., 2009, Ajami et al., 2007). The seasonal input error model has compensated for other sources of uncertainty in the calibration process better than the daily input error model. This can be tested by the correlation of input errors with the hydrological model parameters. In the
Daily_input_error method, the correlation of mean of precipitation multiplier with the 
a__ESCO.bsn is -0.31 and in the Seasonal_input_error method, the correlation of 
May_Jun_mult with a__ESCO.bsn is -0.50. Since the input error model accounts for the 
precipitation uncertainty, it may affect the movement of water in the hydrologic system 
and compensate for the model structural uncertainty. The reduction in modeling errors 
has resulted in better streamflow simulation in Seasonal_input_error method than in 
Daily_input_error method.

The simulated streamflow with 95% confidence interval is described here as the 
prediction uncertainty due to total uncertainty. The prediction uncertainty for the total 
errors include uncertainty due to parameter, input and other sources of uncertainty. The 
uncertainty in streamflow prediction due to total uncertainty and parameter uncertainty in 
the calibration period are presented in Figure 5.7. The streamflow prediction uncertainty 
due to total uncertainty is reduced in the Seasonal_input_error method compared to the 
other two calibration methods. In the Seasonal_input_error method, at some time steps, 
the streamflow prediction uncertainty is quantified solely by the model parameter 
uncertainty. This is also an indication of the improvement in parameter estimation in 
Seasonal_input_error method. Quantitatively, the percentages of observed streamflow 
data covered by prediction uncertainty due to parameter uncertainty are 14.4%, 12.7% 
and 9.4% in the Seasonal_input_error, Daily_input_error and Standard methods, 
respectively. The percentage of observed streamflow data covered by total 95% 
predictive interval is 95.2% in all of the calibration methods. Jin et al. (2010) developed 
an index called the Average Relative Interval Length (ARIL) to measure the quality of 
data coverage by the prediction uncertainty. The difference between the upper limit and
Figure 5.7: Streamflow prediction uncertainty due to total uncertainty and parameter uncertainty in the calibration period.
the lower limit of the confidence interval at any time step divided by the corresponding observed data is termed as the relative interval length and the average value over the time period is termed as ARIL (Jin et al., 2010). A smaller ARIL value and a larger percentage of data coverage indicate a better performance of the prediction method. The value of ARIL in the calibration period is 71.7 for the Seasonal_input_error method while it is 75.3 for the Standard and Daily_input_error methods.

For quantifying the uncertainty in streamflow prediction during model validation period, the uncertainty in input data and SWAT model parameters estimated by the calibration process are propagated through the model simulation. Therefore, the SWAT model parameters and the input error model parameters used in validation are selected from the posterior probability distribution. The streamflow prediction uncertainty due to total uncertainty and due to parameter uncertainty during the validation period are presented in Figure 5.8. This figure shows that the prediction uncertainty due to total uncertainty is lower in the Seasonal_input_error method than the other two methods. Quantitatively, the percentages of observed streamflow data covered by prediction uncertainty due to parameter uncertainty are 8.9%, 8.0% and 6.2%, while the percentages of observed streamflow data covered by 95% predictive interval of total uncertainty are 95.4%, 95.0% and 94.4% in the Daily_input_error, Seasonal_input_error and Standard methods, respectively. The values of ARIL are 151.5, 158.2 and 158.4 in the Seasonal_input_error, Daily_input_error and Standard methods, respectively. Therefore, in the validation period, the overall performance of the Seasonal_input_error method for estimating prediction uncertainty can be considered equivalent to the Daily_input_error method and better than the Standard method. During the validation period, the value of
Figure 5.8: Streamflow prediction uncertainty due to total uncertainty and parameter uncertainty in the validation period.
NS for daily streamflow simulation using the parameters obtained at the highest posterior probability density is 0.41 in Seasonal_input_error method while it is 0.40 in Daily_input_error method and 0.39 in Standard method. The values of NS using monthly streamflow data is 0.80 in Seasonal_input_error and 0.75 in Daily_input_error and Standard methods.

For assessing the consistency of total prediction uncertainty with the observed streamflow, the predictive QQ plots are used in this study. The predictive QQ plots in different calibration methods are shown in Figure 5.9 for the model parameters obtained at the maximum posterior density. Figure 5.9 shows that the uncertainty in streamflow prediction is underestimated in all methods during both calibration and validation periods. In quantitative terms, the values of reliability index are 0.69 in both Standard and Daily_input_error methods during calibration and validation. The values of reliability index are 0.67 and 0.71 in calibration and validation, respectively in Seasonal_input_error method. Hence, the prediction uncertainty quantified by the Seasonal_input_error method can be considered reliable in comparison with other methods.

5.4.5 Test of residual errors

The assumptions of any statistical error model need to be tested. While formulating the posterior probability density functions, the residuals are assumed to be independent, Gaussian with zero mean and constant variance. The QQ plot is used to verify the type of distribution of the residual errors while the ACF is used to test the correlation of the residual errors. The QQ plot of standardized residuals and the ACF plot of the residuals in different calibration methods are shown in Figure 5.10. The residuals
Figure 5.9: Predictive QQ plot in calibration and validation periods.
Figure 5.10: (a) QQ plot of standardized residuals and (b) ACF of residuals with 95% probability limits during calibration.
are calculated as the difference between the observed and simulated streamflow and are standardized by the standard deviation estimated by the different calibration methods. The QQ plot shows that the residuals in all of the calibration methods are far from the theoretical line and are correlated. If the assumption of normality is satisfied, the QQ plot would follow the theoretical line. The QQ plot shows the probability distribution of the residuals is peaked in all other method. The slope of the QQ plot is steeper than the theoretical line indicating that the high streamflows are underestimated by any calibration method. Similar observations were stated in the previous sections.

The ACF plot of residuals shows that the correlations are significant at lag 1 and lag 2 in all of the calibration methods, even though the value of ACF is lower in the Seasonal_input_error method than the two other methods. However, for testing the homoscedasticity of residuals, the standardized residuals are presented with simulated streamflow in Figure 5.11. There is a systematic bias in low streamflow and the variability of residuals increases with the increase of streamflow. This indicates the non-homogeneity of residuals in the calibration methods. Therefore, a heteroscedastic output error model (Thyer et al., 2009) needs to be considered for further improvement in the modeling results. The heteroscedastic output error model can be developed by considering the measurement errors of observed streamflow. However, in this dissertation, the explicit methods of uncertainty analysis do not consider the heteroscedastic output error model. The correlated errors may arise from the model structural uncertainty that can be accounted for by the appropriate autoregressive models. Vrugt et al. (2009) used the first order autoregressive model to the residual errors to
Figure 5.11: Test of homoscedasticity of standardized residuals during calibration.
account for the structural uncertainty and used the storm multiplier model to account for the input uncertainty.

5.5 Uncertainty analysis by storm input error model based calibration method

5.5.1 Identifying the precipitation events

The SWAT model is simulated on the daily timescale and daily precipitation data are used as inputs to the model. Hence, the observed daily rainfall hyetograph and daily streamflow hydrograph of the Canard River watershed are used for identifying the significant storm-events. To reduce the dimension of posterior probability distribution, the precipitation events are selected in a way so that the number of precipitation multipliers are low, but representative for the study area. At least three successive days with zero precipitation are considered for the separation of the precipitation events. On the other hand, if precipitation occurs, but no significant response is observed in the streamflow hydrograph, no storm-event is considered. This behaviour is observed during the July-August months when the evapotranspiration demand is very high in the study area. During the model calibration period from 1991 to 1993, a total of 38 precipitation events are selected. The identification of the precipitation events from February, 1992 (Feb-92) to January, 1993 (Jan-93) is shown in Figure 5.12. There are 16 significant precipitation events during the period.

5.5.2 Convergence of Markov Chains

The prior distribution of precipitation multipliers is assumed to be Gaussian with mean one and unknown variance. The variance is assumed to follow an inverse gamma distribution. As an initial approximation, the values of $\nu_0$ and $s_0$ are assumed to be 1 and 0.01, respectively. The lower scale parameter is considered initially for achieving faster
convergence of Markov Chains. Five parallel Markov Chains are used and the parameter inferences are made by the Metropolis-Hastings Monte Carlo sampler (Hastings, 1970; Metropolis et al., 1953). The stationary distribution is reached after 210,000 iterations considering the Gelman-Rubin criteria (Gelman and Rubin, 1992). To increase the dimension of posterior distribution, the number of iterations required for the convergence of Markov Chain are increased almost 15 times from that of the Seasonal_input_error method. The optimum parameter values are obtained by maximizing the posterior density function. Kavetski et al. (2006a) recommended a value of $s_0$ within 0.2 to 0.3. When the value of scale parameter increased from 0.01 to 0.1, the Markov Chains did not converge.
at 210,000 iterations. Hence, the results obtained from \( \nu_0 = 1, s_0 = 0.01 \) are presented in the following section for identifying the extent of the ‘storm input error’ model to the distributed hydrological modeling.

5.5.3 Estimation of precipitation uncertainty

After the convergence of the Markov Chains, the marginal posterior pdf of precipitation multipliers is estimated using 380,000 samples and the distribution is shown in Figure 5.13. This figure shows that the distribution is not unimodal. The mean and variance of the precipitation multipliers are 0.99 and 2.4e-5, respectively. This reveals that the corrected precipitation is lower than the observed precipitation. This finding is similar to that of seasonal input error model and daily input error model. However, the estimated bias is 1% in the measured precipitation, which can be neglected. Therefore, it is expected that the model prediction using the Storm_input_error model is equivalent to that of the Standard calibration method. In the next section, a comparison is made with the Standard calibration method in terms of estimated parameter uncertainty and the prediction uncertainty. The performance of optimum parameter values is evaluated during model calibration and validation periods. Some posterior diagnostics are carried out with the residuals obtained at the maximum posterior density.

