



**UNIVERSITY OF LEEDS**

**QUANTIFYING UNCERTAINTIES OF MULTI-MODEL CLIMATE  
CHANGE SCENARIOS ON THE WATER CRISIS IN MALAYSIA**

by

Hafizul Aimme bin Che Hamid

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

University of Leeds

School of Civil Engineering

December 2022

## ACKNOWLEDGEMENTS

Along my doctoral degree journey, there are many people have contributed to make it happen. Here, I would love to express my sincere gratitude to all of them for their non-stop support.

I would like to express my deepest gratitude to my supervisors Prof. Martin Tillotson and Dr. Mark Trigg for their kindness, continuous support, and a great guidance along my doctorate journey. They have never failed to support me from the very beginning and provided me with plenty of academic opportunities to accomplish my PhD research work.

I would like to acknowledge to my scholarship donor, Government of Malaysia through Public Service Department, providing financial support throughout the entire PhD program. I am also thankful to School of Engineering, University of Leeds for providing a fund to attend the SHETRAN training in Newcastle University and presenting my work on the SWAT Conference 2022 in Prague.

My deepest thanks to Director General of Department of Irrigation and Drainage Malaysia, Dato' Ir. Dr. Md Nasir bin Md Noh, Senior Director Mejar (K) Dato' Ir. Dr. Ahmad Anuar bin Othman, former Director General Dato' Ir. Nor Hisham bin Mohd Ghazali and Dato' Ir. Abdullah bin Isnin for supporting me in pursuing my PhD. Special thanks also to Deputy Director General of National Water Research Institute of Malaysia (NAHRIM) Ir. Mohd Zaki bin Mat Amin for guidance during my study.

Special thanks to Public Service Department officer, Khairil Azwan bin Abu Mansor, Rais bin Shaari bin Mansor, for their support especially related to the sponsorship. To my best friends, Mohd Sani bin Mat Isa, Mohammad Fikry bin Abdullah, Ahmad Shamsul Irwan bin Mohd Zaki, Muhammad Zaki bin Mashud, Mazwina binti Meor Hamid, Ir. Gs. Ts. Herman bin Tawil, Ir. Ts. Dr. Tajul Ariffin bin Norizan, Nur Wazni binti Mohammad, Mohd Syhrizan bin Mat Ghani, Noor Hidayah binti Md Ibrahim, Zulraimie bin Abdul Jabar, and Isyamuddin bin Ahmad for always morally supportive to me directly or indirectly from the starts until complete of my study even from a long distance.

Finally, I am really grateful to my family from the bottom of my heart. I owe a lot to my mother, Hasnah binti Zakaria, my father, Che Hamid bin Che Hassan, my lovely wife, Yusniza binti Halim and to all my lovely kids, Nur Aisyah Qalesya, Muhammad Haziq Rifqi, Nur Aliya Qaisara, Muhammad Hazim Rizqi, Muhammad Haider Ruzain, and Nur Aina Safiyya for their uncountable sacrifices. No words can describe how I feel gratitude to them. Their constant love, support and encouragement have always been there for me. I want to offer special thanks to my mother and father-in-law who always pray for my success.

## ABSTRACT

Malaysia has a relatively limited capacity to deal with the effects of climate change while being one of the most vulnerable nations to its effects. As a developing country, the lack of a consistent temporal and spatial data source has always been an issue, and the region is also considered data-scarce. This study's primary goal is to evaluate the effects of climate change on Malaysia's water resources, particularly the Selangor River Basin (SRB). Instead of using a single source input dataset, cross-combined datasets from multiple sources were used in order to optimise the hydrological model. Five input variables, including precipitation, temperature, solar radiation, relative humidity, and wind speed, were used to define seven scenarios using single and cross-combined method. To improve the hydrological model multi-site calibration method is employed. Finally, climate change prediction data from several Global Climate Models (GCMs) is utilised to assess the effects of climate change on SRB water supplies. The CFSR and CMADS global reanalysis datasets show a highly significant relationship on precipitation, with an *r-value* of 0.81 for both datasets. However, for temperature data, CMADS surpasses CFSR on maximum and minimum temperatures, with 0.6 and 0.7, respectively. In the SWAT model, most of the scenarios achieved a 'good' performance range on the calibration and validation processes. However, SWAT model with CFSR as input data achieved an 'unsatisfactory' range with  $R^2$  of 0.35, NSE of 0.16, Pbias of 0.00, KGE of 0.50, and RSR of 0.92. For a cross-combined approach, the result shows the combination of the observed and CMADS datasets performed better than the combination of the observed and CFSR datasets. The sequential technique outperformed the simultaneous and basin-by-basin techniques by achieving 'satisfactory' range at all outlets. The SRB's assessment of climate change predicted an increase in precipitation and temperature from 2030 to 2050. Climate data from 'ensemble average' realisation predicted SRB would receive a huge amount of precipitation in November and April every year, and high temperatures from February to June. Additionally, a few sub-basins are expected to have water availability greater than 5 m<sup>3</sup>/s for three consecutive years.

# Contents

|   |     |
|---|-----|
| ACKNOWLEDGEMENTS .....  | i   |
| ABSTRACT .....  | iii |
| LIST OF FIGURES.....  | iv  |
| LIST OF TABLES .....  | xii |
| LIST OF ABBREVIATIONS.....                                      | xiv |
| CHAPTER 1 INTRODUCTION.....                                     | 1   |
| 1.1 Background .....  | 1   |
| 1.2 Research Gap Analysis.....                                  | 3   |
| 1.3 Research Questions.....                                     | 4   |
| 1.4 Project Aims and Objectives .....                           | 5   |
| 1.5 The scope of the study.....                                 | 5   |
| 1.6 Limitations of the study .....                              | 6   |
| 1.7 Dissertation structure .....                                | 7   |
| 2 CHAPTER 2 LITERATURE REVIEW .....                             | 8   |
| 2.1 Global satellite/reanalysis data .....                      | 8   |
| 2.1.1 Climate Forecast System Reanalysis.....                   | 10  |
| 2.1.2 China Meteorological Assimilation Driving Dataset.....    | 11  |
| 2.2 Data performance .....                                      | 12  |
| 2.3 Cross-combined dataset .....                                | 14  |
| 2.4 Regionalisation.....  | 19  |
| 2.5 Hydrological models.....                                    | 20  |
| 2.6 Soil and Water Assessment Tool .....                        | 21  |
| 2.7 Model calibration and validation.....                       | 23  |
| 2.7.1 Multi-site calibration.....                               | 24  |
| 2.7.2 SWAT Calibration and Uncertainty Program Software.....    | 25  |
| 2.8 Climate change .....  | 26  |
| 2.8.1 Impacts of climate change.....                            | 26  |
| 2.9 General Circulation Model.....                              | 29  |
| 2.10 Regional Climate model .....                               | 30  |
| 2.10.1 Regional Hydro Climate Model of Peninsular Malaysia..... | 31  |
| 2.11 Ensemble approach .....                                    | 33  |
| 2.12 Conclusion .....   | 35  |
| 3 CHAPTER 3 STUDY AREA AND DATA.....                            | 36  |
| 3.1 Malaysia.....   | 36  |
| 3.1.1 Geography and Demography Characteristic .....             | 37  |

|        |   |     |
|--------|---|-----|
| 3.1.2  | Population.....   | 38  |
| 3.1.3  | Climate .....   | 40  |
| 3.2    | State of Selangor .....   | 40  |
| 3.3    | Selangor River Basin (SRB).....   | 42  |
| 3.3.1  | Dam and reservoir .....   | 43  |
| 3.3.2  | Surface water contribution zones .....  | 45  |
| 3.4    | Water Resources .....   | 46  |
| 3.4.1  | Water Resources in Malaysia.....  | 47  |
| 3.5    | Water Demand in Malaysia .....  | 48  |
| 3.5.1  | Water Demand in Selangor .....  | 50  |
| 3.6    | Water Crisis in Selangor .....  | 51  |
| 3.7    | Hydrological data .....   | 53  |
| 3.8    | Meteorological data.....  | 56  |
| 3.9    | Global satellite/reanalysis data .....  | 58  |
| 3.9.1  | Climate Forecast System Reanalysis (CFSR).....                                  | 58  |
| 3.9.2  | China Meteorological Assimilation Driving Datasets for SWAT Model (CMADS) ..... | 59  |
| 3.10   | Future climate data .....   | 60  |
| 3.11   | Landscape features.....   | 61  |
| 3.11.1 | Land use .....  | 61  |
| 3.11.2 | Soil Type.....  | 62  |
| 3.11.3 | Digital Elevation Model (DEM).....  | 64  |
| 3.12   | Conclusion .....  | 65  |
| 4      | CHAPTER 4 METHODOLOGY.....  | 66  |
| 4.1    | Overall methodology .....   | 66  |
| 4.2    | Analysis input data performance.....  | 68  |
| 4.2.1  | Data selection.....   | 68  |
| 4.2.2  | Evaluation methods .....  | 69  |
| 4.2.3  | Verification method.....  | 72  |
| 4.3    | Investigate of cross-combine dataset.....                                       | 74  |
| 4.3.1  | Data preparation.....   | 74  |
| 4.3.2  | Hydrological model setup .....  | 75  |
| 4.3.3  | Model Calibration.....  | 80  |
| 4.3.4  | Model Validation .....  | 84  |
| 4.4    | Multi-site calibration evaluation .....   | 86  |
| 4.5    | Analysis multi-site calibration technique .....                                 | 87  |
| 4.6    | Climate change impact assessment .....  | 88  |
| 4.6.1  | Regional Hydro Climate Model for Peninsular Malaysia .....                      | 90  |
| 4.6.2  | Ensemble data of GCMs model.....  | 91  |
| 4.6.3  | Climate simulation .....  | 91  |
| 5      | CHAPTER 5 DATASETS PERFORMANCE ANALYSIS .....                                   | 93  |
| 5.1    | Evaluation of global reanalysis datasets.....                                   | 93  |
| 5.1.1  | Data screening .....  | 94  |
| 5.1.2  | Double mass curve.....  | 96  |
| 5.1.3  | Point-to-point analysis .....   | 98  |
| 5.2    | Analysis on cross-combined datasets.....  | 104 |
| 5.2.1  | Model simulation.....   | 107 |

|       |   |     |
|-------|---|-----|
| 5.2.2 | Sensitivity analysis .....                                      | 108 |
| 5.2.3 | Model evaluation .....  | 109 |
| 5.3   | Summary of the results .....                                    | 113 |
| 6     | CHAPTER 6 MULTI-SITE CALIBRATION ASSESSMENT .....               | 115 |
| 6.1   | Single-site calibration (SSC) method .....                      | 116 |
| 6.1.1 | SSC model performance .....                                     | 116 |
| 6.2   | Multi-site calibration (MSC) method .....                       | 117 |
| 6.2.1 | MSC model performance.....                                      | 117 |
| 6.3   | Model performances and uncertainty analysis.....                | 136 |
| 6.3.1 | Summary of the results.....                                     | 141 |
| 7     | CHAPTER 7 CLIMATE CHANGE ASSESSMENT .....                       | 143 |
| 7.1   | First Generation Coupled General Circulation Model (CGCM1)...   | 144 |
| 7.2   | Integrated Regional-Scale Hydrologic/Atmospheric Model (IRSHAM) | 145 |
| 7.3   | The Fifth-Generation Mesoscale Model (MM5).....                 | 146 |
| 7.4   | Ensemble approach .....   | 146 |
| 7.5   | Regional Hydro Climate Model of Peninsular Malaysia .....       | 148 |
| 7.6   | Simulation on future changes in hydrological components.....    | 152 |
| 7.6.1 | Hydrological model simulation.....                              | 154 |
| 7.7   | Summary of the results .....                                    | 172 |
| 8     | CHAPTER 8 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS              | 174 |
| 8.1   | Summary.....  | 174 |
| 8.1.1 | Input data.....   | 174 |
| 8.1.2 | Calibration and validation techniques.....                      | 175 |
| 8.1.3 | Climate change assessment .....                                 | 176 |
| 8.2   | Conclusions.....  | 178 |
| 8.3   | Contributions of this study.....                                | 179 |
| 8.4   | Recommendations .....   | 180 |
| 8.4.1 | Based on the study results .....                                | 180 |
| 8.4.2 | For future research .....                                       | 180 |
|       | REFERENCES.....   | i   |
|       | APPENDIX A – DOUBLE MASS CURVE ANALYSIS.....                    | i   |

## LIST OF FIGURES

| <u>Figure</u> | <u>Title</u>  | <u>Page</u> |
|---------------|---|-------------|
| Figure 2-1:   | simulation results (a) observation station; (b) CMADS; (c) Observation+CMADS in the Hailiutu River basin during the period 2008–2014 (source: Shao et. al., (2018)).....  | 16          |
| Figure 2-2:   | Statistical figures of the simulation results at the daily scale for (a) precipitation; (b) temperature; (c) a solar radiation data input scenario of the three datasets (GD, CMADS and CFSR) (source: Gu et. al., (2020)).....   | 17          |
| Figure 2-3:   | Streamflow simulation model result using three different precipitations (a) GD; (b) CMADS; (c) CFSR, and (d) output analysis of water balance for models (source: Gu et al., (2020)).....   | 18          |
| Figure 2-4:   | Streamflow simulation model result using three different temperature (a) GD; (b) CMADS; (c) CFSR, and (d) output analysis of water balance for models (source: Gu et al., (2020)).....  | 19          |
| Figure 2-5:   | SWAT development history and model application (source: Gassman et. al., (2007)).....   | 22          |
| Figure 2-6:   | Three-dimensional grid of General Circulation Model (GCM) (source: <a href="https://www.ipcc-data.org/">https://www.ipcc-data.org/</a> ).....   | 30          |
| Figure 2-7:   | Schematic depiction of Regional Climate Modelling and the applications (source: Giorgi, (2019)) .....   | 31          |
| Figure 2-8:   | The modelling domains of Peninsular Malaysia hydro climate model: The green box covered by 54 km x 54 km grids is the larger outer domain; the orange box covered by 18 km x 18 km grids is the intermediate domain; and the red box is covered by 6 km x 6 km grids was the inner modelling domain (source: Amin et. al., (2019))..... | 32          |
| Figure 3-1:   | Map of Malaysia ( <i>source: google website</i> ).....  | 37          |
| Figure 3-2:   | Map of Malaysia in global view (source: <a href="https://site.ieee.org/">https://site.ieee.org/</a> )...  | 38          |

|  |    |
|--|----|
| Figure 3-3: Population projections for Peninsular Malaysia (source: NWRS (2011)).....  | 39 |
| Figure 3-4: Population projections for East Malaysia (source: NWRS (2011))<br>.....  | 39 |
| Figure 3-5: Map of state of Selangor (source: Yahya (2021)).....   | 41 |
| Figure 3-6: Selangor River Basin boundary catchment (source: LUAS (2015))<br>.....   | 43 |
| Figure 3-7: Location map of (a) Peninsular Malaysia, (b) State of Selangor and<br>(c) Selangor River basin.....                          | 44 |
| Figure 3-8: Selangor river basin surface water contribution zones (source:<br>state of the river report (2015)).....                     | 45 |
| Figure 3-9: Distribution of Earth’s Water (source: Gleick, P. H., (1996)).....   | 46 |
| Figure 3-10: Water demand for all sector compared with the available runoff<br>(source: NWRS (2011)).....                                | 49 |
| Figure 3-11: Water demand for various sectors in Selangor (source: NWRS<br>(2011)).....  | 51 |
| Figure 3-12: Drought classification system. (source: (Hasan et al., 2019))..   | 52 |
| Figure 3-13: Hydrological stations in Sg. Selangor basin (source: DID<br>Malaysia).....  | 54 |
| Figure 3-14: Rainfall pattern of Selangor River Basin from 1971–2015 (source:<br>LUAS (2011)) .....                                      | 55 |
| Figure 3-15: SRB annual average rainfall from 2008-2018.....   | 56 |
| Figure 3-16: Meteorological stations in SRB .....  | 57 |
| Figure 3-17: Mean annual (a) maximum and (b) minimum temperature at the<br>Tennamaran Estate meteorological station from 2012-2019 ..... | 57 |