5.5.4 Comparison with Standard calibration method

The marginal posterior pdfs of SWAT model parameters obtained in the Storm_input_error method are compared with that of the Standard calibration method. For estimating the marginal posterior pdf in the Standard method, 10,000 samples are used after the convergence of the Markov Chains. The posterior distributions are shown in Figure 5.14. This figure shows that the distribution patterns are similar in both
methods. Hence, the mean and standard deviation of the aggregate model parameters are the same for all parameters except the a__CN2.mgt parameter (Table 5.3). This indicates that the effects of the storm multiplier model on the parameter estimation is justifiable in the distributed hydrological modeling. Moreover, the correlation between the estimated model parameters (Table 5.4) are insignificant in any method.

![Seasonal_input_error](image)

**Figure 5.13: Marginal posterior pdf of precipitation multipliers.**

For illustrating the efficiency of the parameter values obtained at the maximum posterior density in model prediction, the NS criteria is used. The values of NS during model calibration and validation period are shown in Table 5.5. The performance of the optimum parameter values are almost equivalent in the two calibration methods. Hence, the simulated streamflow hydrographs are overlapped in the two methods during the calibration and validation periods (Figure 5.15).

The uncertainty in model parameters and precipitation estimated in the calibration period is propagated in the validation period for quantifying streamflow prediction uncertainty by the Storm_input_error method. The mean and variance of precipitation
Figure 5.14: Marginal posterior pdf of SWAT model parameters.
### Table 5.3: Comparison of mean (standard deviation) of SWAT model parameters

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>a_CN2.mgt</th>
<th>a_SOL_AWC (.sol)</th>
<th>a_EPCO.bsn</th>
<th>a_ESCO.bsn</th>
<th>a_CN2.mgt</th>
<th>a_SOL_AWC (.sol)</th>
<th>a_EPCO.bsn</th>
<th>a_ESCO.bsn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0.42 (2.61)</td>
<td>0.032 (0.009)</td>
<td>0.001 (0.027)</td>
<td>-0.046 (0.003)</td>
<td>1.19 (2.38)</td>
<td>0.032 (0.009)</td>
<td>0.001 (0.025)</td>
<td>-0.050 (0.002)</td>
</tr>
</tbody>
</table>

### Table 5.4: Correlation between estimated model parameters

<table>
<thead>
<tr>
<th></th>
<th>a_CN2.mgt</th>
<th>a_SOL_AWC (.sol)</th>
<th>a_EPCO.bsn</th>
<th>a_ESCO.bsn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>1.00</td>
<td>-0.10</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>Storm_input_error</td>
<td>1.00</td>
<td>0.12</td>
<td>0.21</td>
<td>0.09</td>
</tr>
</tbody>
</table>

### Table 5.5: Efficiency of SWAT model parameters obtained at the maximum posterior density

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>NS value on Daily timescale</th>
<th>NS value on Monthly timescale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calibration period</td>
<td>Validation period</td>
</tr>
<tr>
<td>Standard</td>
<td>0.47</td>
<td>0.87</td>
</tr>
<tr>
<td>Storm_input_error</td>
<td>0.47</td>
<td>0.89</td>
</tr>
<tr>
<td>Standard</td>
<td>0.39</td>
<td>0.76</td>
</tr>
<tr>
<td>Storm_input_error</td>
<td>0.40</td>
<td>0.77</td>
</tr>
</tbody>
</table>
Figure 5.15: Streamflow prediction using the parameter values at the maximum posterior density
multipliers estimated in the calibration period are used to generate precipitation multipliers for the validation period assuming the multipliers are normally distributed. The percentage of observed streamflow data coverage by 95% prediction uncertainty due to model parameter uncertainty and total uncertainty is presented in Table 5.6 and is compared with that of the Standard method. For illustrating the quality of data coverage, the values of ARIL (Jin et al., 2010) are calculated and are presented in Table 5.7. These tables depict that both the quantity and quality of observed data coverage by the Storm_input_error method are almost equivalent to that of the Standard method. Therefore, the streamflow prediction uncertainty quantified by implementing the 'storm input error' model in the calibration process of a distributed hydrological model is justifiable.

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration period</th>
<th>Validation period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Due to parameter uncertainty</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>9.3</td>
<td>6.1</td>
</tr>
<tr>
<td>Storm_input_error</td>
<td>9.7</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>Due to total uncertainty</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>95.2</td>
<td>95.4</td>
</tr>
<tr>
<td>Storm_input_error</td>
<td>95.1</td>
<td>95.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>75.3</td>
<td>158.4</td>
</tr>
<tr>
<td>Storm_input_error</td>
<td>75.6</td>
<td>158.8</td>
</tr>
</tbody>
</table>
As expected, the parameter uncertainty and streamflow prediction uncertainty quantified in the Storm_input_error method are similar to that of the Standard method, where no precipitation uncertainty is considered in the calibration process. Due to the low variance of precipitation multipliers, the Storm_input_error method behaves like the traditional calibration method. The inclusion of 38 precipitation multipliers seems to allow more degrees of freedom in the calibration process. Despite this freedom, the Storm_input_error method has produced reasonable results that are comparable with the Standard calibration method. This emphasizes the applicability of the storm input error model to the distributed hydrological modeling for quantifying input uncertainty. However, the posterior diagnostics of the error models need to be carried out. The QQ plot of standardized residuals (Figure 5.16) and the ACF plot of residuals (Figure 5.17) show that the residuals are non-normal and autocorrelated. The values of autocorrelation functions in Storm_input_error method coincide with that of the Standard method. The residuals obtained by the parameters values at the maximum posterior density are used for the posterior checks. The variance of residuals are observed to be heteroscedastic in the Storm_input_error method.

5.5.5 Comparison with seasonal input error model and daily input error model based calibration methods

The results obtained from the Storm_input_error method assuming low variance of precipitation multipliers as a prior are compared with that of the Seasonal_input_error and Daily_input_error methods. The comparison is made in terms of i) estimation of optimum parameters, ii) uncertainty in parameter estimation, iii) the efficiency of optimum parameter values in model prediction, iv) estimation of precipitation uncertainty, v) correlation between the estimated SWAT model parameters, vi) model
prediction uncertainty due to parameter uncertainty and total uncertainty, vii) convergence of Markov Chains, viii) difficulties of implementing input error model to the distributed hydrological modeling and ix) the tests of residuals errors. The comparison is summarized in Table 5.8. This table shows that the results obtained from the seasonal input error model are either better than or equivalent to that of the daily input error and storm input error models on the basis of the above quantitative and qualitative criteria. Moreover, the Seasonal_input_error method is easier to apply and computationally less expensive than the Storm_input_error method.

Figure 5.16: QQ plot of standardized residuals in Standard and Storm_input_error methods.
5.6 Summary

A season-dependent multiplicative input error model has been developed for quantifying the input uncertainty explicitly in a distributed hydrological modeling. The results show that accounting for input uncertainty in the calibration process improves the parameter estimation and model prediction. The parameter uncertainty and prediction uncertainty quantified by the Seasonal_input_error method are compared with that of the Daily_input_error and Standard calibration methods. The estimated model parameters are observed to be uncorrelated in any calibration method. The marginal posterior distribution of model parameters is similar in both Seasonal_input_error and Daily_input_error methods, even though the percentage of observed streamflow data covered by parameter uncertainty is the highest in the Seasonal_input_error method in the calibration period. In the validation period, the percentage of observed streamflow data
Table 5.8: Comparison of three calibration methods based on multiplicative input error model

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Seasonal_input_error</th>
<th>Daily_input_error</th>
<th>Storm_input_error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Optimum parameter values and 95% confidence limits of parameter estimation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a_CN2.mgt</td>
<td>-0.04 (-4.70, 5.03)</td>
<td>4.73 (-3.95, 5.69)</td>
<td>2.24 (-3.48, 5.86)</td>
</tr>
<tr>
<td>a__SOL_AWC (.sol)</td>
<td>0.05 (0.03, 0.05)</td>
<td>0.01 (-0.01, 0.04)</td>
<td>0.04 (0.014, 0.05)</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>0.02 (-0.05, 0.05)</td>
<td>-0.02 (-0.05, 0.05)</td>
<td>0.04 (-0.05, 0.05)</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>-0.05 (-0.06, -0.02)</td>
<td>-0.05 (-0.05, -0.03)</td>
<td>-0.04 (-0.05,-0.04)</td>
</tr>
<tr>
<td>2. Efficiency of optimum parameter values in daily streamflow simulation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NS value (calibration)</td>
<td>0.51</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>NS value (validation)</td>
<td>0.41</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>3. Estimation of precipitation uncertainty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall mean of precipitation multipliers</td>
<td>0.97</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>4. Correlation between estimated SWAT model parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>Insignificant</td>
<td>Insignificant</td>
<td>Insignificant</td>
</tr>
<tr>
<td>5. Ninety five percent prediction uncertainty due to parameter uncertainty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of data coverage (calibration)</td>
<td>14.4</td>
<td>12.7</td>
<td>9.7</td>
</tr>
<tr>
<td>% of data coverage (validation)</td>
<td>8.0</td>
<td>8.9</td>
<td>6.7</td>
</tr>
<tr>
<td>6. Ninety five percent prediction uncertainty due to total uncertainty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of data coverage (calibration)</td>
<td>95.1</td>
<td>95.3</td>
<td>95.1</td>
</tr>
<tr>
<td>% of data coverage (validation)</td>
<td>95.0</td>
<td>95.4</td>
<td>95.4</td>
</tr>
<tr>
<td>Value of ARIL ((calibration)</td>
<td>71.7</td>
<td>75.3</td>
<td>75.6</td>
</tr>
<tr>
<td>Value of ARIL ((validation)</td>
<td>151.5</td>
<td>158.2</td>
<td>158.8</td>
</tr>
<tr>
<td>7. Convergence of Markov Chain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chain converged at</td>
<td>15,000</td>
<td>5,000</td>
<td>210,000</td>
</tr>
<tr>
<td>8. Implementation of input error model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identification of storm events</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Difficulty of implementation</td>
<td>Easy</td>
<td>Easy</td>
<td>Complicated to identify storm-events and computationally expensive</td>
</tr>
<tr>
<td>9. Tests of residual errors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance test</td>
<td>Not-satisfied</td>
<td>Not-satisfied</td>
<td>Not-satisfied</td>
</tr>
<tr>
<td>Normality test</td>
<td>Not-satisfied</td>
<td>Not-satisfied</td>
<td>Not-satisfied</td>
</tr>
<tr>
<td>Lag-1 autocorrelation</td>
<td>0.54</td>
<td>0.56</td>
<td>0.57</td>
</tr>
</tbody>
</table>
covered by parameter uncertainty is almost equivalent to that of the Daily_input_error method. The overall input uncertainty quantified by the seasonal input error model is numerically very close to that quantified by the daily input error model. Both of the models quantifies that the corrected precipitation is less than the observed precipitation. The precipitation estimated by the seasonal input error model is observed to be independent of the measured precipitation. The research identifies that there exists model structural uncertainty in modeling the study area (Canard River watershed) and the seasonal input error model compensates for structural uncertainty better than the daily input error model. Hence the modeling errors are reduced in the Seasonal_input_error method and the model prediction is improved. The efficiency of streamflow prediction using the optimal parameter values is higher in the Seasonal_input_error method than the other methods. The quality of data coverage by the prediction uncertainty is higher in the Seasonal_input_error method than the other methods in both calibration and validation period. Moreover, the predictive QQ plot shows the prediction uncertainty quantified by the Seasonal_input_error method is reliable in comparison with other methods.

The results obtained from the posterior checks of residual errors show that the distribution of residual errors follows a peaked distribution with the underestimation of high flows. The autocorrelation of errors are significant at lag 1 and lag 2 in all of the calibration methods, but the value of ACF is lower in the Seasonal_input_error method than the other methods. Based on the performance of the Seasonal_input_error method, it can be concluded that the effects of seasonal input error model on parameter uncertainty and prediction uncertainty are quantifiable in the distributed hydrological modeling.
The storm-event based original multiplicative input error model using a low variance of precipitation multiplier is implemented in the calibration process to evaluate the performance of the input error model in the distributed hydrological modeling. Due to the low variance of precipitation multipliers, the Storm_input_error calibration method yields results similar to the Standard calibration method. This result shows the applicability of storm multiplier model to the highly parameterized distributed model. The comparison between the seasonal input error, daily input error and storm input error models show that the prediction uncertainty due to parameter uncertainty is reduced in the Seasonal_input_error method and the quality of data coverage by the prediction uncertainty due to total uncertainty is improved. However, the width of prediction uncertainty bounds is usually high in any input error model based calibration method. This indicates that high model structural uncertainty exists in streamflow simulation for the watershed.

In comparison with other methods, the Seasonal_input_error method is easier to apply and computationally less expensive. Even though the Daily_input_error method is easy to implement and the computational cost is low, the method is limited to low variance of precipitation multipliers. To test the robustness of the seasonal input error model in parameter estimation and model prediction, the performance of the model needs to be further evaluated for different watersheds with different hydrologic and climatic conditions.
5.7 Conclusions

Based on the findings of this chapter, the following conclusions can be drawn:

- Explicit methods of uncertainty analysis are applicable to distributed hydrological modeling.
- Precipitation estimated by the explicit input error models is lower than the measured precipitation for the study area.
- The modeling results based on the developed seasonal input error model are identical to other input error models.
- DRMSE is lower in seasonal input error model based calibration method than that of daily input error model based method.
- Model prediction uncertainty is underestimated by both traditional and explicit methods of calibration.
- None of the adopted likelihood functions has provided unique solution in the parameter space.
- Even though uniform prior distribution is assumed for seasonal precipitation multipliers, parameter inferences are not ill-posed.
- This leads to identical posterior parameter distribution in the daily and seasonal input error models based calibration methods.
- Model structural uncertainty dominates over model input and parameter uncertainties. Therefore, all explicit calibration methods show similar performance in terms of analyzing uncertainty in hydrological modeling.
- The advantage of seasonal input error model over the existing explicit error models is low dimension of posterior distribution and less computational cost.
• The major limitation of storm input error model is to identify the storm events and high dimension of posterior distribution and high computational cost.

• The daily input error model is limited to small range of variance of precipitation multipliers.

• The seasonal input error model needs to be evaluated for different hydrologic and climatic conditions.
CHAPTER VI
APPLICATION OF SEASONAL INPUT ERROR MODEL

6.1 Introduction

In the previous chapter, the seasonal input error model has been evaluated by comparing its performance for the Canard River watershed with that of daily input error model and storm input error model based calibration methods and with that of Standard calibration method. For further evaluation, the seasonal input error model is applied to another watershed, i.e., Ruscom River watershed of Southwestern Ontario, having hydrologic and climatic conditions similar to that of the Canard River watershed. The hydrological model parameter uncertainty, input data uncertainty and streamflow prediction uncertainty are quantified for the Ruscom River watershed by implementing the seasonal input error model during the calibration process of SWAT model. The results are compared with that of the Standard calibration method. The assumptions of the residual errors and input error models are tested. Moreover, the streamflow prediction uncertainty estimated by the explicit Seasonal_input_error method is compared with that of the implicit method for both of the Canard River and Ruscom River watersheds. The findings of this chapter are summarized and conclusions are drawn at the end of this chapter.

6.2 Evaluation of seasonal input error model for the Ruscom River watershed

6.2.1 Study area, model and data

The Ruscom River watershed (Figure 6.1) is located in the Essex region, Southwestern Ontario, Canada. The area of the watershed is 175 km$^2$ and consists of mainly clay soils with some sandy soils in the southern part of the watershed. Its
Figure 6.1: Location of the Ruscom River watershed

topography is described as level to slightly undulating. The major land use of the watershed is agriculture. The major components of water budget of the area are precipitation, evapotranspiration, surface runoff, tile drain and groundwater flow. The climatic conditions of the watershed are similar to that of the Canard River watershed. Based on the climatic normal record of Environment Canada at Windsor Airport, the annual average precipitation in the watershed is 920 mm for the period of 1971 to 2000 and an average annual rainfall of 805 mm. Most of the snowfall occurs during the winter
months of December - February. The daily average temperature is -4.5°C in January and 22.7°C in July.

For SWAT model simulation, the necessary Geographic Information System (GIS) data, such as watershed boundary, Digital Elevation Model (DEM), land use and soil are obtained from the Essex Region Conservation Authority (ERCA), Ontario, Canada. ArcSWAT delineates the watershed into 31 sub-basins and extracts model input data for each sub-basin. The delineation of the watershed into 31 sub-basins is shown in Figure 6.2. Based on the information of elevation, land use and soil, each sub-basin is divided into a number of HRUs and the SWAT model simulates water balance at a HRU level.

There is no weather station located within the watershed boundary. Hence, 'Woodslee', the weather station (Figure 6.2) closest to the watershed is selected for the climatic data. The climate data for the Woodslee station are obtained from the Environment Canada's website. There is one streamflow gauging station (Figure 6.2) in the Ruscom River watershed. For the calibration of the SWAT model, the daily streamflow data of this gauging station obtained from the Environment Canada website are used. The daily streamflow data for the period from 1990 to 1993 are used for SWAT model calibration and the daily streamflow data for the period from 1980 to 1983 are used for evaluating the streamflow prediction by the estimated SWAT model parameters and seasonal input error model parameters. In each case, one year prior to these periods is considered as warm-up period to stabilize the initial state variables of the SWAT model.
Figure 6.2: Delineation of the Ruscom River watershed into sub-basins
6.2.2 Methodology

Four aggregate model parameters [curve number (CN), available water holding capacity (AWC), the plant uptake compensation factor (EPCO) and soil evaporation compensation factor (ESCO)] are the most sensitive parameters for simulating streamflow of the Ruscom River watershed. These parameters are estimated by two calibration methods: Seasonal_input_error method and Standard calibration method.

To implement the seasonal input error model in the calibration process of SWAT model for the Ruscom River watershed, the seasonal precipitation multipliers are identified on the basis of observed seasonal variation of streamflow in the watershed. The measured streamflow data are assumed to be exactly known. The seasonal variation of precipitation and streamflow in the Ruscom River watershed are similar to that of the Canard River watershed. Therefore, five seasonal precipitation multipliers are selected for the Ruscom River watershed. These are Jan_Apr_mult, May_Jun_mult, Jul_Aug_mult, Sep_Oct_mult and Nov_Dec_mult. The multipliers correct the measured daily precipitation for the seasons corresponding to January to April, May to June, July to August, September to October and November to December, respectively.

The seasonal input error model parameters are sampled in together with the SWAT model parameters. The prior distribution of SWAT model parameters and the seasonal input error model parameters are assumed to be uniform and the prior ranges of these parameters are presented in Table 6.1. The parameter inferences are made using the SCEM-UA algorithm. The computational flowchart (Figure 3.3) is similar to that of the Canard River watershed in any method. To perform the MCMC analysis, the SWAT model is simulated 30,000 times in GNU OCTAVE environment using the text format input data files generated by the ArcSWAT interface for each HRU. In each case of the
MCMC simulations, five parallel Markov Chains are used for sampling and the optimum parameter values are estimated by the SCEM-UA algorithm. The posterior probability distribution of parameters are analyzed using the samples after the chain has reached the stationary distribution.