|  |     |
|--|-----|
| Figure 3-18: Land use (2013) in the Sg. Selangor watershed.....  | 61  |
| Figure 3-19: Local soil series in the Sg. Selangor watershed .....   | 63  |
| Figure 3-20: Sg. Selangor elevation by MERIT DEM.....  | 64  |
| Figure 4-1: Flowchart of an overall procedure.....   | 67  |
| Figure 4-2: Location map of hydro-meteorological stations and gridded point<br>of global reanalysis datasets in Selangor River Basin .....     | 69  |
| Figure 4-3: Microsoft excel macro-enable platform to generate the input<br>database in Microsoft Access database.....                          | 75  |
| Figure 4-4: Flowchart of SRB SWAT model.....   | 77  |
| Figure 4-5: Calibration techniques implemented in SRB SWAT model for (a)<br>single-site and (b) multi-site .....                               | 87  |
| Figure 4-6: Multi-site calibration method used on SRB SWAT model.....  | 88  |
| Figure 4-7: Schematic description of the climate modelling approach .....  | 90  |
| Figure 5-1: Missing data analysis for temperature data record from 2008-2019<br>.....  | 94  |
| Figure 5-2: Monthly temperature data (2011-2016) after using the linear<br>interpolation method to fill in the missing data .....              | 95  |
| Figure 5-3: Double Mass Curve analysis on 2008-2019 precipitation data..   | 96  |
| Figure 5-4: Double Mass Curve analysis on 2008-2019 streamflow data....  | 97  |
| Figure 5-5: Point of gridded data from CFSR, and CMADS.....  | 98  |
| Figure 5-6: Scatter plots on monthly CMADS and CFSR to gauged stations:<br>(a) Precipitation; (b) Max. temperature; (c) Min. temperature ..... | 100 |
| Figure 5-7: Monthly precipitation data analysis on CMADS and CFSR to<br>gauged station.....  | 102 |

|  |     |
|--|-----|
| Figure 5-8: Bias correction on CMADS precipitation data.....   | 102 |
| Figure 5-9: Bias correction on CFSR precipitation data.....  | 103 |
| Figure 5-10: Bias correction on CMADS temperature data.....  | 103 |
| Figure 5-11: Bias correction on CFSR temperature data.....   | 104 |
| Figure 5-12: SWAT database prepared using Microsoft Access for 7 scenarios<br>.....                    | 105 |
| Figure 5-13: SRB sub-basins delineated using SWAT model.....   | 106 |
| Figure 5-14: 171 Hydrological response units (HRUs).....   | 106 |
| Figure 5-15: SRB single-site calibration and validation method.....                                    | 107 |
| Figure 5-16: Global sensitivity analysis for SRB SWAT model parameters<br>.....                        | 109 |
| Figure 5-17: Single-site calibration results on seven scenario datasets ....                           | 110 |
| Figure 5-18: Single-site validation results on seven scenario datasets.....                            | 111 |
| Figure 5-19: Simulated streamflow on CMADS individual and cross-combined<br>datasets.....              | 112 |
| Figure 6-1: SRB calibration and validation method.....   | 116 |
| Figure 6-2: multi-site calibration for basin-by-basin technique .....                                  | 119 |
| Figure 6-3: Calibrated and validated results using basin-by-basin technique in<br>basin-1 (BB:B1)..... | 121 |
| Figure 6-4: Calibrated and validated results using basin-by-basin technique in<br>basin-3 (BB:B3)..... | 122 |
| Figure 6-5: Calibrated and validated results using basin-by-basin technique in<br>basin-3.....         | 123 |
| Figure 6-6: multi-site calibration for simultaneous technique .....                                    | 125 |

|  |     |
|--|-----|
| Figure 6-7: simulated graph on simultaneous techniques at outlet-1.....  | 126 |
| Figure 6-8: simulated graph on simultaneous techniques at outlet-12.....   | 127 |
| Figure 6-9: the best validated streamflow results from each simultaneous technique at outlet-1 .....   | 128 |
| Figure 6-10: validated results for simultaneous techniques at outlet-12 ....   | 129 |
| Figure 6-11: validated result of simultaneous technique at outlet-4.....   | 130 |
| Figure 6-12: multi-site calibration approach in sequential method.....   | 131 |
| Figure 6-13: calibrated results for sequential techniques at outlet-12.....  | 132 |
| Figure 6-14: calibrated results for sequential techniques at outlet-1 .....  | 133 |
| Figure 6-15: validated result of sequential technique at outlet-12.....  | 134 |
| Figure 6-16: validated result of sequential technique at outlet-1 .....  | 135 |
| Figure 6-17: validated result of sequential technique at outlet-4.....   | 136 |
| Figure 6-18: single-site and multi-site validation results at outlet-1.....  | 139 |
| Figure 6-19: multi-site validation result at outlet-4.....   | 140 |
| Figure 6-20: multi-site validation results at outlet-12.....   | 140 |
| Figure 7-1: CGCM1 grid data used in the RegHCM-PM. Blue represents the ocean grids. While green represents the land grids. (source; Chen et al., (2006)).....  | 145 |
| Figure 7-2: The modelling domains of the Peninsular Malaysia hydroclimate model: The green box covered by 54 km x 54 km grids is the larger outer domain; the orange box covered by 18 km x 18 km grids is the intermediate domain; and the red box is covered by 6 km x 6 km grids, which is the inner modelling domain. (source: Amin et al., (2019))..... | 150 |
| Figure 7-3: Schematic description of the modelling approach .....  | 151 |

|   |     |
|---|-----|
| Figure 7-4: RegHCM-PM grid point downscale at 6 km spatial resolution in the SRB. Purple dots represent the centroid of the grid cell. ....   | 152 |
| Figure 7-5: Annual precipitation on individual projections of the ensemble data .....   | 153 |
| Figure 7-6: Annual temperature on an individual projection of the ensemble data .....   | 154 |
| Figure 7-7: Annual mean flow projections in the Muda watershed for (a) individual and (b) ensemble averages using a 10-year moving average and a 95% confidence band. (source: Amin et al., (2017)) ..... | 155 |
| Figure 7-8: Linear trend analysis of precipitation for SRB using the ‘ensemble average’ .....   | 156 |
| Figure 7-9: Annual rainfall variation of ‘ensemble average’ from 2030-2050 using the Mann-Kendall test with Sen’s slope.....  | 156 |
| Figure 7-10: Precipitation projection from the ‘ensemble average’ realisation .....   | 157 |
| Figure 7-11: Rainfall Anomaly Index (RAI) prediction on the ‘ensemble average’ realisation. ....  | 158 |
| Figure 7-12: Linear trend analysis of average temperature for the SRB using SRES ‘ensemble average’ .....   | 159 |
| Figure 7-13: Average temperature variation on ‘ensemble average’ from 2030-2050 using the Mann-Kendall test with Sen’s slope.....   | 159 |
| Figure 7-14: Temperature projection from the ‘ensemble average’ realisation .....   | 160 |
| Figure 7-15: Average monthly precipitation from 2030 to 2050 on the ‘ensemble average’ realisation.....   | 160 |
| Figure 7-16: Average monthly temperature from 2030 to 2050 on the ‘ensemble average’ realisation.....   | 161 |

|  |     |
|--|-----|
| Figure 7-17: Water availability in the SRB sub basin using the ‘ensemble average’ realisation. ....                        | 162 |
| Figure 7-18: Mean monthly flow simulation on ‘ensemble average’ datasets for year 2030, 2040 and 2050 at outlet-1. ....    | 163 |
| Figure 7-19: Mean monthly flow simulation on ‘ensemble average’ datasets for year 2030, 2040 and 2050 at outlet-4. ....    | 163 |
| Figure 7-20: Mean monthly flow simulation on ‘ensemble average’ datasets for year 2030, 2040 and 2050 at outlet-12. ....   | 164 |
| Figure 7-21: Linear trend analysis of precipitation for SRB using the SRES ccsm3a1fi. ....                                 | 164 |
| Figure 7-22: Annual rainfall variation of ‘ccsm3a1fi’ from 2030-2050 using the Mann-Kendall test with Sen’s slope .....    | 165 |
| Figure 7-23: Precipitation projection from SRES ccsm3a1fi realisation .....  | 165 |
| Figure 7-24: Rainfall Anomaly Index (RAI) prediction on the ‘ccsm3a1fi’ realisation. ....                                  | 166 |
| Figure 7-25: Linear trend analysis of temperature for the SRB using SRES ccsm3a1fi .....                                   | 167 |
| Figure 7-26: Average temperature variation on ‘ccsm3a1fi’ from 2030-2050 using the Mann-Kendall test with Sen’s slope..... | 167 |
| Figure 7-27: Temperature projection based on SRES ccsm3a1fi implementation. ....   | 168 |
| Figure 7-28: Monthly precipitation averages from 2030 to 2050 based on the SRES ccsm3a1fi realisation. ....                | 169 |
| Figure 7-29: Monthly average temperature on the SRES ccsm3a1fi realisation from 2030 to 2050. ....                         | 169 |

Figure 7-30: Water availability in the SRB sub basin using the SRES ccsm3a1fi realisation..... 170

Figure 7-31: Mean monthly flow simulation on SRES ccsm3alfi datasets for the years 2030, 2040 and 2050 at outlet-1..... 171

Figure 7-32: Mean monthly flow simulation on SRES ccsm3alfi datasets for the years 2030, 2040 and 2050 at outlet-4..... 171

Figure 7-33: Mean monthly flow simulation on SRES ccsm3alfi datasets for the years 2030, 2040 and 2050 at outlet-12..... 172

## LIST OF TABLES

| <b><u>Table</u></b> | <b><u>Title</u></b>  | <b><u>Page</u></b> |
|---------------------|--|--------------------|
| Table 3-1:          | Total consumptive water demand against total surface water availability for all sectors (source: NWRS (2011)).....   | 48                 |
| Table 3-2:          | Total water demand for various sectors in Selangor (source: NWRS (2011)).....  | 50                 |
| Table 3-3:          | Total water demand in state of Selangor (mm rainfall per year) (source: NWRS (2010)).....                            | 53                 |
| Table 3-4:          | List of hydrological station in Sg. Selangor basin (source: DID Malaysia).....                                       | 54                 |
| Table 3-5:          | Special Report on Emission Scenarios (SRES) of 15 different future climate realisations. (source: (IPCC, 2014))..... | 60                 |
| Table 3-6:          | Sg. Selangor land use information .....  | 62                 |
| Table 3-7:          | Sg. Selangor soil series area .....  | 63                 |
| Table 4-1:          | Cross-combine input data .....   | 74                 |
| Table 4-2:          | SWAT parameters used by other researchers.....   | 79                 |
| Table 5-1:          | Double Mass Curve on yearly rainfall data from 2008-2019.....  | 97                 |
| Table 5-2:          | the point-to-point statistical analysis of CMADS and CFSR.....   | 101                |
| Table 5-3:          | List of parameters used in the calibration process.....  | 108                |
| Table 5-4:          | Calibrated and validated results on single-site method (SSC)..   | 110                |
| Table 6-1:          | Selection criteria for MSC analysis .....  | 118                |
| Table 6-2:          | multi-site calibrated results using the basin-by-basin technique. ....   | 120                |

|  |     |
|--|-----|
| Table 6-3: multi-site validated results using the basin-by-basin technique.  | 120 |
| Table 6-4: Fitted parameters in calibrated model basin-2 and regionalisation method from basin-1 and basin-3.....              | 124 |
| Table 6-5: multi-site calibrated results using the simultaneous technique.   | 125 |
| Table 6-6: multi-site validated results using the simultaneous technique. .  | 127 |
| Table 6-7: multi-site calibrated result using the sequential technique.....  | 131 |
| Table 6-8: multi-site validated result for sequential technique. ....  | 133 |
| Table 6-9: Single-site and multi-site validation results .....   | 137 |
| Table 6-10: Minima range of performance rating and uncertainty analysis  | 138 |
| Table 6-11: SRB SWAT model parameters .....  | 142 |
| Table 7-1: Special Report on Emission Scenarios (SRES) of 15 different future climate realisations. source: (IPCC, 2014) ..... | 147 |

## LIST OF ABBREVIATIONS

|           |   |
|-----------|---|
| AR4       | Fourth assessment report of IPCC                                    |
| AR5       | Fifth assessment report of IPCC                                     |
| AR6       | Six assessment report of IPCC                                       |
| BB        | Basin-by-Basin  |
| CARD      | Center for Agricultural and Rural Development                       |
| CAU       | China Agricultural University                                       |
| CC        | Climate Change  |
| CCCma     | Canadian Center for Climate Modelling and Analysis                  |
| CFSR      | Climate Forecast System Reanalysis                                  |
| CHIRP     | Climate Hazards Group Infrared Precipitations                       |
| CLDAS     | China Meteorological Administration Land Data Assimilation System   |
| CMADS     | China Meteorological Assimilation Driving Dataset                   |
| CMIP5     | Coupled Model Intercomparison Project 5                             |
| CMORPH    | Climate Prediction Center Morphing technique                        |
| CREAMS    | Chemicals, Runoff, and Erosion from Agricultural Management Systems |
| DEM       | Digital elevation model   |
| DID       | Department of Irrigation and Drainage                               |
| DMC       | Double Mass Curve   |
| DOSM      | Department of Statistics Malaysia                                   |
| EPIC      | Environmental Policy Integrated Climate Model                       |
| ENSO      | El Nino-Southern Oscillation  |
| FAO       | Food and Agriculture Organization of the United Nations             |
| GCM       | General Circulation Model   |
| GCP       | Gridded climate products  |
| GDP       | Gross domestic product  |
| GIS       | Geographical Information System                                     |
| GLEAMS    | Groundwater Loading Effects of Agricultural Management Systems      |
| GLUE      | Generalised Likelihood Uncertainty Estimation                       |
| GSMaP     | Global Satellite Mapping of Precipitation                           |
| HRU       | Hydrological response unit  |
| HWSD      | Harmonized World Soil Database                                      |
| IAHS      | International Association of Hydrological Sciences                  |
| IMERG     | Integrated Multi-Satellite Retrievals for GPM                       |
| IOD       | Indian Ocean Dipole   |
| IPCC      | Intergovernmental Panel on Climate Change                           |
| IRSHAM    | Integrated Regional-Scale Hydrologic/Atmospheric Model              |
| JRB       | Johor River Basin   |
| KGE       | King-Gupta Efficiency   |
| KRB       | Kelantan River Basin  |
| LUAS      | Lembaga Urus Air Selangor   |
| MAE       | Mean absolute error   |
| MCM       | Million cubic metre   |
| MCMC      | Monte Carlo Markov Chain  |
| MERIT-DEM | Multi-Error-Removed Improved-Terrain DEM                            |
| MIKE SHE  | MIKE system Hydrological European Model                             |

|                |   |
|----------------|---|
| MJO            | Madden-Julian Oscillation   |
| MLD            | Million litre per day   |
| MMD            | Malaysia Meteorological Department  |
| MNRE           | Ministry of Natural Resource and Environmental  |
| MOA            | Ministry of Agriculture   |
| MPI            | Max Planck Institute for Meteorology  |
| MRB            | Muda River Basin  |
| MSC            | Multi site calibration  |
| NAHRIM         | National Water Research Institute of Malaysia   |
| NC2            | Second National Communication   |
| NCEP           | National Centers for Environmental Prediction   |
| NSE            | Nash–Sutcliffe Efficiency   |
| NWRS           | National Water Resource Study   |
| Pbias          | Percentage error  |
| ParaSol        | Parameter Solution  |
| PERSIANN       | Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network |
| PSO            | Particle Swarm Optimisation   |
| PUB            | Prediction in Ungauged Basins   |
| r              | Correlation Coefficient   |
| R <sup>2</sup> | Correlation of Determination  |
| RBMP           | River Basin Management Plan   |
| RCM            | Regional climate model  |
| RCP            | Representative Concentration Pathways   |
| RegHCM         | Regional Hydro Climate Model  |
| RMSE           | Root Mean Square Error  |
| RSR            | RMSE-observations standard deviation ratio  |
| SE             | Sequential  |
| SM             | Simultaneous  |
| SRB            | Selangor River Basin  |
| SRES           | Special Report on Emission Scenarios  |
| SSC            | Single site calibration   |
| SUFI-2         | Sequential Uncertainty Fitting  |
| SWAT           | Soil and Water Assessment Tool  |
| SWAT-CUP       | SWAT Calibration and Uncertainty programme software                                       |
| Tmax           | Maximum Temperature   |
| Tmin           | Minimum Temperature   |
| TNC            | Third National Communication report   |
| UCAR           | University Corporation for Atmospheric Research   |
| USDA-ARS       | United States Department of Agriculture-Agricultural Research Service                     |
| TRMM           | Tropical Rainfall Measuring Mission   |
| WPP            | World Population Projection   |
| 96PPU          | 95 percent prediction uncertainty   |

# CHAPTER 1

## INTRODUCTION

### 1.1 Background

Water resources are an essential natural resource for all life on Earth. There are many diverse types of water resources, but access to freshwater is crucial for human survival. Water is utilised for environmental, agricultural, industrial, domestic, and recreational purposes. Water resources are becoming increasingly susceptible globally as a result of increased demand for water brought on by population development, industry expansion, increased food production, pollution from numerous human activities, climate change, and effects on land use (Carroll et al., 2013; Odusanya et al., 2019). In total global water composition, about 97% is saltwater and the balance 3% is freshwater. From 3% of the freshwater, two thirds or 68.7% is categorised as glaciers and ice caps, 30.1% is groundwater, 0.9% others and only 0.3% is fresh surface water. Despite 0.3% of fresh surface water, 87% is from a lake, 11% swamps and only 2% from rivers (Du Plessis, 2017). The modest amounts of available freshwater can have a significant impact on living things if it is not being monitored and managed properly. Most countries in the world rely on rivers for their water supply. According to the United Nations' 2018 World Water Development report, the world's water consumption is expected to increase by about one-third by 2050. More than 2 billion people worldwide will be affected by it due to the continuously increasing global population (Azoulay & Hougbo, 2018). For future sustainability, this projection highlights how important it is to manage and monitor water resources.