### Table 6.1: The prior ranges of parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Upper bound</th>
<th>Lower bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>a__CN2.mgt</td>
<td>5.00</td>
<td>-5.00</td>
</tr>
<tr>
<td>a__SOL_AWC ( ).sol</td>
<td>0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>Jan_Apr_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
<tr>
<td>May_Jun_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
<tr>
<td>Jul_Aug_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
<tr>
<td>Sep_Oct_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
<tr>
<td>Nov_Dec_mult</td>
<td>0.25</td>
<td>1.50</td>
</tr>
</tbody>
</table>

### 6.2.3 Estimation of parameter uncertainty and input uncertainty

After the convergence of the Markov Chains, 10,000 samples are analyzed to estimate the uncertainty in SWAT model parameters and seasonal input error model parameters. The marginal posterior probability distribution of the aggregate SWAT model parameters for the Standard calibration method and Seasonal_input_error method are presented in Figure 6.3. The posterior probability distribution of two model parameters, a__EPCO.bsn and a__ESCO.bsn, is changed from that of the Standard method when the seasonal input error model is implemented in the calibration process of the Ruscom River watershed. The distribution patterns of a__CN2.mgt and a__SOL_AWC ( ).sol are almost similar in the Seasonal_input_error and Standard calibration methods. The posterior distribution of a__CN2.mgt is far from a unimodal
Figure 6.3: Marginal posterior pdfs of SWAT model parameters in Standard and seasonal input error model based calibration methods
normal distribution. This indicates that there might exist local optimum values in the parameter space and the inference process is inadequate to find the global optimum value. However, two SWAT parameters, a__CN2.mgt and a__EPCO.bsn, are highly correlated in the Seasonal_input_error method. The correlation coefficient between a__CN2.mgt and a__EPCO.bsn is 0.54 in the Seasonal_input_error method (Table 6.2). The highest correlation in Standard calibration method is observed between a__SOL_AWC () .sol and a__ESCO.bsn and the value of correlation coefficient is 0.25 (Table 6.2). The presence of multiple optima and the correlation between the parameters are the results of highly nonlinear hydrologic system behaviour. The consideration of model structural uncertainty in the calibration process can improve the results.

<table>
<thead>
<tr>
<th>SWAT model parameters</th>
<th>a_CN2.mgt</th>
<th>a__SOL_AWC () .sol</th>
<th>a__EPCO.bsn</th>
<th>a__ESCO.bsn</th>
</tr>
</thead>
<tbody>
<tr>
<td>a_CN2.mgt</td>
<td>1.00</td>
<td>-0.04</td>
<td>0.01</td>
<td>-0.07</td>
</tr>
<tr>
<td>a__SOL_AWC () .sol</td>
<td>1.00</td>
<td>-0.03</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>1.00</td>
<td>-0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>Seasonal_input_error</th>
</tr>
</thead>
<tbody>
<tr>
<td>a_CN2.mgt</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>a__SOL_AWC () .sol</td>
<td>-0.18</td>
<td>1.00</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>0.54</td>
<td>-0.09</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>-0.05</td>
<td>0.09</td>
</tr>
</tbody>
</table>

150
The seasonal input error model parameters are inferred along with SWAT model parameters. The marginal posterior probability distribution of input error model parameters are generated by the 10,000 samples after the stationary distribution is achieved. The marginal posterior probability distribution of seasonal input error model parameters is shown in Figure 6.4. This figure shows that the distribution of each seasonal precipitation multiplier is almost normal, but the mean value of any seasonal multiplier is not equal to one. This clearly indicates the existence of errors in precipitation input to the model. The errors vary from 0.9% in January-April months to 12.5% in July-August months. The overall mean of the seasonal precipitation multipliers is 0.98 which indicates that on average, the actual precipitation conditioned on the observed discharge data is less than the observed precipitation. To check whether the estimated precipitation by the Seasonal_input_error method is independent of the measured precipitation, the deviation between the estimated precipitation and measured precipitation are graphically presented in Figure 6.5. This figure shows that the estimated precipitation conditioned on the observed streamflow does not have any correlation with the measured precipitation. The value of estimated precipitation is obtained from correcting the observed precipitation with the values of seasonal precipitation multipliers at the maximum posterior density.

To evaluate the estimated precipitation uncertainty at the optimal input error model parameters, a comparison of observed and estimated precipitation and observed and simulated streamflow during the calibration period is shown in Figure 6.6. This figure shows that the SWAT model underestimates the streamflow peaks in both of the Seasonal_input_error and Standard calibration methods demonstrating the uncertainty of
Figure 6.4: Box plots of marginal posterior probability distribution of seasonal input error model parameters
(ends of box represent 25% and 75% quantiles, vertical bars indicate 5.0% and 95.0% quantiles, horizontal bars indicate median values and the circles indicate the mean values of seasonal precipitation multipliers)
Figure 6.5: Deviation of estimated precipitation by seasonal input error method against the measured precipitation.
Figure 6.6: Comparison of observed and estimated precipitation and observed and simulated streamflow in seasonal input error model based calibration method and Standard calibration method.
model structure to simulate the high flows. Similar findings are observed when SWAT model is applied to the Canard River watershed. However, the streamflow peaks simulated by the Seasonal_input_error method in the months from January to April coincides with that of the Standard calibration method. The marginal posterior pdf of Jan_Apr_mult shows that the mean of the precipitation multiplier from January to April is 1.0. This indicates that there is no uncertainty in observed precipitation data during the months. Hence, the streamflow simulations by the Seasonal_input_error and the Standard methods are similar during these months. Moreover, the efficiency of optimum parameter values in simulating streamflow during the calibration period is similar in both methods. The value of NS for daily streamflow simulation is 0.55 in both methods. The NS value on monthly timescale is 0.79 and 0.81 in the Standard and Seasonal_input_error methods, respectively.

6.2.4 Estimation of prediction uncertainty

When the seasonal input error model is implemented in the calibration process, the probability distribution function of Daily Root Mean Square Error (DRMSE) is changed from that of the Standard calibration method (Figure 6.7). In the Standard calibration method, the modeling errors are captured by the parameter uncertainty only, while in the Seasonal_input_error method, the modeling errors are represented by both parameter uncertainty and input uncertainty. Thus, consideration of input uncertainty has resulted in reduction in modeling errors by compensating for other sources of uncertainty in the calibration process. This has been tested by the correlation of input errors with the hydrological model parameters. The correlation of May_Jun_mult with a__ESCO.bsn is very high (-0.71). This indicates that the seasonal input error model can compensate for
the model structural uncertainty. This can be confirmed by evaluating streamflow prediction uncertainty estimated by Seasonal_input_error method during the model validation period.

![Probability distribution function of DRMSE](image)

**Figure 6.7: Probability distribution function of DRMSE**

The streamflow prediction uncertainty with 95% limits due to total uncertainty and parameter uncertainty in the calibration period are presented in Figure 6.8. It is noticed that at some time steps, the streamflow prediction uncertainty is quantified solely by the model parameter uncertainty in the method. This shows an improvement in parameter estimation in the Seasonal_input_error method. Quantitatively, the percentages of observed streamflow data covered by prediction uncertainty due to parameter uncertainty are 14.7% and 9.7% in Seasonal_input_error and Standard methods, respectively. The percentage of observed streamflow data covered by total 95% prediction interval is 95.2% and 95.0% in Seasonal_input_error and Standard methods, respectively.
To quantify the uncertainty in streamflow prediction during model validation period, the uncertainty in precipitation and SWAT model parameters estimated by the calibration process are propagated through the model simulation. The streamflow prediction uncertainty due to total uncertainty and the prediction uncertainty due to parameter uncertainty in the validation period are presented in Figure 6.9. The percentages of observed streamflow data covered by prediction uncertainty due to parameter uncertainty are 9.5% and 5.2%. The percentages of observed streamflow data covered by total 95% prediction interval are 94.9% and 95.0% in the Seasonal_input_error and Standard methods, respectively. Even though some of the streamflow peaks during the validation period are not captured by 95% prediction uncertainty in the Seasonal_input_error method, the value of ARIL is 63.7, while it is 64.2 in the Standard method. Therefore, during the validation period, the overall performance of the Seasonal_input_error method can be considered better than the Standard method. However, during the validation period, the value of NS for daily streamflow simulation using the parameters obtained at the highest posterior probability density is 0.69 in both Seasonal_input_error and Standard methods. The values of NS using monthly streamflow data is 0.78 in the Seasonal_input_error method and 0.80 in the Standard method.

To assess the consistency of total prediction uncertainty with the observed streamflow, the predictive QQ plots for the model parameters obtained at the maximum posterior density in the two calibration methods are shown in Figure 6.10. The figure shows that the uncertainty in streamflow prediction is underestimated in both methods during calibration and validation periods. In quantitative terms, the values of the reliability index are 0.48 in Standard method during calibration and 0.58 in validation. In
Figure 6.8: Streamflow prediction uncertainty due to total uncertainty and parameter uncertainty in the calibration period.
Figure 6.9: Streamflow prediction uncertainty due to total uncertainty and parameter uncertainty in the validation period.
Figure 6.10: Predictive QQ plot in calibration and validation periods.
Seasonal_input_error method, the values of reliability index are 0.50 in calibration and 0.59 in validation. Hence, the prediction uncertainty quantified by the Seasonal_input_error method can be considered consistent with that of the Standard method.

6.2.5 Test of residual errors

The assumptions of the statistical error models are tested in this section. In both calibration methods, the residual errors are assumed to be independent, Gaussian with zero mean and constant variance. The QQ plot shown in Figure 6.11 is used to verify the type of distribution of the residual errors and the ACF plot shown in Figure 6.12 is used to test the correlation of the residual errors. The QQ plot is drawn with the standardized residuals and the ACF plot is drawn with the residuals obtained at the maximum posterior density in both calibration methods. The residuals are calculated as the difference between the observed and simulated streamflow and are standardized by the standard deviation estimated by the respective calibration methods.

The QQ plot shows that the residuals in any calibration method are far from the 'Theoretical' line. The slope of the QQ plot is steeper than the theoretical line indicating that the high streamflows are underestimated in any method. The ACF plot of residuals shows that the residuals are correlated and the correlations are significant at lag 1 in both of the calibration methods. For testing the homoscedasticity of residuals, the standardized residuals are plotted against simulated streamflow in Figure 6.13. This figure shows that the variability of residuals increases with the increase of streamflow. This indicates that the variance of the residuals is not constant. Therefore, a heteroscedastic error model needs to be considered. The correlated errors may arise from the model structural
Figure 6.11: QQ plot of standardized residuals during calibration.
Figure 6.12: ACF plot of residuals with 95% probability limits during calibration.
Figure 6.13: Test of homoscedasticity of standardized residuals during calibration.
uncertainty. To reduce the correlated errors, the autoregressive models can be adopted to account for model structural uncertainty. This approach was implemented in the works of Vrugt et al. (2009).