Climate change significantly affects water resources because of the close relationship between the hydrological cycle and climate. It changes the patterns of evapotranspiration and precipitation (Ghulami, 2018; Tsanis et al., 2011). This is seen in the shifting patterns of water supply, with diminishing glaciers and altered patterns of precipitation raising the risk of drought and flood. According to IPCC (2014), all assessed emission scenarios projected an increase in surface temperatures, and large areas of Europe, Asia, and Australia will experience more frequent and prolonged heat waves, while many other regions will experience more intense and frequent extreme precipitation events. Climate variability and changes in land use are the two primary causes of the hydrological cycle altering. According to Kundzewicz et al., (2008), the key factor influencing changes in streamflow volume, peak flow, and flow routing time is climate variability. Changes in land use, on the other hand, can affect annual mean discharge, baseflow, flood frequency, and surface runoff (Huntington, 2006). As suggested by Farsani et al., (2018), research on the impacts of climate change on water resources is essential for the sustainable management of water resources for long-term planning.

The implications of climate change on water resources, water resource planning, and watershed management are commonly evaluated using hydrological models (Bai et al., 2017). Hydrological models help with understanding, predicting, and managing water resources by condensing complex real-world processes into computerised applications. Even though model accuracy may be questioned, modelling applications are the greatest answer to time-consuming factors (Rivas-Tabares et al., 2019). In reality, when provided particular information about the catchment region, hydrological models may be used as important decision support system tools for sustainable water resource management and affordable procedures (Combalicer et al., 2010; Odusanya et al., 2019). Studying streamflow with scientific methodologies and procedures can aid in identifying environmental issues such as soil degradation, land use changes, and climate change (Ang & Oeurng, 2018). However, because of a lack of on-the-ground observations and inadequate or low-quality data, hydrological modelling in data-scarce regions is a difficult endeavour (Odusanya et al., 2019).

## 1.2 Research Gap Analysis

Data scarcity (i.e., ungauged watersheds) has always been a major issue when developing good hydrological models and limits the ability to assess and understand local hydrology. It is a fundamental problem faced by hydrology and water resources management to predict hydrologic responses, especially on ungauged watersheds (Farmer, 2016; Sivapalan et al., 2003). Meteorological data exhibits a respectable degree of geographical and temporal precision when included in the hydrological model, and it has a significant impact on the local weather. Unfortunately, there are only a limited number of meteorological stations available (Duan et al., 2019). For the hydrological community, it is tough to determine the proper response to a threat to the water resources and water environment. Prediction accuracy and reliability are becoming increasingly important in order to take the necessary steps to prevent and manage the water resources sustainably, to prevent and manage natural disasters, and to manage ecosystems around us (Sivapalan et al., 2003). Nowadays, several methods are available for predicting basin reactions such as methods suitable for ungauged basins including extrapolating response data from gauged to ungauged basins, remote sensing measurements, applications of process-based hydrological models where the climate inputs are either specified or measured, and application of combined meteorological-hydrological models without the requirement of specifying precipitations inputs (Sivapalan et al., 2003). The International Association of Hydrological Sciences (IAHS) started an effort in 2013 to look into potential solutions for Prediction in Ungauged Basins (PUB). This initiative was designed to concentrate on a better understanding of climatic and landscape elements that regulate hydrological processes occurring at all scales, to enhance the capability to forecast water fluxes in ungauged/data-scarce basins, and to be able to better predict their uncertainties (Hrachowitz et al., 2013). According to Bloschl, (2016), the most precise method of establishing such estimations at any specific place is to measure them over a prolonged period. However, this is not feasible due to logistical and financial constraint, or even just because one is interested in how the hydrological variables will develop in the future.

Thus, for a country like Malaysia, limited understanding about climate change impacts and lack of information from ground observations shows there is a need to be addressed. For example, in the IPCC (Intergovernmental Panel on Climate Change) AR4 WGII report, there were 75 research conducted to examine the effects of global warming on physical and biological systems, however none of these studies originated in Southeast Asia (IPCC, 2007b). In reality, the IPCC AR4 report does not specifically cover climate change in Malaysia because there is little published information about it in this region (T. Tangang et al., 2012). NWRS (2011), stated that due to the unpredictability of rain and how climate change affects the weather, the country needs to act quickly to save water and find and develop other sources of water.

### **1.3 Research Questions**

In this research, the following questions will be answered.

- i. In a hydrological model, how well do observed station datasets and global satellite/reanalysis datasets work?
- ii. Which datasets best perform in a hydrological model, individually or cross-combined?
- iii. How does the multi-site calibration method improve an ungauged watershed compared to the single calibration method?
- iv. Which calibration techniques in multi-site perform the best when multiple datasets (observed station, global satellite/reanalysis data, and cross-combined) are used?
- v. What are the existing and prediction of the climate change impacts scenarios on water resources utilising multi-model datasets with multi-site calibration modelling approach?

## 1.4 Project Aims and Objectives

This study aims to assess the impacts of climate change on water resources in Malaysia by employing a rainfall-runoff model that is driven by multi-model climate datasets and calibrated using a multi-site calibration method, as well as addressing an ungauged sub-catchment by using cross-combined datasets as input data. To achieve these aims, the specific objectives of this study are found as follows:

- I. To evaluate the performance of global satellite/reanalysis datasets to observed station.
- II. To investigate the application of cross-combine dataset from multiple sources in hydrological model
- III. To evaluate the performance of single-site calibration and multi-site calibration methods
- IV. To analyse the performance of basin-by-basin, simultaneous and sequential calibration techniques.
- V. To assess the impacts of climate change on water sources in Selangor River Basin (SRB)

## 1.5 The scope of the study

1. Historical hydro-meteorological data such as precipitation, temperature, solar radiation, relative humidity, wind speed, and streamflow were collected from various agencies, which are the Department of Irrigation and Drainage Malaysia (DID), and the Malaysian Meteorological Department (MET Malaysia), as well as from global datasets.
2. The performance of historical hydro-meteorological data from global satellite/reanalysis data was analysed using the *point-to-point* method against the observed station and verified using statistical analyses of the correlation coefficient ( $r$ ), mean absolute error (MAE), and root-mean square error (RMSE).

3. Seven scenarios with input data from multiple sources were employed in the SWAT (Soil and Water Assessment Tool) model to define the optimum hydrological model for SRB.
4. The performance of the hydrological model was enhanced by using two calibration methods: the single-site and multi-site.
5. The evaluation of hydrological models using multi-source datasets and multi-site calibration methods was determined using five different objective functions, including Kling-Gupta efficiency (KGE), Nash-Sutcliffe efficiency (NSE), coefficient of determination (R<sup>2</sup>), Percent bias (Pbias), and the RMSE-observations standard deviation ratio (RSR).
6. An ensemble approach of multiple future projections from multiple General Circulation Models (GCMs) based the United Nations Intergovernmental Panel on Climate Change (IPCC) Forth Assessment Report (AR4) provided by NAHRIM (National Water Research Institute of Malaysia) were used to assess the climate change impact in SRB.
7. A calibrated and validated SWAT model was used to simulate the climate change impact on SRB water resources.
8. The impacts of climate change on water resources in the SRB were predicted for 10-year periods starting from 2030, 2040, and 2050 under 15 different future climate realisations.

## **1.6 Limitations of the study**

1. Input data from dam operations was not considered due to the data restriction and availability.
2. The hydrological model has been setup up to the streamflow station at Rantau Panjang due to the uncertain data within the agricultural area.
3. This study applied climate change projection data from AR4 instead of AR5 and AR6 due to the data availability from the managing agency in Malaysia.

4. The effects of climate change impacts on water resources were based on modifications in climatic factors including precipitation and temperature.

## **1.7 Dissertation structure**

The dissertation contains 8 chapters covering the scope of the study. A brief description of each chapter is as follows:

Chapter 1 includes the background of the study, research gap analysis, research question, project aim and objectives, scope of the study, and limitation of the study.

Chapter 2 provides a literature review of the application of global reanalysis datasets, the cross-combined approach, hydrological modeling, multi-site calibration method, climate change and water resources, GCMs (General Circulation Models) and downscaling techniques.

Chapter 3 gives the necessary information about the study area such as, its location hydro-climatic data, and landscape features.

Chapter 4 provides the details of the methods which were used to achieve the various objectives of the study.

Chapter 5 presents the results of the performance evaluation of global satellite/reanalysis datasets for the Selangor River Basin.

Chapter 6 presents the hydrological model performance from the single-site and multi-site calibration method.

Chapter 7 provides the results of climate change assessment on SRB water resources.

Chapter 8 highlights the main results for each objective of the study, concludes the findings, and makes recommendations for further studies.

## **CHAPTER 2**

### **LITERATURE REVIEW**

In this chapter, the most recent studies concerning hydro-meteorological input data, model calibration and validation methods, as well as the modelling of the impact of climate change, are reviewed to achieve the objectives of this study. In terms of hydro-meteorological input data, datasets from individual and cross-combined sources consisting of observation stations and satellite datasets were reviewed. The regionalisation, single-site, and multi-site methodologies were examined in order to enhance the performance of the rainfall-runoff model. Additionally, an evaluation of the effects of climate change on water resources using the Regional Climate Model and ensemble technique was also included. This study proposed Selangor River Basin, Malaysia (SRB) as study area since there are limited studies on water resources in Malaysia, especially SRB, and the importance of SRB as a water supply catchment to the major city (LUAS, 2015).

#### **2.1 Global satellite/reanalysis data**

Reliable hydro-meteorological data is crucial for describing the hydrological condition in a watershed. Traditionally, data from ground observation stations has been employed as an input for watershed modelling simulations. However, due to the sparse distribution of ground observation stations and the malfunction of equipment that causes errors or missing data, during collection, it is unable to accurately depict the spatial variability (Guo et al., 2018). To address these difficulties, a large number of scientists from around the world have produced gridded observational datasets for the globe and regional domains. High geographical and temporal resolution, even over highland regions, and less inhabited areas is a benefit of these datasets. (Nhi et al., 2019). Nowadays, global satellite datasets are frequently employed by researchers as they are accessible to the public (Ghulami, 2018) and significantly complement the lack of gauged data resulting in increased model confidence (Worqlul et al., 2014).

Several well-known and publicly accessible satellite data are the Tropical Rainfall Measuring Mission (TRMM) (Z. Liu et al., 2012), Multi-satellite Precipitation Analysis (TMPA) (Huffman et al., 2007), Climate Prediction Center morphing technique product (CMORPH) (Joyce et al., 2004), Climate Hazards Group Infrared Precipitations (CHIRP) (Funk et al., 2015), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network (PERSIANN) (Nguyen et al., 2018), the Climate Forecast System Reanalysis (CFSR) (NCAR, 2017) and the new satellite data for East Asia region is the China Meteorological Assimilation Driving Dataset (CMADS) (Xian-Yong Meng et al., 2015). The TRMM, TMPA, CMORPH, CHIRP, and PERSIANN are the satellite-based datasets. Meanwhile, the CMADS and CFSR are the reanalysis precipitation datasets created to provide better estimates and they were calibrated against ground-based observations (Dao et al., 2021).

Satellite-based datasets collect precipitation information via visible data, infrared imaging, and passive microwave detection. However, the retrieval algorithms (Lo Conti et al., 2014), sampling technique (Nijssen & Lettenmaier, 2004), and measurement apparatus (Villarini et al., 2009) are always imperfect. As opposed to reanalysis datasets, where the observation of forcing data, the data assimilation technique, and the prediction model all contribute to the accuracy of the datasets (Li et al., 2018). Observed data and model predictions are combined to produce these datasets (Seyyedi et al., 2015). According to Zhang et. al., (2013); the best alternative for presenting weather data is the global reanalysis datasets. In fact, outside of China, particularly in Southeast Asia, these datasets have not been effectively used for hydro-meteorological research (Dao et al., 2021). Therefore, it is an opportunity to explore the potential of these datasets in order to enhance the hydrometeorological model performance, especially in Malaysia. Furthermore, several studies have demonstrated a strong correlation between the observing system and the assimilated data and the correctness of reanalysis datasets (Inoue & Matsumoto, 2004).

### **2.1.1 Climate Forecast System Reanalysis**

The National Centers for Environmental Prediction's third-generation reanalysis product, CFSR, was introduced in August 2004. (NCEP) (NCAR, 2017). It contains three components: (i) assimilation of satellite radiances, (ii) an interactive sea-ice model, and (iii) coupling of the atmospheric and ocean during the generation of the 6-hour guess field (NCAR, 2017). The CFSR is built using contemporary data assimilation methods and a prediction model that extrapolates non-seen characteristics from observed data obtained from diverse sources, such as rain gauges, ships, water balloons, and satellites (Tan, Gassman, et al., 2017). The five elements of the climate in CFSR are temperature, precipitation, wind speed, relative humidity, and solar radiation.

The CFSR is a global, high resolution, coupled system of the atmosphere, ocean, land surface, and sea ice that provides the most precise assessment of the condition of these linked domains from 1979 to the present (36 years) (Saha et al., 2014). The CFSR has a worldwide ocean resolution of 0.25° at the equator and 0.5° globally beyond the tropics with 40 levels. With 64 levels in the vertical, the global atmospheric resolution is around 38 km (T382). The global land surface model features four levels of soil compared to the global sea ice model's three levels (NCAR, 2017). With higher spatial temporal coverage, it provides daily high-resolution rainfall data (Nhi et al., 2019).

According to Tan et. al., (2021), CFSR is one of the most often used products in SWAT modelling. CFSR has been used since 2014 at the upstream area of Three Gorges Reservoir, China (Y. Yang et al., 2014), the Blue Nile River Basin, Ethiopia (Dile & Srinivasan, 2014) and at five selected watersheds, four in US and one in Ethiopia (Fuka et al., 2014), based on the SWAT literature database (CARD, 2021). These studies demonstrated successful streamflow simulation using the SWAT model. The CFSR gave precise rainfall estimates in the Lake Tana basin, Ethiopia, and the Logone watershed, Africa, according to other research by Worqlul et al. (2014) and Nkiaka et al. (2017). Some research has also been conducted in tropical and sub-tropical regions using CFSR as input data, such as the studies from Bressiani et. al., (2015) and Monteiro et. al., (2016).

In Malaysia, there have been limited studies determining the feasibility of using CFSR gridded datasets as an input data in a rainfall-runoff model. Tan et al. (2014) used the CFSR gridded datasets as an input to a streamflow model at the daily and monthly scales for the Johor River basin (JRB). Tan et al. (2017) assessed the competence of three gridded climate products (GCPs), including CFSR at the Johor River basin (JRB) and Kelantan River basin (KRB), and Zhang et al. (2020) assessed the capability of CFSR and CMADS for simulating streamflow in the Muda River basin (MRB).

### ***2.1.2 China Meteorological Assimilation Driving Dataset***

Dr Xianyong Meng from China Agricultural University (CAU) developed the CMADS (China Meteorological Assimilation Driving Dataset), a new public dataset based on CLDAS (China Meteorological Administration Land Data Assimilation System) data assimilation technology. In the CLDAS assimilation system, data from multiple sources, such as satellite observation, land surface observation, and numerical products, were combined (Xian-yong Meng et al., 2016). Data loop layering, resampling, and bilinear interpolation were used in the development of CMADS, which combined the technologies of the Local Analysis and Prediction System (LAPS) and the Space-Time Multiscale Analysis System (STMAS) (Xianyong Meng et al., 2018). Among the data sources for the CMADS series are the over 40,000 regional automated stations that are part of China's 2,421 national automatic and business evaluation centres (Xian-yong Meng et al., 2016). CMADS includes the following elements: daily mean temperature, daily maximum temperature, daily minimum temperature, daily cumulative precipitation (20–20 hours), daily mean relative humidity, daily mean specific humidity, daily mean solar radiation, daily mean wind, daily mean atmospheric pressure, and soil temperature and soil moisture (Xianyong Meng & Wang, 2017). Hourly data on temperature, pressure, and wind speed obtained from 29,452 local weather stations and 2,421 national weather stations were combined using the ambient field of the European Centre for Medium-Range Weather Forecasts (ECMWF) and the Space and Time Multiscale Analysis System (STMAS). Combining observed data from meteorological stations with precipitation

reanalysis data from the NOAA was done using the CPC MORPHing technique (CMORPH). Using the discrete-ordinate radiative transfer (DISTORT) model, solar radiation data from the International Satellite Cloud Climatology Project (ISCCP) radiance data were combined with data from the FY-2E (FengYun-2) satellite (Tian et al., 2018).