6.3 Comparison with the results of the Canard River watershed

The seasonal input error model is applied to the Canard River and Ruscom River watersheds (Figure 6.14) of the Essex region, Southwestern Ontario, Canada. The Canard River is the largest watershed while the Ruscom River is the second largest watershed in the Essex region. In this section, the results of these watershed modeling studies are summarized. The seasonal streamflow and precipitation characteristics of the watersheds based on the observed data for the period of 1981-2000 are presented in Table 6.3. This table shows that the precipitation input to the Ruscom River watershed is less than that of the Canard River watershed. However, the seasonal patterns of streamflow are similar in both watersheds. The comparison of results for the seasonal input error model in these two watersheds is presented in Table 6.4. This table shows that there are some differences and some similarities in the application of seasonal input error model to these two watersheds.

The seasonal input error model shows that the true precipitation is lower than the observed precipitation in both watersheds. The test of dependence of input error model residuals shows that the estimated precipitation is independent of measured precipitation for any of the watershed modeling. The estimated uncertainty in precipitation data is low in the Ruscom River watershed. Hence, the efficiency of SWAT model parameters at the maximum posterior density in the Seasonal_input_error method is similar to that of the Standard calibration method where no input data uncertainty is considered during the
Figure 6.14: Location of the Canard River and Ruscom River watersheds
Table 6.3: Streamflow and precipitation characteristics of the watersheds for the period of 1981-2000

<table>
<thead>
<tr>
<th></th>
<th>Jan-Apr</th>
<th>May-Jun</th>
<th>Jul-Aug</th>
<th>Sep-Oct</th>
<th>Nov-Dec</th>
<th>Annual</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ruscom River</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area (km$^2$)</td>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average precipitation (mm)</td>
<td>53</td>
<td>79</td>
<td>80</td>
<td>84</td>
<td>68</td>
<td>841</td>
</tr>
<tr>
<td>Average streamflow (m$^3$/s)</td>
<td>1.9</td>
<td>0.8</td>
<td>0.3</td>
<td>0.7</td>
<td>1.4</td>
<td>1.2</td>
</tr>
<tr>
<td><strong>Canard River</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area (km$^2$)</td>
<td>348</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average precipitation (mm)</td>
<td>69</td>
<td>86</td>
<td>83</td>
<td>88</td>
<td>76</td>
<td>947</td>
</tr>
<tr>
<td>Average streamflow (m$^3$/s)</td>
<td>2.6</td>
<td>1.1</td>
<td>0.8</td>
<td>1.1</td>
<td>1.9</td>
<td>1.7</td>
</tr>
</tbody>
</table>
Table 6.4: Comparison of results with seasonal input error model for two watersheds

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Ruscom River watershed</th>
<th>Canard River watershed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Optimum parameter values and 95% confidence limits of parameter estimation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a_CN2.mgt</td>
<td>-4.54 (-5.48, 4.03)</td>
<td>-0.04 (-4.70, 5.03)</td>
</tr>
<tr>
<td>a__SOL_AWC (.).sol</td>
<td>0.05 (0.04, 0.05)</td>
<td>0.05 (0.03, 0.05)</td>
</tr>
<tr>
<td>a_EPCO.bsn</td>
<td>-0.02 (-0.03, 0.03)</td>
<td>0.02 (-0.05, 0.05)</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>0.04 (0.02, 0.05)</td>
<td>-0.05 (-0.06, -0.02)</td>
</tr>
<tr>
<td>2. Efficiency of optimum parameter values in daily streamflow simulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NS value (calibration)</td>
<td>0.55</td>
<td>0.51</td>
</tr>
<tr>
<td>NS value (validation)</td>
<td>0.69</td>
<td>0.41</td>
</tr>
<tr>
<td>3. Estimation of precipitation uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall mean of precipitation multipliers</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>4. Mean and 95% confidence limits of seasonal precipitation multipliers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan_Apr_mult</td>
<td>1.01 (0.96, 1.05)</td>
<td>1.17 (1.12, 1.22)</td>
</tr>
<tr>
<td>May_Jun_mult</td>
<td>0.94 (0.86, 1.02)</td>
<td>0.93 (0.84, 1.02)</td>
</tr>
<tr>
<td>Jul_Aug_mult</td>
<td>0.88 (0.82, 0.93)</td>
<td>0.68 (0.61, 0.74)</td>
</tr>
<tr>
<td>Sep_Oct_mult</td>
<td>1.07 (1.0, 1.14)</td>
<td>1.07 (0.99, 1.15)</td>
</tr>
<tr>
<td>Nov_Dec_mult</td>
<td>1.02 (0.94, 1.10)</td>
<td>1.01 (0.94, 1.09)</td>
</tr>
<tr>
<td>4. Correlation between estimated SWAT model parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between a_CN2.mgt and a__EPCO.bsn</td>
<td>0.54</td>
<td>0.12</td>
</tr>
<tr>
<td>5. Correlation between precipitation multipliers and SWAT model parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between May_Jun_mult and a__ESCO.bsn</td>
<td>-0.71</td>
<td>-0.50</td>
</tr>
<tr>
<td>6. Ninety five percent prediction uncertainty due to parameter uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of data coverage (calibration)</td>
<td>14.7</td>
<td>14.4</td>
</tr>
<tr>
<td>% of data coverage (validation)</td>
<td>9.5</td>
<td>8.0</td>
</tr>
<tr>
<td>7. Ninety five percent prediction uncertainty due to total uncertainty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of data coverage (calibration)</td>
<td>95.2</td>
<td>95.1</td>
</tr>
<tr>
<td>% of data coverage (validation)</td>
<td>94.9</td>
<td>95.0</td>
</tr>
<tr>
<td>8. Prediction of streamflows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High streamflows</td>
<td>Underestimated</td>
<td>Underestimated</td>
</tr>
<tr>
<td>Reliability index of prediction</td>
<td>0.58</td>
<td>0.71</td>
</tr>
<tr>
<td>9. Convergence of Markov Chain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chain converged at</td>
<td>16,000</td>
<td>15,000</td>
</tr>
<tr>
<td>10. Tests of residual errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance of errors</td>
<td>Heteroscedastic</td>
<td>Heteroscedastic</td>
</tr>
<tr>
<td>Distribution of errors</td>
<td>Non-normal</td>
<td>Non-normal</td>
</tr>
<tr>
<td>Lag 1 autocorrelation</td>
<td>0.20</td>
<td>0.54</td>
</tr>
<tr>
<td>10. Dependence of estimated precipitation on measured values</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated precipitation</td>
<td>Independent</td>
<td>Independent</td>
</tr>
</tbody>
</table>
calibration period. However, the estimated uncertainty in seasonal precipitation is higher in the Canard River watershed than that of the Ruscom River. The efficiency of optimum parameter values in streamflow prediction during the validation period is higher in the Ruscom River watershed. The reason might be that the model structural uncertainty is relatively low for the Ruscom River watershed modeling. The ACF plots of residuals show that the lag1 correlation coefficient is smaller in the Ruscom River watershed than that of the Canard River watershed. The model structural uncertainty may result in correlated errors and the correlation of errors can be increased if there exists input uncertainty in hydrological modeling. Hence, in the Canard River watershed, the correlation of residuals is observed to be reduced when the seasonal input error model is implemented in the calibration process.

The marginal posterior distributions of SWAT model parameters show that the posterior pdf of a_ESCO.bsn is different for the two watersheds. The values of a_ESCO.bsn are lower in the Canard River watershed modeling. This indicates that the model is extracting evaporative demand from the lower levels of soils. The seasonal input error model quantifies that the corrected precipitation is lower than the observed precipitation in both watersheds and quantifies higher precipitation uncertainty in the Canard River watershed modeling. The correlation between May_Jun_mult and a_ESCO.bsn is higher in the Ruscom River watershed than in the Canard River watershed. The higher correlation between the seasonal precipitation multiplier with hydrological model parameter indicates that the seasonal input error model has compensated for the model structural uncertainty in the Ruscom River watershed. However, the reliability index based on the predictive QQ plot shows that the streamflow
prediction is more reliable in the Canard River watershed than in the Ruscom River watershed. The uncertainty in streamflow prediction due to parameter uncertainty during the validation period is marginally lower in the Ruscom River watershed while the streamflow prediction uncertainty due to total uncertainty is almost identical in both watersheds. However, the distribution of modeling errors in the Ruscom River watershed is different from that of the Canard River watershed. In both watersheds, streamflow prediction uncertainty due to parameter uncertainty in the Seasonal_input_error method is lower than that of the Standard calibration method.

The convergence of Markov Chains for simulating streamflow is slightly slower in the Ruscom River watershed. The reason may be the existence of different local maxima in the model parameter spaces. Therefore, the correlation between a_CN2.mgt and a_EPCO.bsn is observed to be higher in the Ruscom River watershed in comparison with the Canard River watershed. The posterior checks of residuals errors show that the errors are non-normal, heteroscedastic and correlated when the seasonal input error is applied to both watersheds. However, the correlation of residuals is lower in the Ruscom River watershed indicating less model structural and input uncertainty in the calibration process.

6.4 Comparison of seasonal input error model and AR(1) model based calibration methods

6.4.1 Methodology

The Seasonal_input_error method accounts for precipitation uncertainty explicitly in the model calibration process while the AR(1)_model method accounts for precipitation uncertainty implicitly, lumping it with other sources of uncertainty in the calibration process of SWAT model. This section provides a comparison of the results
obtained from the Seasonal_input_error and the AR(1) model for both of the Ruscom River and Canard River watersheds. The posterior probability distribution is analyzed by the SCEM-UA algorithm. The prior distributions of AR(1) model parameter, the seasonal precipitation multipliers and SWAT model parameters are assumed to be uniform. The prior ranges of these parameters used for uncertainty analysis of hydrological models of both watersheds are shown in Table 6.5. Five parallel Markov Chains are used for parameter inferences and the uncertainty in model parameters and inputs are estimated by the 10,000 samples after the convergence of Markov Chains. The simulations are performed under the GNU OCTAVE environment.

Table 6.5: The prior ranges of parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Ruscom River watershed</th>
<th>Canard River watershed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SWAT model parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a__CN2.mgt</td>
<td>U(5.00,-5.00)</td>
<td>U(5.00,-5.00)</td>
</tr>
<tr>
<td>a__SOL_AWC (.).sol</td>
<td>U(0.05,-0.05)</td>
<td>U(0.05,-0.05)</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>U(0.05,-0.05)</td>
<td>U(0.05,-0.05)</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>U(0.05,-0.05)</td>
<td>U(0.05,-0.05)</td>
</tr>
<tr>
<td><strong>AR(1) model parameter</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi )</td>
<td>U(0.00,1.00)</td>
<td>U(0.00,1.00)</td>
</tr>
<tr>
<td><strong>Seasonal precipitation multipliers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan_Apr_mult</td>
<td>U(0.25,1.50)</td>
<td>U(0.25,1.50)</td>
</tr>
<tr>
<td>May_Jun_mult</td>
<td>U(0.25,1.50)</td>
<td>U(0.25,1.50)</td>
</tr>
<tr>
<td>Jul_Aug_mult</td>
<td>U(0.25,1.50)</td>
<td>U(0.25,1.50)</td>
</tr>
<tr>
<td>Sep_Oct_mult</td>
<td>U(0.25,1.50)</td>
<td>U(0.25,1.50)</td>
</tr>
<tr>
<td>Nov_Dec_mult</td>
<td>U(0.25,1.50)</td>
<td>U(0.25,1.50)</td>
</tr>
</tbody>
</table>
6.4.2 Uncertainty analysis of the Ruscom River watershed modeling

The input error model implemented in the Seasonal_input_error method has quantified that the mean values of seasonal precipitation multipliers vary from almost 1.0 in January-April months to 0.88 in July-August months. On average, the observed precipitation is higher than the corrected precipitation. However, the uncertainty in precipitation data is low in the Ruscom River watershed. The marginal posterior distribution of SWAT model parameters shown in Figure 6.15 reveals that consideration of precipitation uncertainty explicitly has influenced the parameter inferences and made the distribution of a__EPCO.bsn unimodal in the Seasonal_input_error model. The posterior distribution of a__CN2.mgt parameter is not normal in any calibration method. The findings indicate that none of the methods could search the unique value of a__CN2.mgt from the parameter space, even though it is a very sensitive parameter for SWAT model simulation. The marginal posterior distribution of AR(1) model parameter is shown in Figure 6.16. This figure shows that the distribution is unimodal and the mode of the first order AR model parameter is 0.22.

The values of SWAT model parameters obtained at the maximum posterior probability and the uncertainty in parameter estimation with 95% confidence limits are shown in Table 6.6. Table 6.7 demonstrates that the estimated model parameters such as a__EPCO.bsn and a__CN2.mgt are more correlated in the Seasonal_input_error method while the parameters a__SOL_AWC ( ).sol and a__ESCO.bsn are more correlated in AR(1)_model calibration method. The efficiency of the optimum parameter values in streamflow simulation during calibration and validation periods is shown in Table 6.8. This table shows that the efficiency of parameters at the maximum posterior density is the same for daily streamflow simulation in both implicit and explicit methods. If the
Figure 6.15: Marginal posterior pdfs of SWAT model parameters for the Ruscom River watershed
Figure 6.16: Marginal posterior pdf of AR(1) model parameters for the Ruscom River watershed

Table 6.6: Optimum values of SWAT model parameters with 95% confidence limits for the Ruscom River watershed

<table>
<thead>
<tr>
<th>SWAT model parameters</th>
<th>AR(1)_model</th>
<th>Seasonal_input_error</th>
</tr>
</thead>
<tbody>
<tr>
<td>a_CN2.mgt</td>
<td>-4.99 (-5.81, 3.93)</td>
<td>-4.54 (-5.48, 4.03)</td>
</tr>
<tr>
<td>a__SOL_AWC (.sol)</td>
<td>0.05 (0.03, 0.05)</td>
<td>0.05 (0.04, 0.05)</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>0.01 (-0.05, 0.05)</td>
<td>-0.02 (-0.03, 0.03)</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>0.03 (0.01, 0.04)</td>
<td>0.04 (0.02, 0.05)</td>
</tr>
</tbody>
</table>
Table 6.7: Correlation of SWAT model parameters for the Ruscom River watershed

<table>
<thead>
<tr>
<th>SWAT model parameters</th>
<th>a_CN2.mgt</th>
<th>a__SOL_AWC ().sol</th>
<th>a__EPCO.bsn</th>
<th>a__ESCO.bsn</th>
</tr>
</thead>
<tbody>
<tr>
<td>a_CN2.mgt</td>
<td>1.00</td>
<td>0.05</td>
<td>0.00</td>
<td>-0.05</td>
</tr>
<tr>
<td>a__SOL_AWC ().sol</td>
<td></td>
<td>1.00</td>
<td>0.00</td>
<td>0.28</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td></td>
<td></td>
<td>1.00</td>
<td>-0.03</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 6.8: Efficiency of optimum parameter values for streamflow simulation in the Ruscom River watershed

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>NS value on</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily timescale</td>
<td>Monthly timescale</td>
<td></td>
</tr>
<tr>
<td>Calibration period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>0.55</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>0.55</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Validation period</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>0.69</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>0.69</td>
<td>0.78</td>
<td></td>
</tr>
</tbody>
</table>

simulated streamflow is expressed on a monthly timescale, the efficiency of model parameters is high in the AR(1)_model method during calibration period while it is close to that of the Seasonal_input_error method during validation period. The probability distribution function of standard deviation of errors in the AR(1)_model calibration method and the Daily Root Mean Square Error (DRMSE) in the Seasonal_input_error method are shown in Figure 6.17. This figure shows that the consideration of input uncertainty explicitly has resulted in reduction in modeling errors. This leads to the conclusion that the seasonal input error model can compensate for other sources of
Figure 6.17: Posterior pdfs of standard deviation of errors in AR(1) model based calibration method and DRMSE in seasonal input error model based calibration method in the Ruscom River watershed.
uncertainty in the calibration process. This might cause an increase in the correlation of parameters a__EPCO.bsn and a__CN2.mgt in the Seasonal_input_error method.

The estimated uncertainty in SWAT model parameters and the input error model parameters are propagated during the validation period for quantifying streamflow prediction uncertainty in the Seasonal_input_error method, while the estimated uncertainty in SWAT model parameters are used to quantify prediction uncertainty in the AR(1)_model method. The percentages of observed streamflow data covered by the 95% prediction uncertainty due to parameter uncertainty and total uncertainty are presented in Table 6.9.

Table 6.9: Percentage of observed streamflow data covered by 95% prediction uncertainty in the Ruscom River watershed

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration period</th>
<th>Validation period</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)_model</td>
<td>11.2</td>
<td>6.1</td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>14.7</td>
<td>9.5</td>
</tr>
<tr>
<td><strong>Due to total uncertainty</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>94.9</td>
<td>94.9</td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>95.2</td>
<td>94.9</td>
</tr>
</tbody>
</table>

Table 6.9 shows that the streamflow prediction uncertainty due to parameter uncertainty is low in the Seasonal_input_error method while the percentages of observed streamflow data bracketed by total uncertainty are equal in both methods. This indicates that the parameter and streamflow prediction uncertainty quantified by the Seasonal_input_error method is comparable with the method where different sources of errors are lumped together in the AR(1)_model. Furthermore, for quantifying the reliability of streamflow prediction, the predictive QQ plot is examined. The predictive QQ plots for the calibration and validation period in both methods are shown in Figure 6.18. This figure
Figure 6.18: Predictive QQ plots in calibration and validation periods in the Ruscom River watershed.
reveals that both methods underestimate the streamflow prediction uncertainty. The reliability indices of streamflow predictions for these methods are presented in Table 6.10. This table shows that the streamflow prediction uncertainty is reliable in Seasonal_input_error method and comparable to that of the AR(1)_model.

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)_model</td>
<td>0.49</td>
<td>0.58</td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>0.50</td>
<td>0.57</td>
</tr>
</tbody>
</table>

The QQ plot of the standardized residuals of the AR(1)_model and Seasonal_input_error methods are shown in Figure 6.19 to verify the normal distribution of the modeling errors. This figure shows that the residuals do not follow the normal distribution. The ACF plots of residuals (Figure 6.20) show that the residuals are not correlated in the AR(1)_model while the autocorrelation parameter is significant at lag 1 with 95% confidence in the Seasonal_input_error method. Moreover, the standardized residuals have non-constant variance in both calibration methods (Figure 6.21). The residuals are standardized by the standard deviation of errors estimated by the calibration method. The residuals are the differences between the observed and simulated streamflow obtained at the maximum posterior density. The cumulative periodograms with 95% confidence limits for the residuals of the AR(1)_model and Seasonal_input_error methods are shown in Figure 6.22 to verify the randomness of the modeling errors. This figure shows that the cumulative periodogram is within the 95% confidence limits for the AR(1)_model. This reveals that the representation of errors by the AR(1) process has
accounted for model input and structural uncertainty adequately so that the residuals are independent. Since the residual errors are uncorrelated, the parameter inferences based on the AR(1) process can be considered reliable for the Ruscom River watershed. Due to inadequate representation of model structural uncertainty in the Seasonal_input_error method, the residuals are observed to be correlated. However, the value of ACF is 0.2 at lag 1 which is not very high. The application of seasonal input error model for uncertainty analysis of the Ruscom River watershed modeling indicates that the season-dependent input error model is capable of identifying low precipitation uncertainty. Even though the assumptions of modeling errors are not fully satisfied in the Seasonal_input_error method, the uncertainty in parameter estimation and model prediction are equally reliable to that of the AR(1)_model.