The CMADS dataset covers the whole of East Asia and started being used in 2016 by Meng et. al., (2016) at the Heihe River Basin, China, to improve hydrologic modelling. The CMADS dataset continued being used by Meng et. al., (2017) in the Manas River basin to evaluate the water cycle situation in a basin covered with glaciers and snow. In total, 34 papers were found in the Web of Science database and 32 papers in the SWAT literature database, with most of the studies being conducted in China's river basin. Only three of those studies were conducted outside of China, with Vu et. al., (2018) applying the CMADS dataset to the Han River Basin in Korea, Dao et. al., (2021) to the Cau River Basin in Vietnam, and Zhang et. al., (2020) to the Muda River Basin in Malaysia. In general, most of the studies conducted using the CMADS dataset show satisfactory performance in simulating streamflow, except for Vu et. al., (2018) where the result is moderate performance.

## **2.2 Data performance**

As homogeneous datasets that can be employed immediately to simulate streamflow, CFSR and CMADS are more reliable than other satellite derived precipitation products. These two reanalysis datasets have this significant benefit over satellite precipitation products, which lack linked temperature information and have a range of timescales (Dao et al., 2021). For a group of surface variables, such as precipitation, surface air temperature, and surface heat fluxes, Wang et al. (2011) investigated the climatic variability of the CFSR dataset in 2011. The results showed that CFSR successfully replicated the observed mean precipitation spatial pattern with better precipitation distribution across the tropical north-western Pacific, the South Pacific Convergence Zone, South America, and south-eastern Pacific compared to R1 (NCEP/NCAR reanalysis), R2 (NCEP/DOE reanalysis), and ERA40 (European Centre for Medium Range Weather Forecast) dataset. Water flux

is also more accurately portrayed because of the enhanced precipitation distribution in the CFSR dataset. Research in the Lake Tana basin, Ethiopia, by Worqlul et. al., (2014) and Dile and Srinivasan (2014) showed the CFSR produced the most accurate rainfall estimates and performed satisfactorily in hydrological modelling. While Fuka et. al., (2014) employed the CFSR at five watersheds that represent various hydroclimate regimes to obtain historical weather and demonstrate the application of the dataset. The results show that the CFSR is as accurate as a weather station at predicting streamflow across a wide range of hydroclimate regimes and watershed models. Additionally, Tan (2014) and Tan et. al., (2017) have used CFSR datasets for the Kelantan and Johor River basins in Malaysia. According to findings from both studies, CFSR shows an acceptable performance in monthly streamflow simulation and shows a strong correlation with daily maximum temperature data. CFSR also performed better in hydrological simulation when solar radiation was included as an input source (Gu et al., 2020). In fact, the application of CFSR datasets in streamflow simulation studies is still limited in Malaysia (Mou Leong Tan, Gassman, et al., 2017).

When compared to TWR (traditional weather station) on monthly and daily time scales, Meng et. al., (2016) found that CMADS is more accurate than CFSR. In terms of a watershed with glaciers and snow cover, such as the Manas River basin (MRB) in China, CMADS performed well on a monthly basis with the runoff simulation following a pattern that is comparable to the observed and the timings of flood occurrence matching well (Xianyong Meng et al., 2017). CMADS can correctly duplicate the gauged precipitation and outperforms the Tropical Rainfall Measuring Mission (TRMM) and the Integrated Multi-satellite Retrievals for GPM (IMERG) in detecting precipitation (Song et al., 2021). Additionally, because of the high spatial resolution of CMADS, integrating SWAT with CMADS in a hydrological simulation shows a higher degree of accuracy when compared to TRMM and IMERG.

A comparison study on two reanalysis datasets (CFSR and CMADS) and two satellite-based datasets (TRMM and PERSIANN-CDR) by Gao et. al., (2018) shows the correlation between reanalysis datasets and gauge observation is better than that of satellite-based datasets. Whereas Vu et. al., (2018) reported that when compared to gauged rainfall data and runoff simulation on the Han River Basin in South Korea, satellite precipitation products from TRMM and CMADS exhibit a greater degree of accuracy than PERSIANN and PERSIANN-CDR (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network-Climate Data Record). Guo et. al., (2018) used two hydrological models (IHACRES and Sacramento) to compare three precipitation datasets (CMADS, TMPA-3B42V7 and gauged interpolated) in simulating daily streamflow in the Lijiang River, Southern China. The findings demonstrate that, compared to the other precipitation products, CMADS was more effective at capturing the peaks in streamflow. This is further supported by the findings of Cao et.al., (2018), who discovered that CMADS was successful in predicting daily streamflow in the Lijiang River Basin, China. A comparison of two popular datasets from Liu et. al., (2018) reveals that CMADS datasets have higher accuracy and quality than CFSR. Similar to CFSR, the application of CMADS datasets in Malaysia is still limited, as one based on the CARD database has been found.

### **2.3 Cross-combined dataset**

Meteorological data are necessary for hydrological models, and the accuracy of the model simulation is strongly correlated with the geographical and temporal resolution of the data (Eini et al., 2019). In the hydrological model, fewer input variables cause uncertainty in the simulation, resulting in lower model accuracy. The meteorological data that powers the hydrological model should ideally have acceptable spatial and temporal precision and be able to accurately represent the local weather. Unfortunately, there are few meteorological stations and the point-based monitoring ranges are not very large (Duan et al., 2019).

The cross-combined dataset approach is a method to combine input variables from multiple data sources in order to enhance the hydrological model's performance. Remote sensing technology advancements now provide abundant and valuable observation information as well as multi-temporal scale meteorological data for hydrological research (Awange et al., 2019). This opens up new research possibilities for hydrological models using sparse meteorological stations (Famiglietti et al., 2015; Sun et al., 2018). Numerous researchers have also conducted relevant studies using this technique (Jiang et al., 2012; Worqlul et al., 2017; Xue et al., 2013), however the bulk of them focus on the influence of precipitation meteorological components on hydrological models (Beck et al., 2017; Fang et al., 2019; Li et al., 2018). When building a hydrological model using this method, focusing just on one variable may cause uncertainty. Therefore, a cross-combined approach can overcome these limitations by combining input variables from observation stations and satellite datasets.

The precise spatial distribution within the watershed can be shown with the help of consistent and adequate hydrological data. The spatial distribution may be difficult to precisely reflect in some locations due to a lack of observation stations/reliable data (Miao et al., 2015; Shao et al., 2018). Most hydrometeorological studies have been conducted using the single-source dataset. As far as this research is concerned, only two studies applied cross-combined datasets from the global reanalysis dataset to enhance model simulation based on the SWAT database and the Clarivate Analytics Web of Science database (Web of Science, 2022). In SWAT model, temperature is a key variable as it affects many hydrological processes such as evapotranspiration, snowmelt, and soil moisture. Meanwhile solar radiation affects hydrological processes, particularly evapotranspiration, relative humidity affects rate of evapotranspiration by measure of the amount of moisture in the air and wind speed affects the rates of evapotranspiration and the dispersion of pollutants in the air. SWAT uses temperature data to simulate the rates of evapotranspiration from vegetation and water bodies, as well as the rate of snowmelt in winter months, uses solar radiation data to simulate the amount of energy available to drive the processes of evaporation

and transpiration from vegetation and water bodies, uses relative humidity data to calculate the vapor pressure deficit, and uses wind speed data to calculate the aerodynamic resistance of vegetation and the rate of dispersion of pollutants in the air. Instead of utilising only one data source, Shao et al. (2018) combined the precipitation values from CMADS and the observed gauge to be utilised in the model to drive the hydrological model for enhanced simulation results. The result demonstrates that the combination datasets of observed gauge and CMADS provide higher Nash-Sutcliffe Efficiency (NSE) values than single-source data, with an average of 0.72 in the calibration period and 0.42 in the validation period, as illustrated in Figure 2-1. As a result, a combined dataset can provide the best simulation result and enhanced performance compared to a single source dataset.

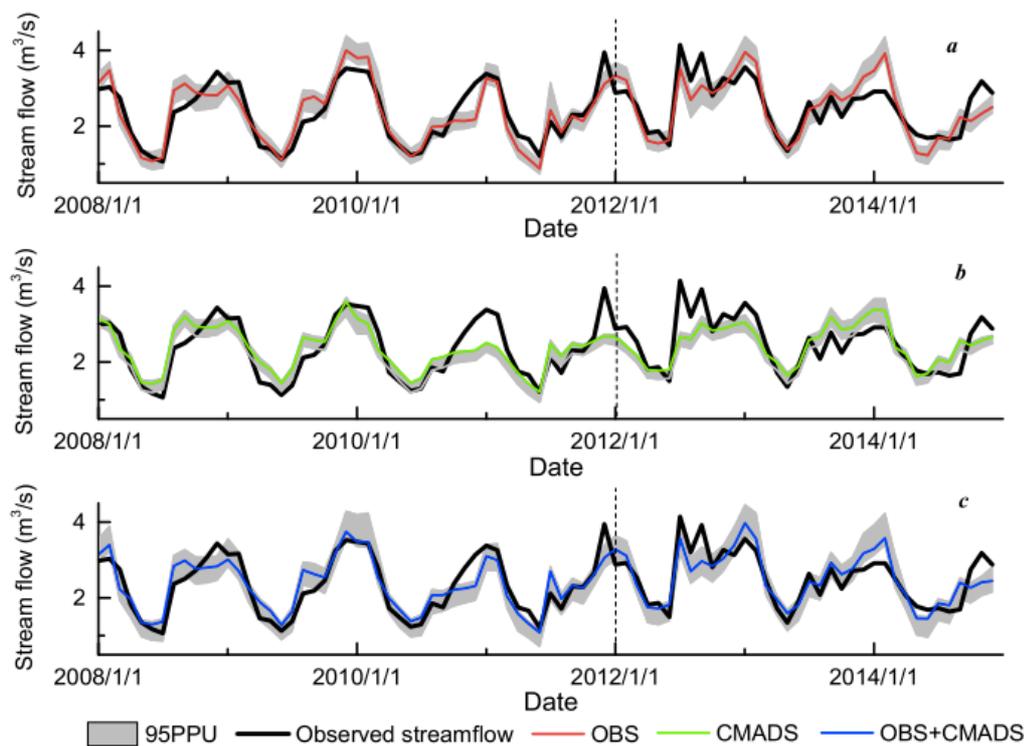


Figure 2-1: simulation results (a) observation station; (b) CMADS; (c) Observation+CMADS in the Hailiutu River basin during the period 2008–2014 (source: Shao et. al., (2018)).

In the Yellow River Source Region, China, Gu et al. (2020) used data from three sources, including the observed gauge, CFSR, and CMADS, to determine the amount of precipitation, temperature, and solar radiation. According to Gu et al., (2018), (i) gauged precipitation consistently performed better in runoff modelling than CMADS and CFSR (ii) in comparison to gauge temperature data, CMADS and CFSR performed better, (iii) the solar radiation of CFSR is more suited for hydrological modelling than CMADS, and (iv) model powered by various sources datasets performed exceptionally well, particularly on observed gauge and CFSR. Figure 2-2 demonstrates how the simulation results of the model are strongly influenced by the precipitation datasets, but the streamflow simulation is only slightly impacted by the temperature and solar radiation datasets. According to Gu et al., (2018), result from observed precipitation performed better, followed by CFSR and CMADS with precipitation input of GD, CMADS and CFSR were 519.24 mm, 412.51 mm and 768.07 mm respectively as illustrated in Figure 2-3. Meanwhile, as shown in Figure 2-4, the input of CFSR temperature perform best, followed by CMADS and GD.

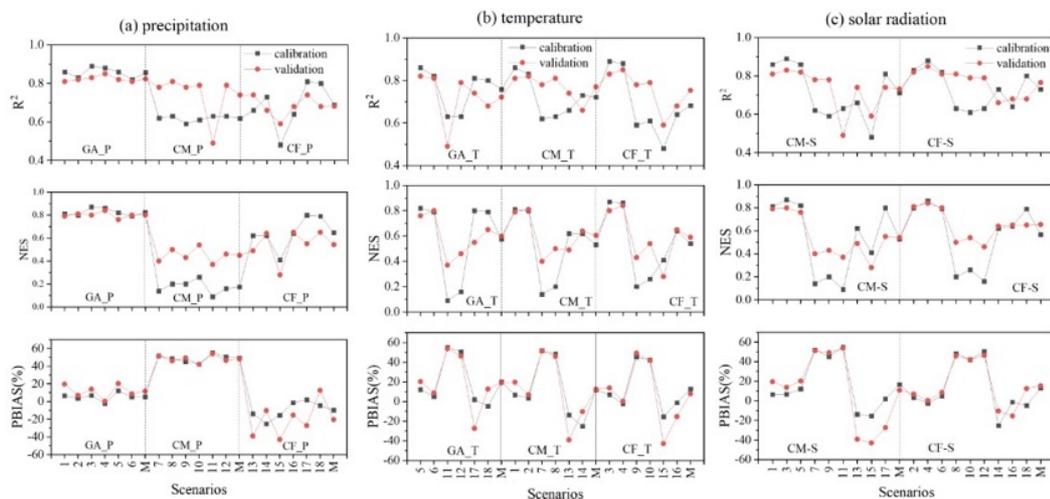


Figure 2-2: Statistical figures of the simulation results at the daily scale for (a) precipitation; (b) temperature; (c) a solar radiation data input scenario of the three datasets (GD, CMADS and CFSR) (source: Gu et al., (2020))

As a result, the combination of multiple datasets influences the model simulation, particularly the elements related to precipitation. Even a temperature variable can yield a good streamflow simulation (Mou Leong Tan, Gassman, et al., 2017). In this work, we present a cross-combination of five input variables, including precipitation, temperature, solar radiation, relative humidity, and wind speed, from multiple datasets between observation station and satellite datasets, as opposed to concentrating on three input variables, such as temperature, precipitation, and solar radiation (CFSR and CMADS). The potential of additional input variables (relative humidity and wind speed) was studied by simulating the basin streamflow.

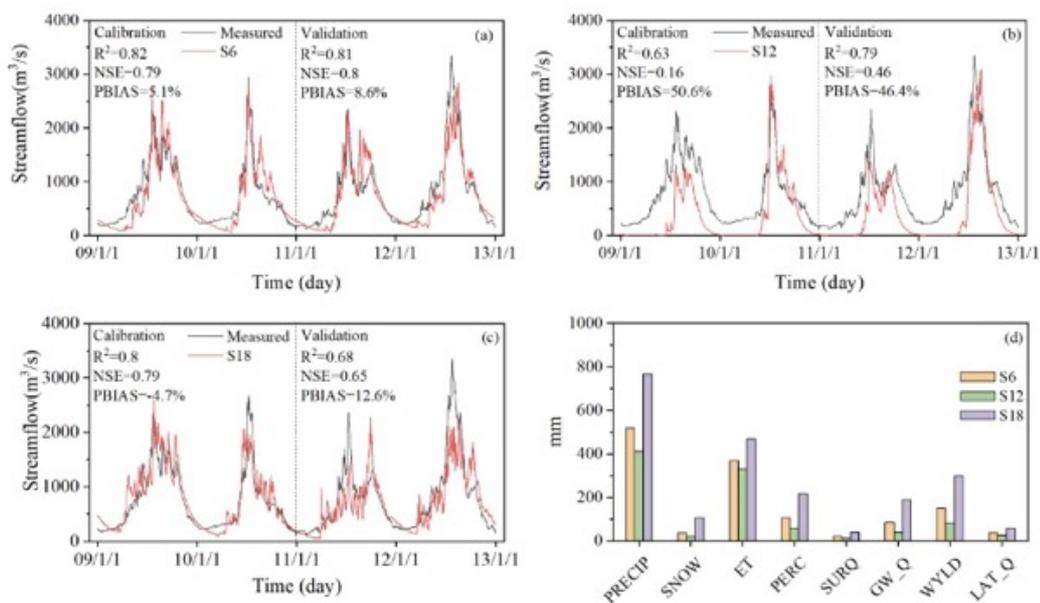


Figure 2-3: Streamflow simulation model result using three different precipitations (a) GD; (b) CMADS; (c) CFSR, and (d) output analysis of water balance for models (source: Gu et al., (2020))

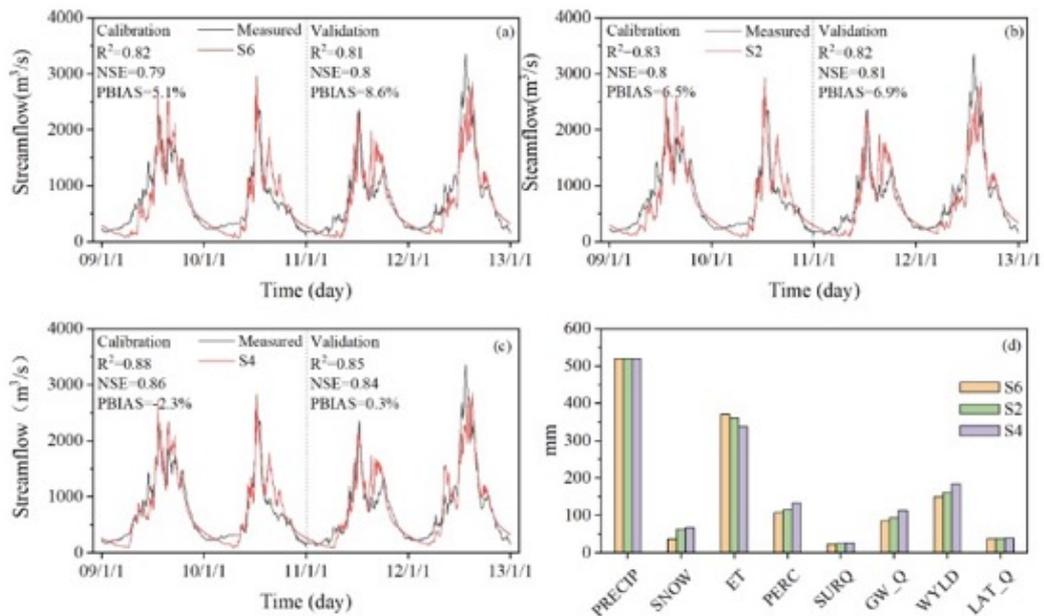


Figure 2-4: Streamflow simulation model result using three different temperature (a) GD; (b) CMADS; (c) CFSR, and (d) output analysis of water balance for models (source: Gu et al., (2020))

## 2.4 Regionalisation

Regionalisation is a procedure that quantifies numerous hydrological processes that occur in ungauged watersheds by transferring an optimal set of parameters from one or more gauged watersheds to an ungauged watershed (Merz & Blöschl, 2004; Swain & Patra, 2017). This method comes after many of the watersheds have limited or no monitoring data available, or that the monitoring data is sporadic or known to be unreliable (Sivapalan et al., 2003). According to Gitau and Chaubey (2010), this issue might be solved by extending and/or generalising model parameters discovered during the calibration of gauged watersheds to ungauged watersheds *within the same region*. Additionally, Swain and Patra (2017), stated that regionalisation is an alternative method for predicting characteristics including streamflow, sediment load, and water quality in an ungauged watershed.