Figure 6.19: QQ plot of standardized residuals in the Ruscom River watershed.
Figure 6.20: ACF of residuals with 95% probability limits in the Ruscom River watershed.
Figure 6.21: Test of homoscedasticity of standardized residuals in the Ruscom River watershed.
Figure 6.22: Cumulative periodogram of residuals with 95% limits in the Ruscom River watershed
6.4.3 Uncertainty analysis of the Canard River watershed modeling

The precipitation uncertainty quantified by the seasonal input error model is higher in the Canard River watershed modeling than that of the Ruscom River watershed modeling. In January-April months, the mean of seasonal precipitation multiplier is 1.17 while in July-August months, the mean of seasonal precipitation multiplier is 0.68. Thus, the variation of precipitation uncertainty in different seasons is high in the Canard River watershed. On average, the model quantifies that the observed precipitation is higher than the corrected precipitation.

The marginal posterior distribution of SWAT model parameters for the AR(1) model and Seasonal_input_error methods are shown in Figure 6.23. This figure reveals that consideration of precipitation uncertainty explicitly has changed the distribution of a__SOL_AWC().sol from unimodal normal in the AR(1) model method to exponential in the Season_input_error method. The distribution of other SWAT model parameters are not normal in any method. The posterior distribution of a__CN2.mgt and a__EPCO.bsn are almost uniform in any method. These findings indicate that the Bayesian inferences based on both implicit and explicit methods produce similar results in terms of parameter sampling efficiency. However, the correlation between the estimated SWAT model parameters (Table 6.11) is low in both implicit and explicit methods. The values of SWAT model parameters obtained at the maximum posterior probability and the uncertainty in parameter estimation with 95% confidence limits are shown in Table 6.12. This table shows the wider range of parameter values. The findings are similar to that depicted in Figure 6.23. The marginal posterior distribution of the AR(1) model parameter is shown in Figure 6.24. This figure shows that the distribution is
Figure 6.23: Marginal posterior pdfs of SWAT model parameters for the Canard River watershed.
Table 6.11: Correlation of SWAT model parameters for the Canard River watershed

<table>
<thead>
<tr>
<th>SWAT model parameters</th>
<th>a_CN2.mgt</th>
<th>a__SOL_AWC (.sol)</th>
<th>a__EPCO.bsn</th>
<th>a__ESCO.bsn</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AR(1)_model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a_CN2.mgt</td>
<td>1.00</td>
<td>-0.02</td>
<td>0.06</td>
<td>-0.02</td>
</tr>
<tr>
<td>a__SOL_AWC (.sol)</td>
<td>1.00</td>
<td></td>
<td>0.00</td>
<td>0.22</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Seasonal_input_error</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a_CN2.mgt</td>
<td>1.00</td>
<td>-0.04</td>
<td>0.12</td>
<td>-0.10</td>
</tr>
<tr>
<td>a__SOL_AWC (.sol)</td>
<td>1.00</td>
<td></td>
<td>0.19</td>
<td>0.07</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.12: Optimum values of SWAT model parameters with 95% confidence limits for the Canard River watershed

<table>
<thead>
<tr>
<th>SWAT model parameters</th>
<th>AR(1)_model</th>
<th>Seasonal_input_error</th>
</tr>
</thead>
<tbody>
<tr>
<td>a_CN2.mgt</td>
<td>4.18 (-4.63, 5.60)</td>
<td>-0.04 (-4.71, 5.03)</td>
</tr>
<tr>
<td>a__SOL_AWC (.sol)</td>
<td>0.01 (-0.02, 0.04)</td>
<td>0.05 (0.03, 0.05)</td>
</tr>
<tr>
<td>a__EPCO.bsn</td>
<td>0.02 (-0.05, 0.05)</td>
<td>0.02 (-0.05, 0.05)</td>
</tr>
<tr>
<td>a__ESCO.bsn</td>
<td>-0.05 (-0.06, -0.02)</td>
<td>-0.05 (-0.06, -0.02)</td>
</tr>
</tbody>
</table>

unimodal and the mode of the first order AR model parameter is 0.60, which is higher than that of the Ruscom River watershed.

The efficiency of the optimum parameter values in streamflow simulation during calibration and validation period is shown in Table 6.13. This table shows that the efficiency of parameters at the maximum posterior density is similar for streamflow prediction during validation period in both implicit and explicit methods. During calibration period, the efficiency of optimum parameters in streamflow simulation is higher in the Seasonal_input_error method. The probability distribution functions of standard deviation of errors in the AR(1)_model calibration method and the DRMSE in the Seasonal_input_error method are shown in Figure 6.25. This figure shows that the
Figure 6.24: Marginal posterior pdf of AR(1) model parameters for the Canard River watershed

Table 6.13: Efficiency of optimum parameter values for streamflow simulation in the Canard River watershed

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>NS value on</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily timescale</td>
<td>Monthly timescale</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Calibration period</td>
<td>Validation period</td>
<td></td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>0.47</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>0.51</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>0.42</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>0.41</td>
<td>0.80</td>
<td></td>
</tr>
</tbody>
</table>
Figure 6.25: Posterior pdfs of standard deviation of errors in AR(1) model based calibration method and DRMSE in seasonal input error model based calibration method in the Canard River watershed.
distribution of modeling errors are narrow in the Seasonal_input_error method. This implies that the seasonal input error model can compensate for other sources of uncertainty in the calibration process. However, for the reliability of streamflow prediction, the predictive QQ plots for both implicit and explicit methods are examined.

The estimated uncertainty in SWAT model parameters and the seasonal input error model parameters are propagated through SWAT model simulation during the validation period and streamflow prediction uncertainty due to parameter uncertainty and total uncertainty is estimated in the Seasonal_input_error method. The percentages of observed streamflow data covered by the 95% prediction uncertainty are presented in Table 6.14. A comparison of prediction uncertainty estimated by the AR(1)_model with the Seasonal_input_error method is also shown in this table.

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration period</th>
<th>Validation period</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)_model</td>
<td>12.8</td>
<td>11.8</td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>14.4</td>
<td>8.0</td>
</tr>
<tr>
<td><strong>Due to total uncertainty</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1)_model</td>
<td>95.2</td>
<td>95.3</td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>95.2</td>
<td>95.0</td>
</tr>
</tbody>
</table>

Table 6.14 shows that the streamflow prediction uncertainty due to parameter uncertainty is lower in the Seasonal_input_error method in calibration period and higher in validation period than that of the AR(1)_model. The percentages of observed streamflow data bracketed by total uncertainty are almost equal in both methods in calibration and validation periods. This reveals that the effects of seasonal input error model on parameter estimation and streamflow prediction can be quantified in the
distributed hydrological modeling. The predictive QQ plots for both methods (Figure 6.26) reveal that both methods underestimate the streamflow prediction uncertainty. However, the reliability indices of streamflow predictions (Table 6.15) show that the reliability of streamflow prediction uncertainty in the Seasonal_input_error method is equivalent to that of the AR(1)_model. This finding confirms that there is no inconsistency in streamflow prediction in the Seasonal_input_error method in comparison with the AR(1)_model. Similar observations are presented when the results of the Seasonal_input_error method are compared with that of the Daily_input_error method for the Canard River watershed (section 5.4.4).

The QQ plot of standardized residuals (Figure 6.27) for the AR(1)_model and Seasonal_input_error methods show that the residuals are not normal in any method. The ACF plots of residuals (Figure 6.28) show that the residuals are correlated in both the AR(1)_model and Seasonal_input_error methods. The autocorrelation parameter is significant at lag 1 and lag 2 with 95% confidence in both methods. However, the lag 1 correlation is higher with the residuals of the Seasonal_input_error method. This reveals that the AR(1) process can represent the modeling errors better than the seasonal input error model. This is an expected result since the likelihood function of the AR(1)_model is based on the assumption of correlated errors while the Seasonal_input_error method is based on the assumption that the input data are not correct and the seasonal input error model accounts for input uncertainty explicitly, which may reduce correlation of modeling errors. The correlation of modeling errors is lower in the Seasonal_input_error in comparison with the Standard calibration method. However, the variance of the standardized residuals is not constant in both the AR(1)_model and the
Figure 6.26: Predictive QQ plots in calibration and validation periods in the Canard River watershed.
Table 6.15: Reliability of streamflow prediction in the Canard River watershed

<table>
<thead>
<tr>
<th>Calibration method</th>
<th>Calibration</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)_model</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>Seasonal_input_error</td>
<td>0.67</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Figure 6.27: QQ plot of standardized residuals in the Canard River watershed
Seasonal_input_error methods (Figure 6.29). The cumulative periodograms with 95% confidence limits (Figure 6.30) for the residuals of the AR(1)_model and the Seasonal_input_error methods show that the residuals are not random in any method. But the non-randomness is high in the Seasonal_input_error method. The cumulative periodogram of residuals of the Seasonal_input_error method is almost similar to that of the Standard calibration method, where no input uncertainty is considered during model calibration. However, the correlation of residual errors is marginally reduced in Seasonal_input_error method (Figure 5.10). The tests of normality, independence and homoscedasticity are performed with the residuals obtained at the maximum posterior density. The application of seasonal input error model and AR(1) model for uncertainty analysis indicates that the model structural uncertainty is high for the Canard River.
Figure 6.29: Test of homoscedasticity of standardized residuals in the Canard River watershed.
Figure 6.30: Cumulative periodogram of residuals with 95% limits in the Canard River watershed
watershed modeling. The uncertainty in precipitation data as quantified by the seasonal input error model is highly variable in different seasons of the Canard River watershed. The uncertainty in model input and structure is partially removed for the watershed modeling by representing the modeling errors with the AR(1) process. However, the assumptions of normality, homoscedasticity and independence of the AR(1)_model residuals are not satisfied. The uncertainty in SWAT model parameter estimation and model prediction estimated by the Seasonal_input_error method for the Canard River watershed is comparable to that of the AR(1)_model.