There are two methods of regionalisation: the hydrological model-dependent method and the hydrological model-independent method (Razavi & Coulibaly, 2013). The most widely used method is regression analysis, which employs model parameters, and physical watershed properties as dependent and independent variables, respectively (Swain & Patra, 2017). For example, to investigate the possibilities of creating regionalized model parameter sets for application in ungauged watersheds, Gitau and Chaubey (2010) compare two regionalisation techniques: global averaging and regression-based parameters. The greatest advancements in regionalisation methodologies have only happened in the last decade, and there are few examples of regionalisation research from the final decade of the 20<sup>th</sup> century (Blöschl, 2016; Hrachowitz et al., 2013).

## **2.5 Hydrological models**

A hydrological model is created utilising a constrained set of measuring methods, a constrained set of spatial and temporal observations, and condensed presumptions about how natural hydrological systems work. (Beven, 2012; Okiria et al., 2022). Numerous hydrological models are available for use in various contexts, including water resource planning and management, flood forecasting, water supply, and water quality assessment (Ghulami, 2018). Black-box, conceptual, and distributed models are the three different categories of hydrological models. The underlying hydrologic process is frequently expressed as an empirical model in "black-box" models, which do not provide a description of it. Conceptual models, as opposed to black-box models, are based on more complex empirical equations, and are intended to depict the underlying physical process. Examples of such models include the Tank model (Sugawara, 1974) and Sacramento model (Burnash, 1979). The equations employed to define the distributed model description are based on the principles of conservation of mass and energy and have a stronger physics underpinning in the underlying hydrological process.

The SWAT (Soil and Water Assessment Tool) model (Arnold, Moriasi, et al., 2012) and MIKE SHE (MIKE System Hydrological European Model) models (Abbott et al., 1986) are examples of distributed models. The SWAT model was employed in this study because it is the most extensively used hydrological model globally (Gassman, Reyes, Green, Arnold, et al., 2007) and needs further reviews for Southeast Asia, where the model has been progressively deployed over the past 10 years (Mou Leong Tan et al., 2019).

## **2.6 Soil and Water Assessment Tool**

Dr. Jeff Arnold created the Soil and Water Assessment Tool (SWAT), also known as a river basin or watershed scale model, for the United States Department of Agriculture's Agricultural Research Service (USDA-ARS) in order to forecast how land management techniques will affect water, sediment, and agricultural chemical yields in significant complex watersheds with changing soils, land use, and management conditions over time. (Neitsch et al., 2011). The SWAT was developed based on more than 30 years of modelling experience at USDA-ARS (Gassman, Reyes, Green, Arnold, et al., 2007). The first SWAT version, 94.2, was issued in the early 1990s. Then, it was modified to versions 96.1, 98.2, 99.2, 2000 (Arnold, Moriasi, et al., 2012), and the most recent version is SWAT2012 revision 684 (SWAT, 2022). The USDA-ARS models such as GLEAMS (Groundwater Loading Effects of Agricultural Management Systems), CREAMS (Chemicals, Runoff, and Erosion from Agricultural Management Systems), and EPIC (Environmental Policy Integrated Climate Model), as well as other models, as indicated in Figure 2-5, contributed important components to the current SWAT models.

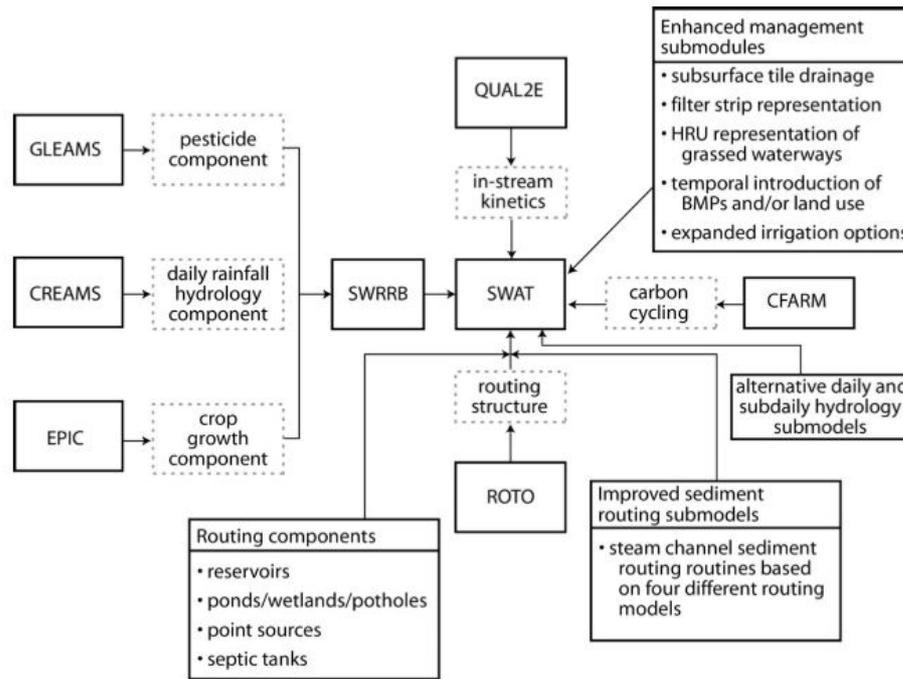


Figure 2-5: SWAT development history and model application (source: Gassman et. al., (2007))

A semi-distributed model called SWAT is capable of simulating time steps every day, every month, or every year (Arnold, Kiniry, et al., 2012; Gassman, Reyes, Green, Arnold, et al., 2007). Researchers investigated hydrological processes, the consequences of climate change, changes in land use, water use management, and evaluations of water quality and quantity using SWAT extensively (Gassman, Reyes, Green, & Arnold, 2007). As of 2022, over 4,300 SWAT-related publications have been published based on the Clarivate Analytics Web of Science database and the Center for Agriculture and Rural Development (CARD), with 126 articles concentrating on Southeast Asia. Vietnam and Thailand were the top two countries in Southeast Asia for SWAT applications, with 28.6% and 22.2%, respectively (Mou Leong Tan et al., 2021). According to Mannschatz et. al., (2016), research on the web-based comparison of modelling tools showed that the SWAT model earned 39.2% of the popularity indexes ( $P_r$ ) of 352 water, soil, and waste management models, which is higher than other models.

The SWAT model is preferred among researchers because of its approachability, versatility, and high level of evaluation, promotion, and support (Mou Leong Tan et al., 2021). In Malaysia, 32 SWAT application studies have been demonstrated since 2012, based on the Clarivate Analytics Web of Science database (Web of Science, 2022), with Abu Hasan (2012) being the first to apply SWAT modelling with a focus on the flow and sediment in the Bukit Merah reservoir. While the most recent SWAT application was presented by Raffar et. al., (2022) with a study on the watershed-scale modelling of the irrigated rice farming at Muda River basin. Within these 32 studies, only one study has been carried out in the Selangor River Basin by Kondo et. al., (2021). However, the study by Kondo et. al., (2021) focused on the fecal contamination for watershed management under tropical climate conditions. Therefore, it is a chance to investigate the SWAT model's capability in order to simulate streamflow in SRB for future water resources management and to contribute to the literature about the Asean environment, particularly in Malaysia.

## **2.7 Model calibration and validation**

Calibration is a technique for more precisely parameterizing a model for a specific set of local variables and lowering the degree of prediction uncertainty. According to Zhang et. al., (2016), it is challenging to select an appropriate set of parameter values that may precisely reflect the hydrology of the watershed. By contrasting the model output with the observed data under the same conditions, the values for the input parameters are carefully selected within their respective ranges (Arnold, Moriasi, et al., 2012). The calibration of a hydrological model is normally carried out at a single site or outlet within a watershed, which is suitable for a small watershed but not for a big watershed (S. Wang et al., 2012). Therefore, it is suggested to use multi-site calibration to get the best model performance since it captures the geographical variability in the (bigger) watershed (Desai et al., 2021) and has a higher level of parameter flexibility than single-site calibration (Shrestha et al., 2016). Calibration used to be done manually, which is more subjective and time-consuming (Mousavi et al., 2012). However, the auto-calibration method

has gained popularity among researchers due to its availability and ability to shorten the entire calibration process (Molina-Navarro et al., 2017).

### **2.7.1 Multi-site calibration**

The multi-site calibration method (MSC) is a technique for calibrating a large watershed by subdividing it into smaller watersheds and applying different model parameters to each and every specified watershed (Bai et al., 2017). MSC has emerged as a significant development path in contemporary hydrological simulation research with the aim of improving model simulation accuracy (Desai et al., 2021; Shrestha et al., 2016; Song et al., 2021), after it was found that single-site calibration (SSC) was ineffective, particularly when it came to the numerous parameters and their expression on the spatial variability of a large watershed (Anderton et al., 2002; Bai et al., 2017). Numerous studies have used the MSC method to enhance the hydrological model. For instance, Li et al. (2010) found that the validation watershed's relative cumulative error of predictions was reduced by 3.5-7.4% for monthly streamflow, 3.5-6.3% for monthly total nitrogen loads, and 4.3-5.9% for monthly total phosphorus loads by using a multi-site weighted average approach in conjunction with multiple optimisation goals. According to Saeidifarzad et. al., (2014), MSC can produce reliable parameters with significant physical meaning as opposed to SSC, which fails to successfully link the model parameters to the physical features and lacks a distinct physical meaning. Wang et. al., (2012) used MSC to calibrate and validate the Chaohe basin, a significant mountainous watershed in northern China. The findings showed that MSC improved for the third station (Dage station) while somewhat declined for two of the three stations (Xiahui and Daiying). Desai et al. (2021) used MSC to assess the water balance in the semi-arid region of Central India, the Betwa river basin. According to the findings, the Coefficient of Determination ( $R^2$ ), Nash-Sutcliffe efficiency (NSE), RMSE-observations standard deviation ratio (RSR), and percent bias (PBIAS) performance ratings, respectively, varied from 0.83 to 0.92, 0.06 to 0.91, 0.03 to 0.63, and -19.8 to 19.3, respectively.

In order to enhance model calibration in the Lake Chad basin, Nkiaka et al., (2018) utilised a multiple calibration technique to understand the hydrological process in spatially heterogeneous catchments. Nkiaka et al. (2018) found that the sequential calibration approach outperformed the SSC and simultaneous procedures at both daily and monthly time-steps. The MSC approach was also used in this study due to its substantial research to improve the simulation of the hydrological model and to lower the uncertainty of the watershed.

### **2.7.2 SWAT Calibration and Uncertainty Program Software**

Eawag-Swiss Federal Institute of Aquatic Science and Technology created the SWAT calibration and uncertainty programme software, also known as SWAT-CUP, which has automatic calibration and uncertainty analysis capabilities. (Arnold, Moriasi, et al., 2012; Ozdemir & Leloglu, 2019). SWAT-CUP is a standalone programme created for SWAT calibration that has five alternative calibration techniques, functions for validation and sensitivity analysis, and the ability to visualise the research region using Bing Map (K. C. Abbaspour et al., 2015). Generalised Likelihood Uncertainty Estimation (GLUE), Parameter Solution (ParaSol), Sequential Uncertainty Fitting (SUFI-2), Monte Carlo Markov Chain (MCMC) and Particle Swarm Optimization (PSO) are five techniques that are integrated into SWAT-CUP (K. C. Abbaspour et al., 2015). By representing uncertainties from several sources, such as the conceptual model uncertainty, input uncertainty, and parameter uncertainty, SUFI-2 was able to calibrate and validate, according to Rusli et al., (2017). It was also revealed that SUFI-2 was quite successful for laborious calibration on large-scale models (Yang et al., 2008). Thus, this study has utilised the SUFI-2 programme for calibration and an uncertainty analysis model.

## **2.8 Climate change**

According to NAHRIM (2006), rainfall in Peninsular Malaysia is projected to increase about 2.6% by 2041-2050, while river flood flow and low flow would become more extreme. For instance, it was anticipated that the Muda River's and Selangor River's monthly low flows would decline by 79% and 93%, respectively. While MMD (2009) simulated future temperatures and rainfall up to the year 2099 for the entire country. Compared with the base period 1990-1999, the simulated mean annual rainfall for 2020-2029 and 2050-2059 is significantly lower. However, for 2090-2099, the simulated rainfall is significantly higher than 1990-1999 especially for Sarawak and Peninsular Malaysia.

### **2.8.1 Impacts of climate change**

The IPCC's (Intergovernmental Panel on Climate Change) Sixth Assessment Report (AR6) is based on the IPCC's Fifth Assessment Report (AR5) as well as the Special Reports of the Sixth Assessment (AR6) cycle on *Global Warming of 1.5 °C (SR 1.5 °C)* (IPCC, 2019b), *Climate Change and Land (SRCCL)* (IPCC, 2019a), and *the Ocean and Cryosphere in a Changing Climate (SROCC)* (IPCC, 2019c). Since the 2014 release of the IPCC AR5 report, the report assesses new research, methodology developments, and adjustments to climate change mitigation measures. Recent developments and current trends reported in AR6 indicated that the period from 2010 to 2019 experienced the greatest average annual GHG (greenhouse gas) emissions of any previous decade, with  $59 \pm 6.6$  GtCO<sub>2</sub>-eq in 2019, which is 12% more than in 2010 and 54% more than in 1990 (IPCC, 2022b). However, between 2010 and 2019, the average annual growth rate fell from 2.1% between 2000 and 2009 to 1.3% annually. According to AR5, in the coming century, global temperatures will rise by 1.1 to 6.4 °C, and sea levels will rise by 18 to 59 cm. Global warming is expected to reach 1.5 degrees Celsius between 2030 and 2052 if the current pace of growth continues (Masson-Delmotte et al., 2018).

From an Asian perspective, Shaw et al. (2022) reported an increase in surface temperature since the 20th century, as well as a steady increase in the number of hot days and warm nights across the continent, while a steady decrease in the number of cold days and nights, with the exception of the southern part of Siberia. In Asia, the threat of heatwaves, droughts in arid and semiarid regions of West, Central, and South Asia, delays and weakening of the monsoon circulation, floods in monsoon regions of South, Southeast, and East Asia, and glacier melting in the Hindu Kush Himalaya region increased with rising temperatures (Doblas-Reyes et al., 2021). The ability to deal with such consequences has not made much progress in Asia and has had disastrous consequences for the population and economies of the region (Shaw et al., 2022).

Climate change is occurring today and has done so throughout the duration of natural history, regardless of whether it is caused by natural cycles, human causes, or any combination of these elements. The IPCC (2017) predicts that temperature rises, and precipitation variability will have a substantial impact on water-related hazards including floods and droughts. Climate change also has an impact on a number of other elements, including evapotranspiration, infiltration, and surface runoff (Di Baldassarre et al., 2011). The water cycle and climate change interact closely. Precipitation and evapotranspiration patterns are projected to be most significantly impacted as a result of climate change's accompanying alterations to the hydrologic cycle (Tsanis et al., 2011). In addition to the growing human population and the advancing global economy, Oki and Kanee (2006) noted that variations in water availability brought on by climate change are another concern that makes it difficult to utilise water sustainably. For instance, Ligaray et al. (2015) investigated how the hydrological responses to different climate sensitivity and greenhouse gas emission scenarios changed in the Chao Phraya River, Thailand. According to the findings, a considerable shift in streamflow is brought on by an increase in CO<sub>2</sub> build-up. As a result, Gonzalez-Zeas et al. (2019) recommended focusing more on hydro systems in tropical and mountainous regions to ensure sustainable management of water resources in the face of climate change.