6.5 Summary

In this chapter, the performance of seasonal input error model is evaluated through its application to the Ruscom River and Canard River watersheds. The seasonal variation of hydrologic and climatic variables are similar in both of the watersheds. The results reveal that the precipitation uncertainty is high in the Canard River watershed modeling while it is low in the Ruscom River watershed modeling. Due to low precipitation uncertainty in the Ruscom River watershed modeling, the Seasonal_input_error method has resulted in parameter and streamflow prediction uncertainty similar to that of the Standard calibration method. In the Standard calibration method, the precipitation data are assumed to be known exactly and no input error model is used in the calibration process. However, in Canard River watershed modeling, the Seasonal_input_error model performs better than the Standard calibration method. In both watersheds, the seasonal input error model shows that the measured precipitation is higher than the estimated precipitation and the estimated precipitation is shown to be independent of the measured precipitation. The higher flow events are underestimated by
the SWAT model in all of the calibration methods for both of the watershed modeling. This is a result of model structural uncertainty in the calibration process.

For evaluating the reliability of streamflow prediction uncertainty quantified by the Seasonal_input_error method, the results are compared to those of the AR(1)_model for both of the watersheds. In the AR(1)_model, the model input uncertainty is accounted for implicitly with other sources of uncertainty in watershed modeling. The results of the AR(1)_model shows that the AR(1) process adequately represents all sources of uncertainties in Ruscom River watershed modeling while the AR(1) process is not enough to describe the uncertainties in a lumped approach for the Canard River watershed modeling. The uncertainty in parameter estimation, the efficiency of optimum parameter values in streamflow prediction and streamflow prediction uncertainty estimated by the Seasonal_input_error model are comparable to those of the AR(1)_model in both watersheds. The modeling errors of Seasonal_input_error method are observed to be heteroscedastic, correlated and non-normal. The reason might be the dominance of model structural uncertainty over input and parameter uncertainties. For any watershed, the modeling errors are lower in the Seasonal_input_error method than in the Standard method and AR(1)_model. However, the observed streamflow data covered by total uncertainty in the methods are similar in both watersheds. In addition, the streamflow prediction uncertainty due to parameter uncertainty is lower than the prediction uncertainty due to total uncertainty in any uncertainty analysis method.

The results discussed in this chapter show that the uncertainty analysis of SWAT model by implementing seasonal input error model in the calibration process is consistent with the uncertainty analysis of implicit methods. Therefore, the seasonal input error
model can be applied to the distributed hydrological modeling for quantifying input uncertainty in the calibration process.

6.6 Conclusions

Based on the findings of this chapter, the following conclusions can be drawn:

- The seasonal input error model shows that the estimated precipitation is less than the measured precipitation for the two case studies.
- If precipitation uncertainty is low, the seasonal input error model based calibration method behaves like a traditional method of calibration.
- The implicit and explicit methods of uncertainty analyses show insignificant differences in model results. Model structural uncertainty dominates over input and parameter uncertainties.
- The reliability of model prediction using the seasonal input error model is similar to that of implicit method.
- Model prediction uncertainty is underestimated by both implicit and explicit methods of calibration.
- There exits multiple local optima in the parameter space. Therefore, any implicit or explicit method could not provide unique solution of model parameters.
- The seasonal input error model exhibits similar performance for the watersheds with similar hydrologic and climatic conditions.
- For further improvement in parameter inferences, model structural uncertainty needs to be accounted for explicitly in the calibration process.
CHAPTER VII

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

Input uncertainty is a major source of uncertainty in any hydrological modeling. This dissertation has focused on precipitation uncertainty in the distributed hydrological modeling. The uncertainty in hydrological modeling is quantified either implicitly by lumping all sources of errors or explicitly by addressing different sources of errors individually. This dissertation has explored the existing implicit and explicit methods of uncertainty analysis for a physically-based distributed hydrological model. Due to the high-dimensionality and computational problems, explicit methods are not used for quantifying uncertainty in distributed hydrological modeling. Moreover, due to the difficulties involved in explaining the effects of explicit error model on parameter estimation and model prediction, the uncertainty in distributed modeling for different errors is usually expressed in terms of parameter uncertainty. To address these issues in uncertainty analysis, a new seasonal input error model has been developed for quantifying precipitation uncertainty explicitly in distributed hydrological modeling. The newly developed model is based on the multiplicative input error model, but uses the season-dependent precipitation multipliers to quantify precipitation uncertainty. The developed methodology has been applied to two watersheds for analyzing uncertainty of watershed modeling. Both watersheds exhibit similar hydrologic and climatic conditions. The performance of the developed method has been evaluated by the estimation of hydrological model parameter uncertainty, input uncertainty and streamflow prediction uncertainty. The efficiency of optimum parameter values in streamflow prediction is also
examined. The assumptions of the statistical error model are verified by using the standard tools of verification. Based on the findings of the dissertation, the following conclusions can be drawn:

1) While applying the implicit methods for quantifying precipitation uncertainty lumped with other sources of uncertainty, this study reveals that the parameter estimation is biased when the Box-Cox transformation of data is adopted in the likelihood function for addressing the non-homogeneity and non-normality problems of residual errors. Therefore, it is essential to verify the estimated parameters based on data transformation in predicting streamflow and other components of a hydrologic system. The possible alternative to data transformation is to use a heteroscedastic error model for formulating the likelihood function.

2) The parameter and prediction uncertainties estimated by the implicit methods are considered reliable when the assumptions of the statistical error model are satisfied. Therefore, the major challenge of the implicit methods is to describe the modeling errors by an appropriate statistical error model and to formulate the likelihood function for parameter inferences.

3) The uncertainty in streamflow prediction due to model parameter uncertainty is reduced when the autoregressive models are used to represent the correlated errors. However, the uncertainty in streamflow prediction due to total uncertainty is wider than that of parameter uncertainty due to the dominance of additive errors over the errors caused by the estimated parameters.

4) The explicit methods of uncertainty analysis for the distributed hydrological modeling is a challenging task. In this dissertation, precipitation uncertainty is accounted
for explicitly in a physically-based distributed model by two existing input error models (daily input error and storm input error) and the seasonal input error model. No inconsistency in model parameter estimation and model prediction is observed in any explicit method of calibration when the results are compared with that of traditional method of calibration. Moreover, when the estimated precipitation uncertainty is very low, the calibration methods based on explicit input error models perform equivalent to the traditional calibration method. This indicates the applicability of the explicit methods of uncertainty analysis to the highly-parameterized distributed hydrological model.

5) The limitations of existing multiplicative input error models led to the development of the seasonal input error model. The seasonal input error model shows that the measured precipitation data is higher than the true precipitation value. This finding is consistent with that of other multiplicative error models. The uncertainty in precipitation data quantified by the seasonal input error model is also comparable to other input error models.

6) The effects of seasonal precipitation multipliers on parameter estimation and model prediction are described by the correlation of estimated model parameters and the reliability of model prediction uncertainty. Thus, the applicability of the seasonal input error model is justified for a distributed hydrological model.

7) The seasonal input error model is capable of quantifying high as well as low precipitation uncertainty. Therefore, the seasonal input error model can be extended to watersheds with different climatic and hydrologic conditions.
8) Even though the modeling errors are correlated, heteroscedastic and non-normal in seasonal input error model based calibration method, the parameter and predictive uncertainties estimated by the method are consistent with that of implicit methods. This has increased the confidence in the results obtained from the seasonal input error model.

9) By keeping the number of precipitation multipliers equal to the number of distinct seasons, the seasonal input error model has reduced the number of latent variables in the Bayesian modeling and thus reduced the high-dimensionality and computational problems of the existing storm-event based multiplier model.

10) This research reveals that the model structural uncertainty is dominant over the input uncertainty and parameter uncertainty in hydrological modeling. Therefore, the prediction uncertainty due to parameter uncertainty is narrow in comparison with the prediction uncertainty due to total errors. This finding is similar to other uncertainty analysis studies.

11) Due to the presence of high model structural uncertainty, there exists non-uniqueness in parameter estimation, especially for the curve number. The autoregressive models as well as the input error models could not remove this problem. Therefore, for further improvement in the parameter inferences, model structural uncertainty needs to be accounted for explicitly in the calibration process.

7.2 Future work

The newly developed seasonal input error model has been evaluated by its application to two watersheds that exhibit similar hydrologic and climatic conditions. To identify the applicability and limitations, the seasonal input error model needs to be
extended to watersheds of different sizes with different hydrologic and climatic conditions.

The future research work needs to be carried out in quantifying precipitation uncertainty due to the representation errors in the distributed hydrological modeling. Further investigation also needs to be made to clarify the distribution of precipitation pattern errors within a season.

For identifying different sources of errors in hydrological modeling, the explicit methods are the most attractive solution. This study has focused on explicit consideration of precipitation uncertainty only in hydrological modeling. The research needs to be extended to the explicit treatment of model structural uncertainty in distributed hydrological modeling.

To extend the seasonal input error model to different types of calibration problems, the efficiency of the seasonal input error model needs to be evaluated for multi-site and multi-objective calibration of a distributed hydrological model.
REFERENCES


## VITA AUCTORIS

<table>
<thead>
<tr>
<th>NAME:</th>
<th>Arpana Rani Datta</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLACE OF BIRTH:</td>
<td>Barisal, Bangladesh</td>
</tr>
<tr>
<td>YEAR OF BIRTH:</td>
<td>1977</td>
</tr>
<tr>
<td>EDUCATION:</td>
<td>Ph. D Candidate, Civil Engineering</td>
</tr>
<tr>
<td></td>
<td>University of Windsor, Windsor, ON, Canada</td>
</tr>
<tr>
<td></td>
<td>M.Sc in Water Resources Engineering (2005)</td>
</tr>
<tr>
<td></td>
<td>Bangladesh University of Engineering &amp; Technology, Bangladesh</td>
</tr>
<tr>
<td></td>
<td>B.Sc in Civil Engineering (2000)</td>
</tr>
<tr>
<td></td>
<td>Bangladesh University of Engineering &amp; Technology, Bangladesh</td>
</tr>
</tbody>
</table>