In Malaysia, several research on climate change have been carried out. By utilising 10 GCMs under three RCPs, Dlamini et al. (2017) evaluate the impacts of climate change on the hydrological flows of the upper Sg. Bernam basin in Selangor (4.5, 6.0 and 8.5). This study finds that future temperatures and rainfall will increase in all scenarios, with average projected rainfall changes during dry seasons of -2.4%, -3.2%, and -3.7% and wet seasons of 1.0%, 0.8% and 2.4%. During the dry season, RCP 8.5 was the most severe scenario. Shaaban et al. (2011) evaluated the effects of future climate change on water resources using the regional hydro climatic model of Peninsular Malaysia (RegHCM-PM). Tan et al. (2015) investigated how changes in climate and land use impacted hydrological components in the Johor River Basin.

The effects of climate change have been demonstrated to have an influence on Malaysia as well. According to Wan Azli (2010), recent climatic and hydrological events in Malaysia may have demonstrated a deviation from the regular trends and an unstable rainfall pattern in Peninsular Malaysia. In the same journal issue, Tangang et al. (2012) discovered that Malaysia's climatic variability is controlled by the Madden-Julian Oscillation (MJO), the El Niño-Southern Oscillation (ENSO), and the Indian Ocean Dipole (IOD). Both came to the conclusion that Malaysia's temperatures are rising and that precipitation is trending in a more erratic way over the long run. There is therefore an immediate need for action to conserve water supplies and to research and develop alternative sources of water in the nation due to the unpredictable nature of rainfall and climate change (NWRS, 2011). It is critical to discover strategies to either update current water availability models or create new ones since it is quite likely that climate changes will have an effect on water resources and water availability in the near future. Furthermore, as reported in NWRS (2011), SRB currently faces a deficit in water supply, and it will get worse with the climate change impact. For this reason, research on SRB water resources and their effects on climate change is crucial to ensure the long-term viability of these resources.

## 2.9 General Circulation Model

Climate models, also known as general circulation models (GCMs), are numerical simulations that forecast how the global climate system will react to rising greenhouse gas concentrations. The cryosphere, atmosphere, ocean, and land surface physical processes may all be represented by this most advanced technology (Mahmood, 2013). In order to give physically and geographically consistent estimates of regional climate for impact assessments, only GCMs working in tandem with nested regional models are capable of doing so. Figure 2-6 illustrates the three-dimensional grid used by GCMs to simulate the climate, which had horizontal resolution between 250 and 600 km, atmospheric vertical levels between 10 and 20, and 30 layers in the ocean (IPCC, 2022a).

According to Dlamini et. al., (2017) and Prudhomme et. al., (2003), the most appropriate information for evaluating both historical and future climate scenarios is currently produced from GCMs. For a better translation of climate data, a downscaling strategy is usually required because GCMs are typically provided at coarse resolution. Dynamic downscaling and statistical downscaling are the two main downscaling techniques (Nkululeko Simeon Dlamini et al., 2017). The dynamic downscaling is referred to as a mini-GCM because it uses the same boundary constraint as the driving GCM to shrink the horizontal region covered (typically around 25 by 25 km or even less) in an effort to replicate local climate variables. Despite producing high resolution climate data, these methods have not been widely used because of the expense and difficulty of maintaining them. Statistical downscaling methods, which include methods like weather type algorithms, transfer functions and stochastic weather generators are currently the most often used instruments for investigating climate change (R L Wilby et al., 2004). They are adopted because they provide point climate data at a particular site of interest and are easy to implement (Robert L. Wilby et al., 1998).

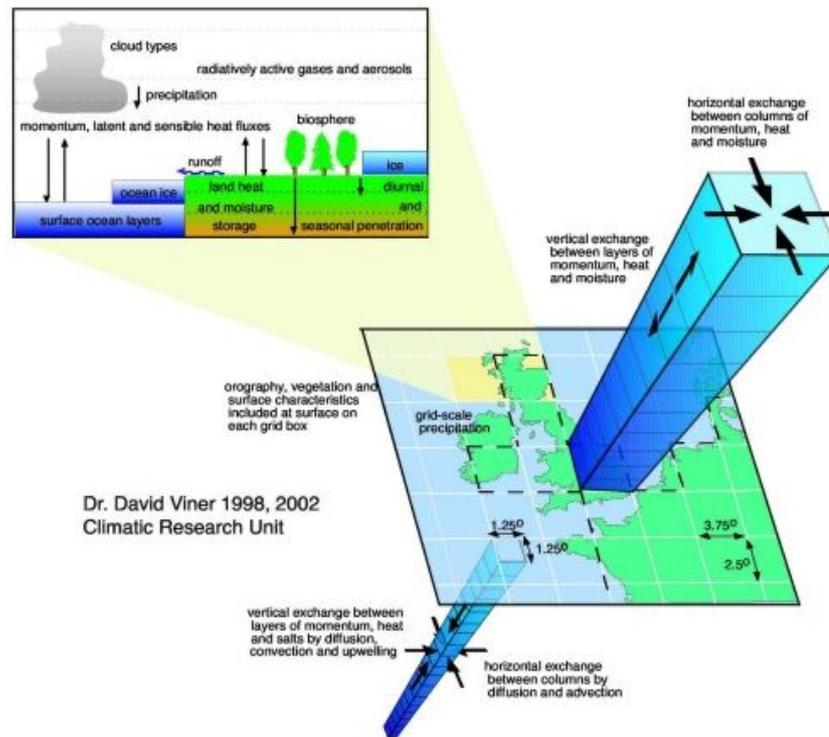


Figure 2-6: Three-dimensional grid of General Circulation Model (GCM)  
 (source: <https://www.ipcc-data.org/>)

## 2.10 Regional Climate model

Regional climate modelling was first developed in the late 1980s (Dickinson et al., 1989; Giorgi & Bates, 1989) and has rapidly expanded over the past three decades with more than 400 papers per year in late 2010s based on the Web of Science database (Giorgi, 2019). In order to reduce the coarse resolution of global climate models (GCMs), regional climate models (RCMs) were created, as shown in Figure 2-7. Grid scales used in RCMs are more suited for research on local phenomena and for use in vulnerability, impacts, and adaptation (VIA) analysis (Giorgi, 2019).

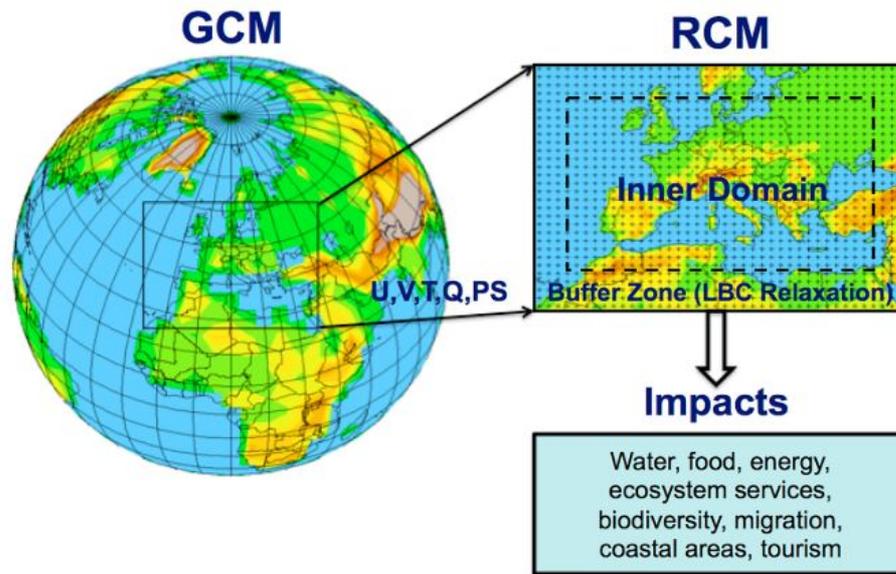


Figure 2-7: Schematic depiction of Regional Climate Modelling and the applications (source: Giorgi, (2019))

### **2.10.1 Regional Hydro Climate Model of Peninsular Malaysia**

The Regional Hydro Climate Model of Peninsular Malaysia, or RegHCM-PM, was developed in order to downscale the global historical and climate change atmospheric database produced by the First Generation Coupled General Circulation Model (CGCM1) to Peninsular Malaysia at a fine spatial resolution of approximately 9 km from a coarse grid resolution (~410 km). At the regional and watershed sizes, it is impossible to assess climate change using a coarse spatial resolution. To explore the potential long-term climatic and hydrologic effects of global warming on the hydrologic regime and water resources, the coarse scale of GCM data must be downscaled to a regional size and watershed (Chen et al., 2006). RegHCM-PM has a mesoscale atmospheric model component and a regional land hydrology model component, just like the original IRSHAM (Integrated Regional-Scale Hydrologic/Atmospheric Model) for Japan. However, RegHCM-PM is a nonhydrostatic (MM5) atmospheric model, in contrast to IRSHAM, which is hydrostatic.

Zaki et al. (2019) used RegHCM-PM to downscale a coarse spatial resolution of 410 km CCCma to a lower gridded resolution of 6 km in order to analyse the effects of climate change on Peninsular Malaysia's hydrology and water resources. The RegHCM-PM downscaling process utilised three layered domains, as shown in Figure 2-8. As its initial and boundary conditions, the first run of the RegHCM-PM used an outer domain of the global scale atmospheric data spanning the entire Peninsular Malaysia region as well as the surrounding areas. The outer domain, with a 54 km grid, covers Southeast Asia. RegHCM-PM second domain (intermediate domain) was run over the inner domain with conditions collected from world scale atmospheric data and interpolated at a 6 km resolution. The second domain RegHCM-PM simulation results were used to determine the boundary conditions. With a spatial grid resolution of 6 km and time steps of an hour, RegHCM-PM simulations using these initials, boundary conditions, and the inner domain generate a complete set of atmospheric data (precipitation, air temperature, radiation, wind, relative humidity, and evapotranspiration).

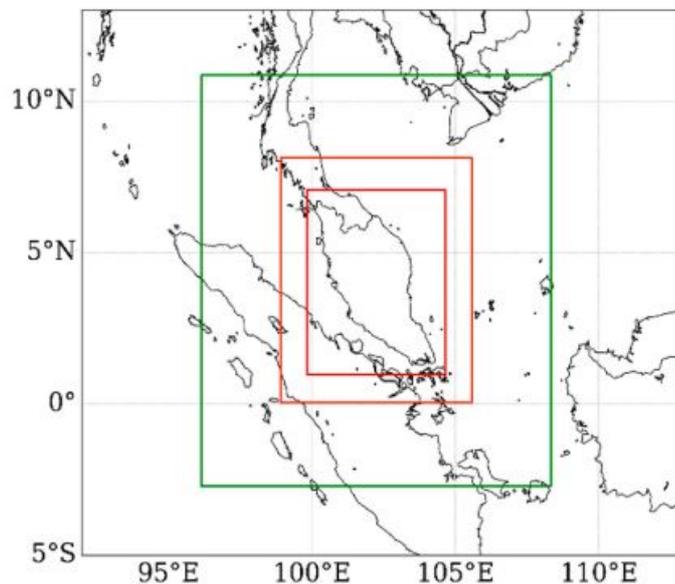


Figure 2-8: The modelling domains of Peninsular Malaysia hydro climate model: The green box covered by 54 km x 54 km grids is the larger outer domain; the orange box covered by 18 km x 18 km grids is the intermediate domain; and the red box is covered by 6 km x 6 km grids was the inner modelling domain (source: Amin et. al., (2019))

There are several studies on historical hydro-climatic observations in the literature, but very few studies on how climate change may affect hydro-climatic systems in the future across Peninsular Malaysia (M.Z.M. Amin et al., 2017). For instance, Suhaila and Jemain (2009) examined historical hydro-climatic measurements to see if models that separate rainfall data into categories depending on the frequency of nearby wet days between 1975 and 2004 gave better fits than models that included all the data. With regard to how anthropogenic warming affects the monsoon and other phenomena like the Madden-Julian Oscillation (MJO), the El Niño-Southern Oscillation (ENSO), and the Indian Ocean Dipole (IOD), Tangang et al. (2012) provide an overview of current knowledge in the physical basis and science of climate change. While Desa et al. (2001) investigated the pattern of excessive rainfall and predicted the likely maximum precipitation (PM) in the Selangor area. Whereas, on hydro-climatic processes, Shaaban et. al., (2011) predicted the changes of water resources in Peninsular Malaysia based on one realisation of future climate change projections, which is the Coupled General Circulation Model of the Canadian Center for Climate Modelling and Analysis (CCCma-CGCM1) and Amin et. al., (2017) investigated future hydrologic conditions over the Muda and Dungun River Basin by analysing river flows in the first and second halves of the 21<sup>st</sup> century using an ensemble of three different GCMs. Cruz et al. (2017) added that using one or a few downscaled GCMs for just one or two climate scenarios may not be enough; instead, a diversity of probable future climatic situations with an uncertainty measure are required to assess the implications of climate change and develop an adaptation plan.

### **2.11 Ensemble approach**

Hawkins and Sutton (2009) identified three elements that affect climate forecast uncertainty: model uncertainty, internal variability of the climate system, and scenario uncertainty. The internal variability of the climate system is defined as natural fluctuations that occur in the absence of any radiative forcing of the planet. The model uncertainty is defined as an uncertainty due to different responses of climate models to the same radiative forcing, and the scenario uncertainty is defined as an uncertainty in future emissions of

greenhouse gases. To assess these uncertainties, the ensemble technique was suggested, which uses diverse future estimates from several GCMs based on numerous scenarios (M.Z.M. Amin et al., 2017). Amin et al. (2017) and Amin et al. (2019) employed the ensemble method of three distinct coupled land-atmosphere-ocean GCMs based on the Fourth Assessment Report (AR4) of the United Nations Intergovernmental Panel on Climate Change (IPCC) to account for model uncertainty in climate change simulation (IPCC, 2007b). The applied models include the fifth generation ECHAM5 GCM from the Max Planck Institute (MPI) for Meteorology in Germany (Roeckner et al., 2006), the Coupled Ocean-Atmosphere General Circulation Model (CGCM2.3.2) from the Meteorological Research Institute (MRI) of Japan (Yukimoto et al., 2001), and the third-generation Community Climate System Model version 3 (CCSM3) from the University Corporation for Atmospheric Research (UCAR) in United States (Collins et al., 2006). Additionally, Amin et al. (2016) used an ensemble of the ECHAM GCM and CGCM2.3.2 to analyse the predicted changes in the hydrologic regime and water resources due to the projected worldwide shift in climate throughout the 21st century, as well as from planned land-use changes in Sabah and Sarawak. AR4, published in 2007, focused mainly on the physical science of climate change, while AR5, published in 2014, included more detailed analyses of impacts, vulnerability, and adaptation as well as a stronger focus on regional impacts. AR6, published in 2021, includes even more detailed analysis of climate science, impacts, and adaptation and mitigation strategies, including an assessment of the latest research on the impacts of 1.5° C warming. Additionally, AR6 includes a stronger focus on the impacts of climate change at the regional level, including an assessment of the specific impacts on cities and small islands. Since there is limited study on this ensemble dataset approach, this study proposes to apply the ensemble dataset from three sources which are the ECHAM5 GCM, CGCM2.3.2, and CCSM3 on SRB water resources projection. Furthermore, this approach has not been used in SWAT modelling in order to simulate water resource prediction in 2030, 2040 and 2050.

## 2.12 Conclusion

From the literature review, most researchers apply individual hydro-meteorological data, such as precipitation, as an input variable to a hydrological model. The issue arises when there are only a few ground stations in a watershed to feed the hydrological model. With the latest technology on satellite datasets, researchers now have an advantage to enhance the hydrological model, especially in the scarcely gauged area. Nowadays, cross-combining datasets is a new approach to enhancing the hydrological model.

However, only a few researchers applied cross-combined datasets from multiple sources, and the result shows a good performance in model simulation. Regionalisation is the optimum method for streamflow, sediment load, and water quality prediction in ungauged basins in order to establish the ideal model parameters. While the multi-site calibration method enhanced the simulation of the hydrological model and reduced the watershed uncertainty. In order to account for model uncertainty, the Regional Hydro Climate Model of Peninsular Malaysia (RegHCM-PM) using an ensemble approach of different future forecasts from numerous GCMs based on diverse scenarios was suggested for Malaysia's environment. Furthermore, this is a novel approach, and the datasets have been validated by NAHRIM, Malaysia's sole agency in charge of managing climate change data. Additionally, this approach has not been tested using the SWAT model and implemented in the SRB.

## CHAPTER 3

### STUDY AREA AND DATA

#### 3.1 Malaysia

Based on a globe map of the Koppen-Geiger climate classification, Malaysia is one of the countries in Southeast Asia that has a tropical rainforest climate. Other countries in the region that share this climate include Indonesia, Thailand, Singapore, the Philippines, and Brunei (Kottek et al., 2006). A tropical rainforest climate is typically hot, extremely humid, and wet with year-round rainfall (Holden, 2011). Located at coordinated 2° 30' N and 112° 30' E in the equatorial region, Malaysia's total land area is about 330,803 km<sup>2</sup>, the 66<sup>th</sup> largest country by total land area. Peninsular Malaysia and East Malaysia had a total area of 132, 631 km<sup>2</sup> and 198, 172 km<sup>2</sup> respectively, and they were divided from one another by the South China Sea as illustrated in Figure 3-1 (DID, 2011a). There are 2,986 river basins in Malaysia, with 189 serving as the primary river basins. 89 rivers out of 189 were in Peninsular Malaysia, 22 rivers were in Sarawak, and 78 rivers were in Sabah, with total length of more than 57,300 km (DID, 2009b). Malaysia is geographically flanked by two huge oceans: The Pacific Ocean to the east and the Indian Ocean to the west. With accompanying natural climatic fluctuation, these seas have a substantial impact on the climate of Malaysia (T. Tangang et al., 2012). Malaysia is categorised as having an equatorial climate with heavy and constant rainfall all year round. Malaysia's surface climate is affected by the southeast and northeast monsoons. From May through August, the southeast monsoon is characterised by low-level south-westerly winds, which is the driest phase. The whole nation of Malaysia is affected by droughts during this time period. While the northeast monsoon is driven by north-easterly winds, the wettest season often begins in November and is characterised by north-easterly winds. From November until February of the following year, there is usually flooding along the east coast of Peninsular Malaysia (Kelantan, Terengganu, and Pahang) (Sa'adin et al., 2016).



Figure 3-1: Map of Malaysia (source: google website)

### 3.1.1 Geography and Demography Characteristic

Malaysia is located at coordinates 2° 30' N and 112° 30 E which is in the Southeast Asia region, with a total land area of about 330,803 km<sup>2</sup> (NWRS, 2011). Malaysia is divided into two parts, which Peninsular Malaysia and East Malaysia. Land area of Peninsular Malaysia is about 132,631 km<sup>2</sup>, consisting of eleven states and two federal territories. The land area for East Malaysia is about 198,172 km<sup>2</sup> consisting of two states and one Federal Territory (NWRS, 2011). Perlis, Kedah, Pulau Pinang, Perak, Selangor, Melaka, Negeri Sembilan, Johor, Pahang, Terengganu, Kelantan, Federal Territory of Kuala Lumpur, and Putrajaya were in Peninsular Malaysia. While Sabah, Sarawak and the Federal Territory of Labuan are in East Malaysia. Figure 3-2 below shows the location of Malaysia on a global map. Pahang and Sarawak, with 36,137 km<sup>2</sup> and 124,450 km<sup>2</sup>, respectively, have the greatest land areas in Peninsular Malaysia and East Malaysia (NWRS, 2011). Approximately 330,803 km<sup>2</sup> of Malaysia's total land area is made up of water, including lakes, rivers, and other internal waters, making up 1,224 km<sup>2</sup> or 0.37% of the total land area. Malaysia is classified as a tropical rainforest country, with land cover approximately 99.63% of the total area and forest covering 58.2% (Saw Swee Hock, 2007).



Figure 3-2: Map of Malaysia in global view (source: <https://site.ieee.org/>)

### **3.1.2 Population**

Malaysia's population has been growing rapidly from 13.75 million in 1980 to 30.7 million in 2015 and is projected to increase to around 41.7 million in 2050 (WPP, 2017). Meanwhile, the Department of Statistics Malaysia (DOSM) projected an increasing population pattern to 49.7 million in 2050 (NWRS, 2011). Melaka has the highest urban population growth rate in more developed states, at 8.1%, followed by Selangor at 7.8%. While the lowest growth rate of only 1.4% is in the Federal Territory of Kuala Lumpur, in less developed states, the highest is 8.0% both in Sabah and the Federal Territory of Labuan and the lowest is in Kelantan with only 0.1 million for 1991-2000 which is 1.0% (NWRS, 2011). Figure 3-3 and Figure 3-4 show the population projections in Peninsular and East Malaysia until 2050.

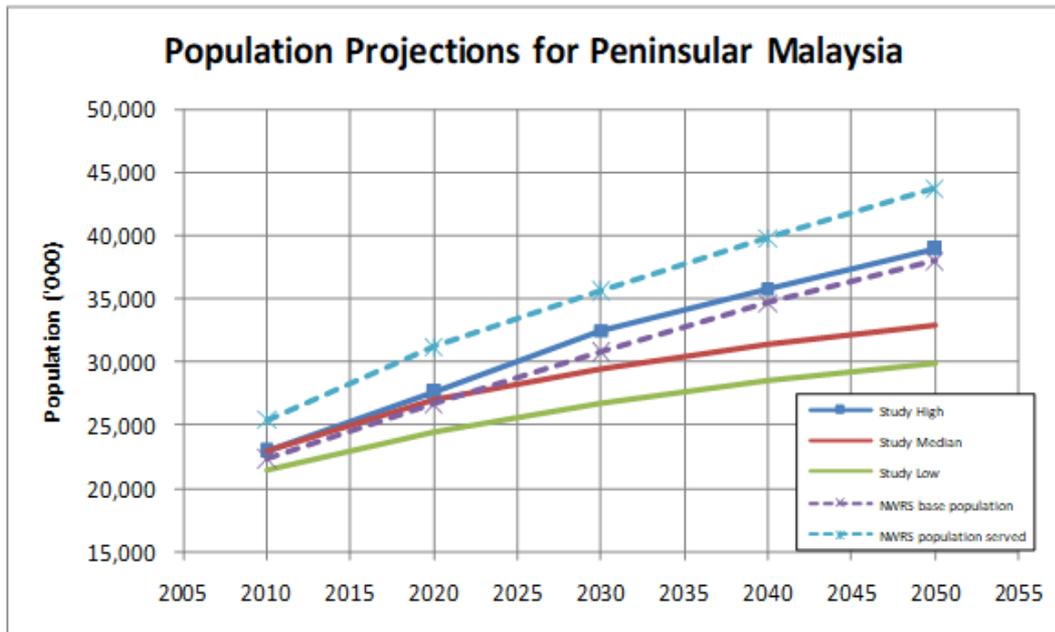


Figure 3-3: Population projections for Peninsular Malaysia (source: NWRS (2011))

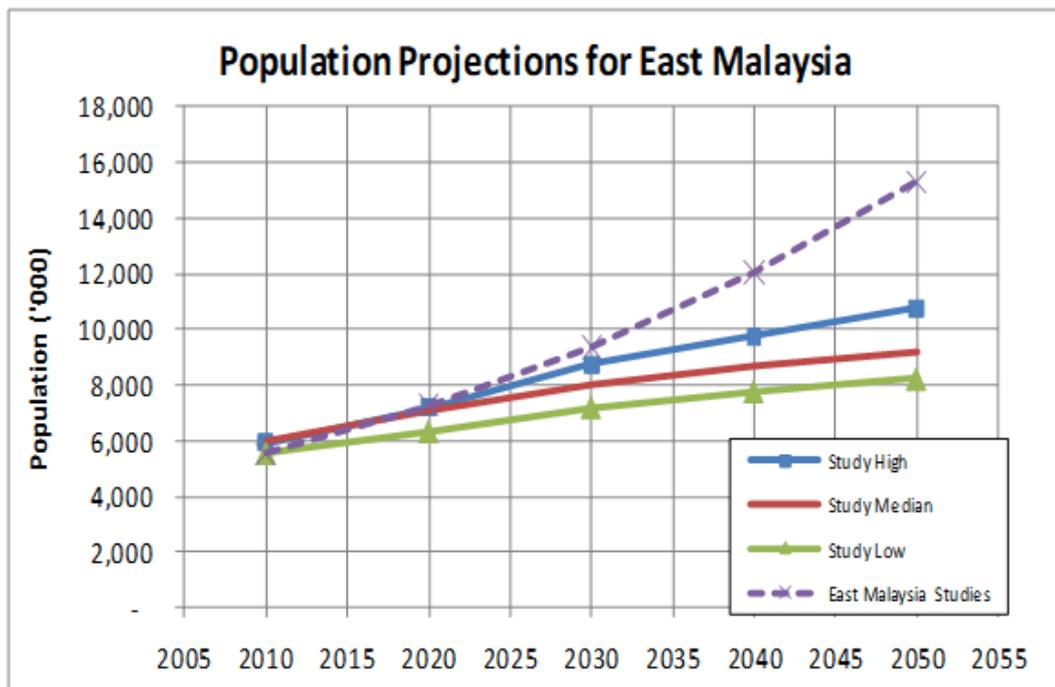


Figure 3-4: Population projections for East Malaysia (source: NWRS (2011))

### **3.1.3 Climate**

Malaysia is a tropical country that is near to the equator and has a monsoon climate with significant and continuous rainfall throughout the year (Sa'adin, Kaewunruen, & Jaroszweski, 2016). Geographically, Malaysia is encircled by the Pacific Ocean to the east and the Indian Ocean to the west. According to Tangang et. al., (2012), the northeast monsoon and the southeast monsoon both have an impact on Malaysia's surface climate. The southeast monsoon which begins in May and lasts through August, is distinguished by low-lying south-westerly winds and is the drier season. The entire nation of Malaysia was experiencing drought at the time. While the northeast monsoon, which lasts from November to February the following year and is the wettest season, is characterised by north-easterly winds. Flooding is unavoidable during this period, particularly on the east coasts of the three Malaysia states of Kelantan, Terengganu, and Pahang (Sa'adin et al., 2016). Malaysia received an average rainfall of about 2,940.60 mm/year with average rainfall of 2,945.50 mm/year in Peninsular Malaysia and 3,238.50 mm/year in East Malaysia (NWRS, 2011). Sa'adin et al. (2016) claim that both regional topography characteristics and seasonal wind flow patterns influence the pattern of rainfall distribution in Malaysia. Due to its proximity to the South China Sea, Peninsular Malaysia's east coast experiences greater rainfall, particularly during the monsoon season. With temperatures ranging from 21°C (70°F) to 32°C (90°F), Malaysia has an average annual evaporation of 1,250.30 mm, with the Federal Territory of Labuan recording the highest amount (1,480 mm) and Johor the lowest (1,130 mm) (NWRS, 2011).

### **3.2 State of Selangor**

The Malaysian federation is made up of a total of 13 states and three federal territories. East Malaysia is made up of two states and one federal territory, while Peninsular Malaysia is made up of 11 states and two federal territories. The State of Selangor is one of the thirteen states that make up Malaysia. It is situated on Peninsular Malaysia's west coast, 3 degrees 20 minutes north and 101 degrees 30 minutes east, directly facing the Straits of Malacca as illustrated in Figure 3-5.

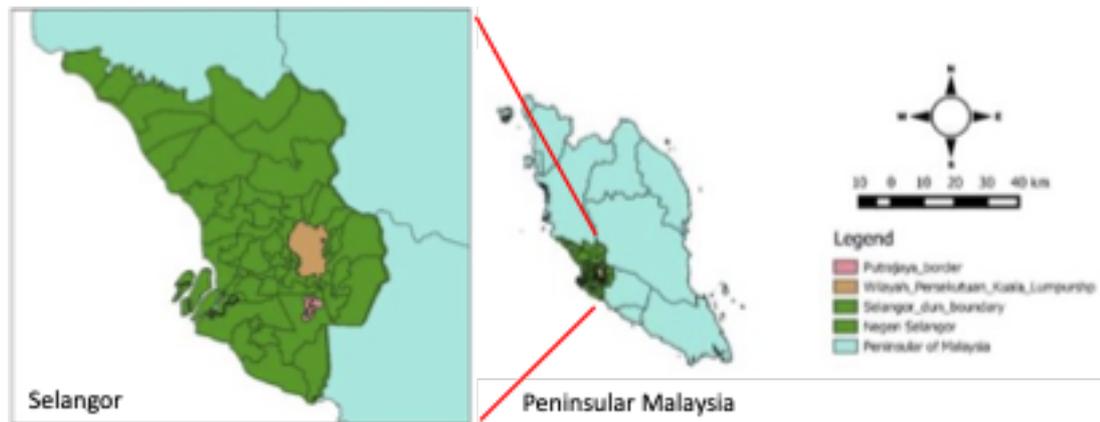


Figure 3-5: Map of state of Selangor (source: Yahya (2021))

Selangor is surrounded by the states of Negeri Sembilan to the south, Perak to the north, Pahang to the east, and the Straits of Malacca to the west. The overall land area amounts to around 8,396 km<sup>2</sup> in total. It includes the Federal Territory of Wilayah Persekutuan Kuala Lumpur and Putrajaya (DID, 2011b). There are seven main river basins in the state of Selangor, and they are named Sg. Langat, Sg. Selangor, Sg. Sepang, Sg. Klang, and Sg. Buloh (DID, 2011b). According to DOSM (2015), the SRB population is approximately 451,219 people, with Bumiputera constituting the majority at 266,219 (59%), Chinese at 76,707 (17%), Indians at 81,220 (18%), and non-citizens at 27,073 (6%). The Sungai Selangor Basin Management Plan 2007–2012 classified the river's principal uses as water supply, ecology, tourism and recreation, sand mining, aquaculture, and inland navigation. The river is renowned for other tourist destinations as well, including the Kampung Kuantan firefly habitat. Selangor is the most developed state with the highest economic contribution in Malaysia (Leeta, 2009). According to DOSM (2021), Selangor's Gross Domestic Product (GDP) in 2019 is RM 345,147 million and RM 326,841 million in 2020 at constant 2015 prices. The main economic pillars of Selangor are trade, industry, and services (NWRS, 2011). In addition, Selangor boasts the best infrastructure, the biggest population, the highest level of life, and the lowest poverty rate in the nation (Leeta, 2009). Total population in Selangor is approximately 6.52 million, with 0.3% average annual population growth rate (DOSM, 2021).

### **3.3 Selangor River Basin (SRB)**

Sg. Selangor or Selangor River Basin (SRB), is located in the upper section of the state of Selangor. It is the third biggest basin in the state of Selangor, after Sg. Langat and Sg. Bernam (LUAS, 2015). As shown in Figure 3-6, the catchment area for the SRB is about 2,200 km<sup>2</sup>, which corresponds to around 28 percent of the state. There are 110 kilometres of river stream within this region. From the slopes of Fraser's Hill to the Straits of Melaka, the SRB crosses the northeast part of Selangor to the southwest. The basin's headwaters are characterised by forested slopes and rugged mountains. Oil palms and rubber plants make up the majority of the vegetation in the largely undulating middle region. Before emptying into the Melaka Straits at Kuala Selangor, the river meanders through a peat swamp forest in low-lying coastal regions (DID, 2011b). SRB has 13 major river tributaries, including Sg. Kanching, Sg. Kerling, Sg. Sembah, Sg. Batang Kali, Sg. Rening, Sg. Luit, Sg. Kul, Sg. Gumut, Sg. Darah, Sg. Kubu, Sg. Gerachi, Sg. Peretak, and Sg. Tinggi. SRB has an additional eight lesser river tributaries in SRB, including Sg. Beletak, Sg. Buloh, Sg. Liam, Sg. Serendah, Sg. Rawang, Sg. Garing, Sg. Kundang, and Sg. Ayer Hitam (LUAS, 2011).

The SRB is divided into four distinct zones: the forest, agricultural (including animal husbandry, aquaculture, oil palm and rubber cultivation), built-up areas (containing residential, commercial, industrial, mixed development, institutions, mining, sand mining, and quarries), and water bodies (rivers, lakes, and ponds) (LUAS, 2015). SRB is a crucial natural resource that supports a variety of economic activities in the state. Securing around 67% (3,018 MLD) of the total water supply requirement for Kuala Lumpur and Selangor through the Sungai Selangor Water Supply Scheme is one of the priorities (LUAS, 2015). This river passes through the six main towns of Kuala Kubu Baru, Rawang, Serendah, Rasa, Bestari Jaya, and Kuala Selangor. Kuala Kubu Baru is the largest of these cities (LUAS, 2011). The SRB consisted of three districts and three administrative bodies: Hulu Selangor, Gombak, and Kuala Selangor. Figure 3-6 shows the boundary of the Selangor River Basin (SRB).



Figure 3-6: Selangor River Basin boundary catchment (source: LUAS (2015))

### 3.3.1 Dam and reservoir

Sg. Tinggi's dam and Sg. Selangor dam were two dams specifically used for water supply in SRB. Sg. Tinggi dam catchment area is about 40 km<sup>2</sup>, located in the Bukit Tarek forest reserve, the upper reaches of Sg. Buloh, as depicted in Figure 3-7(c). It is constructed close to Ladang at the riverheads of Sg. Buloh, the first order tributary of Sg. Selangor, with the dam capacity is about 51,694 MLD (LUAS, 2015). As part of Sungai Selangor Phase 1 (stage 2), the Sg. Tinggi dam was commissioned in 1996 to manage the flow at the current water intake, Batang Berjuntai, located approximately 30 kilometres downstream. A 150 MLD pumping station has been installed to aid in the replenishing of the Sg. Tinggi reservoir from the Sg. Selangor mainstem (DID, 2011b). Sg. Tinggi can now hold 114.5 MCM of water because the height of the dam was raised (LUAS, 2015).

The Sg. Selangor dam, which was most recently constructed in Selangor, is situated in the upper region of SRB, which is in the Kuala Kubu Baru forest reserve and has a catchment area of 197 km<sup>2</sup> (DID, 2011b). It is a rockfill dam with a 26,247 MLD dam capacity. The dam began operations in 2003 as part of Sungai Selangor Phase 3 in order to supply an additional 1,100 MLD (12.7 m<sup>3</sup>/s) and boost the abstraction capacity from Sg. Selangor from 1,900 MLD (22 m<sup>3</sup>/s) to 3,000 MLD (34.7 m<sup>3</sup>/s). The Sg. Selangor dam has a storage capacity of 230 million cubic metres (LUAS, 2015). Both dams were gazetted as water detention dams for water supply in 2009 and were managed by SPLASH (Syarikat Pengeluar Air Sungai Selangor Sdn. Bhd). According to (DID, 2011b), the Sg. Selangor dam and Sg. Tinggi dam will control river flow to generate a total net system output of 2,890 MLD, which is less than the total abstractions from all treatment plants (2,981.5 MLD).

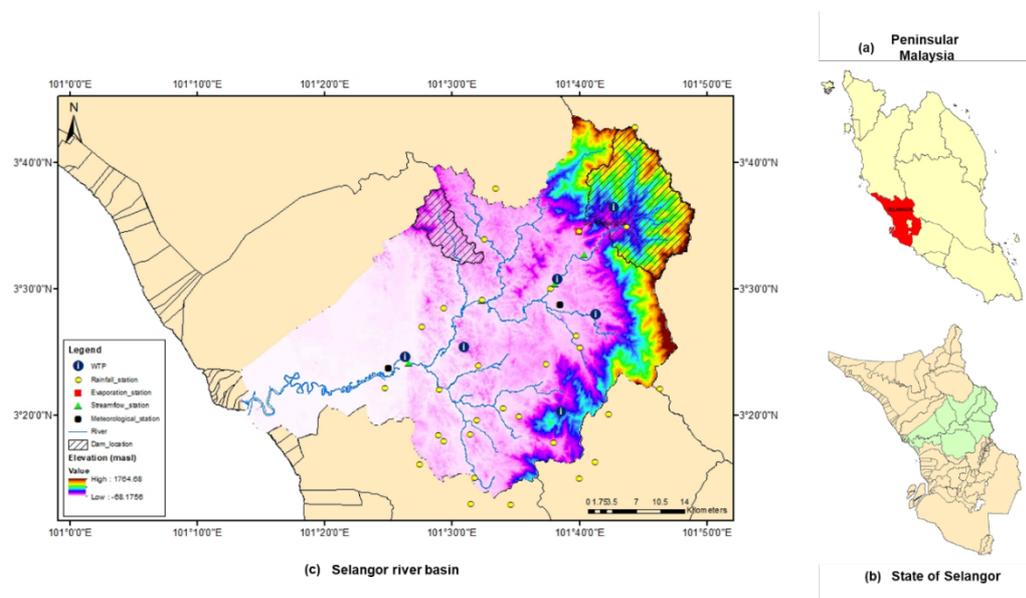


Figure 3-7: Location map of (a) Peninsular Malaysia, (b) State of Selangor and (c) Selangor River basin

### 3.3.2 Surface water contribution zones

The Selangor Water Management Authority (LUAS), the state organisation responsible for regulating water resources, has divided SRB into three zones. This is done to ensure that water resources managers can understand where the river basin's water consumption and storage are concentrated and plan for mitigating circumstances like drought conditions or widespread pollution when they occur (LUAS, 2015). Three zones were established to analyse the contribution of surface water resources: the effective zone, the semi-effective zone, and the non-effective zone. When precipitation is totally captured and stored as much as possible in dam reservoirs, that area is known as the effective zone. As depicted in Figure 3-8, the semi-effective zone is the area where surface water from rainfall contributes to river flow without the benefit of significant storage, while the non-effective zone is covered by catchments that do not contribute to any abstraction for public water supply or major irrigation. This study focused on the effective and semi-effective zone.

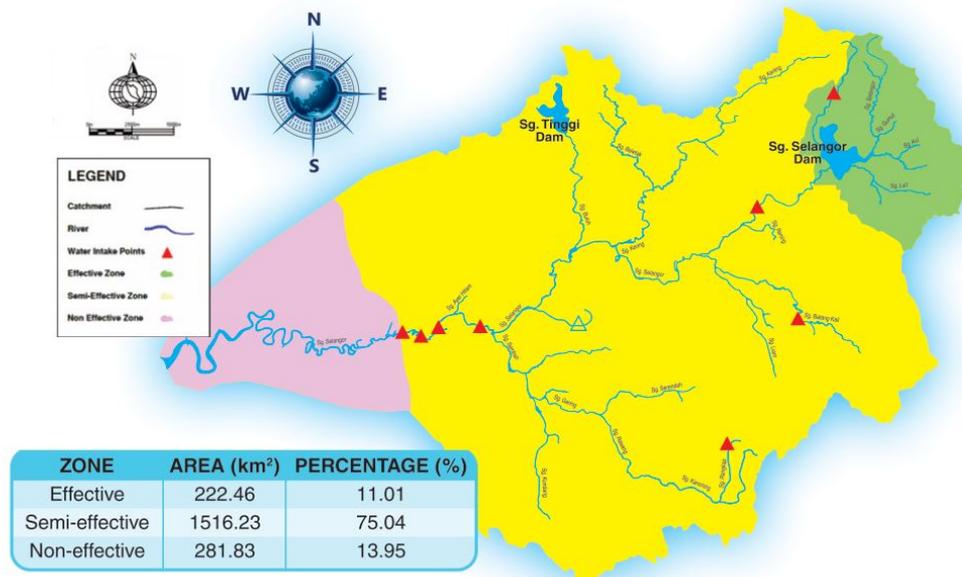


Figure 3-8: Selangor river basin surface water contribution zones (source: state of the river report (2015))

### 3.4 Water Resources

Water resources are a vital natural resource for all life on Earth. All kinds of water resources are important on this planet, but for human life, the accessibility of freshwater is most essential. Water resources should be looked at least from three key perspectives which are (i) water as a resource, (ii) water for everyday activities, and (iii) impacts from and on water resources. When considering water as a resource, one must consider the dynamic processes that replenish catchments and water bodies all the way up to the hydrological regimes and hydrogeological conditions. Meanwhile, the use of water for daily activities shows that water resources serve a variety of purposes that benefit both people and the environment. Floods, pollution, erosion, and siltation are all major hazards cause by and affecting water resources (MNRE, 2012).

In total, saltwater accounts for approximately 97% of the global water composition, with freshwater accounting for the remaining 3%. Two-thirds, or 68.7%, of freshwater is classified as glaciers and ice caps, 30.1% as groundwater, 0.9% as other, and only 0.3% as fresh surface water. Despite 0.3% of fresh surface water, 87% is from lakes, 11% is from swamps and only 2% from rivers (Du Plessis, 2017). Furthermore, the vast majority of countries in the world rely on rivers as their water supply. The percentage distribution of Earth's water is illustrated in Figure 3-9.

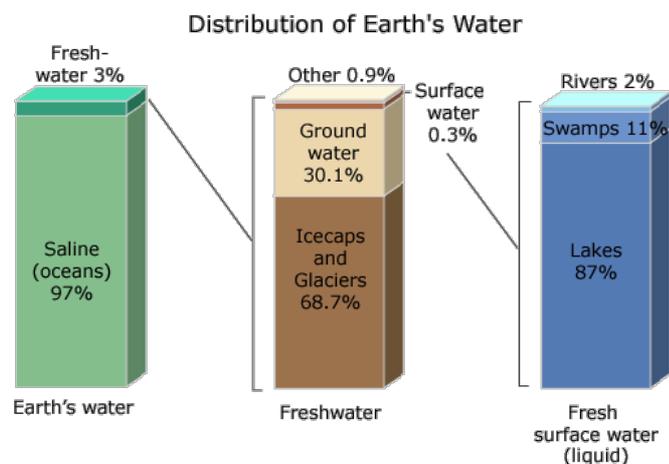


Figure 3-9: Distribution of Earth's Water (source: Gleick, P. H., (1996))

The small amounts of accessible freshwater can have a great impact on living things if they are not monitored and managed in a good manner. According to Ahmed et al. (2014), 166 million people in 18 countries have water scarcity, while an additional 270 million people in 11 countries experience water stress. According to projections in the 2018 United Nations World Water Development Report, the fast-expanding global population will cause the world's water consumption to rise by about one-third by 2050. This will have an impact on more than 2 billion people worldwide (Azoulay Hougbo, 2018), and as water demand rises in large cities (González-Zeas et al., 2019) it will put further strain on water supplies. Water resources across the world are becoming increasingly susceptible as a result of rising water demand brought on by factors like as population growth, industrial development, increased food production, pollution from numerous human activities, climate change, and affects on land use. (Carroll et al., 2013; Odusanya et al., 2019). Consequently, as stated in the River Basin Management Plans (RBMPs), special mitigation and adaptation strategies for the management of water resources are required to balance the water needs of various users (Rivas-Tabares et al., 2019).

#### **3.4.1 Water Resources in Malaysia**

Malaysia receives about 973 billion m<sup>3</sup> of annual rainfall, of which 413 billion m<sup>3</sup> return to the atmosphere as evaporation, 63 billion m<sup>3</sup> move to groundwater recharge, and 495 billion m<sup>3</sup> transform as a surface runoff. It is estimated that about 70% of the 495 billion m<sup>3</sup> of surface runoff is lost to the sea by way of flood discharge (NWRS, 2011). Most of Malaysia's main towns and cities were founded as riverside villages, and rivers have played a significant role in the history of the nation. More than 98% of the water demand in Malaysia comes from rivers and reservoirs, and less than 2% comes from groundwater (SNC, 2015). DID (2009) states that Malaysia has 189 primary river basins, including 89 in Peninsular Malaysia and 100 in East Malaysia. These river basins branch out into 1,800 rivers, with a combined length of more than 57,300 km. Legally, rivers and water resources in Malaysia are

under the jurisdiction of the state government, except for rivers shared by more than one state, where they will be taken over by the federal government.

### 3.5 Water Demand in Malaysia

Malaysia's economy is rapidly developing, and the population is expected to reach approximately 50 million in 2050. According to Gonzalez-Zeas et al. (2019), the future stress on water resources and water demand, particularly in the main cities, would be exacerbated by the expanding population. Increasing population, expansion in urbanisation, industrialisation, and irrigation will contribute to water scarcity in Malaysia. Therefore, the demand for water for daily needs will subsequently increase (NWRS, 2011). According to NWRS (2011), total consumptive water demand for all sectors in Malaysia is 17,205 MCM in 2020 with 13,664 MCM in Peninsular Malaysia and 3,541 MCM in East Malaysia. The demand of water for all sectors including potable water, irrigated paddy, non-paddy crops, livestock and fisheries is predicted to increase 18,233 MCM in 2050 with 14,488 MCM in Peninsular Malaysia and 3,745 MCM in East Malaysia as listed in Table 3-4.

Table 3-1: Total consumptive water demand against total surface water availability for all sectors (source: NWRS (2011))

| States                | Land Area sq km | Total Consumptive Water demand (MCM ) |               |               |               |               |
|-----------------------|-----------------|---------------------------------------|---------------|---------------|---------------|---------------|
|                       |                 | 2010                                  | 2020          | 2030          | 2040          | 2050          |
| Pertlis               | 821             | 306                                   | 299           | 286           | 284           | 281           |
| Kedah                 | 9,500           | 2,922                                 | 2,976         | 2,842         | 2,873         | 2,876         |
| Pulau Pinang          | 1,048           | 765                                   | 829           | 835           | 874           | 894           |
| Kelantan              | 15,099          | 1,632                                 | 1,619         | 1,586         | 1,600         | 1,604         |
| Terengganu            | 13,035          | 884                                   | 975           | 970           | 999           | 1,026         |
| Perak                 | 21,035          | 1,949                                 | 1,923         | 1,798         | 1,801         | 1,811         |
| Selangor              | 8,396           | 2,238                                 | 2,491         | 2,570         | 2,760         | 2,922         |
| Pahang                | 36,137          | 726                                   | 946           | 897           | 911           | 959           |
| Negeri Sembilan       | 6,686           | 340                                   | 361           | 358           | 366           | 374           |
| Melaka                | 1,664           | 323                                   | 366           | 376           | 409           | 439           |
| Johor                 | 19,210          | 715                                   | 881           | 1,033         | 1,164         | 1,301         |
| <b>Pen Malaysia</b>   | <b>132,631</b>  | <b>12,800</b>                         | <b>13,664</b> | <b>13,551</b> | <b>14,040</b> | <b>14,488</b> |
| Sabah                 | 73,631          | 912                                   | 1,356         | 1,392         | 1,442         | 1,469         |
| Sarawak               | 124,450         | 1,054                                 | 2,162         | 2,125         | 2,175         | 2,247         |
| WP Labuan             | 91              | 18                                    | 24            | 26            | 28            | 29            |
| <b>East Malaysia</b>  | <b>198,172</b>  | <b>1,985</b>                          | <b>3,541</b>  | <b>3,542</b>  | <b>3,645</b>  | <b>3,745</b>  |
| <b>Total Malaysia</b> | <b>330,803</b>  | <b>14,785</b>                         | <b>17,205</b> | <b>17,093</b> | <b>17,685</b> | <b>18,233</b> |

The highest water demand in 2020 is from the state of Kedah with 2,976 MCM, followed by the state of Selangor with 2,491 MCM. However, the highest projected water demand in 2050 is from the state of Selangor with 2,922 MCM compared to the state of Kedah with 2,876 MCM. Meanwhile, the lowest water demand is from the Federal Territory of Labuan, with only 24 MCM in 2020 and 29 MCM in 2050. In terms of water sectors, currently the highest water demand in Malaysia comes from the irrigation sector. However, the highest water demand from the potable water sector is expected to reach 28.1 mm/year, higher than irrigated paddy's 21.8 mm/year in 2050. Potable water demand in Malaysia consists of four main sectors, which are domestic, industrial, commercial, and institutional. Figure 3-10 illustrate the projected water demand for various sectors compared to available runoff in Malaysia.

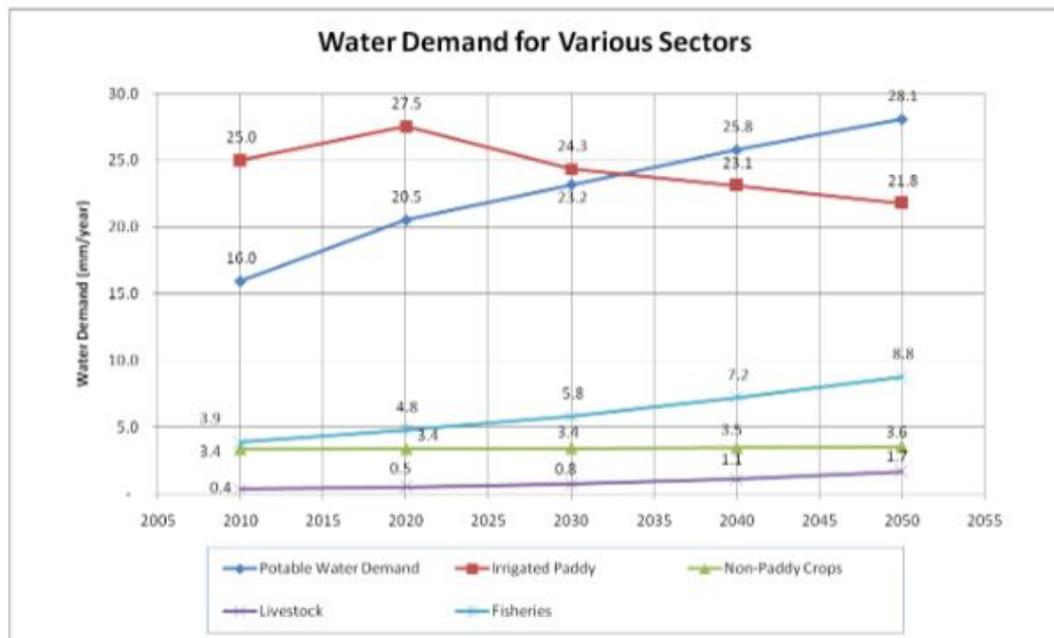


Figure 3-10: Water demand for all sector compared with the available runoff (source: NWRS (2011))

### 3.5.1 Water Demand in Selangor

According to NWRS (2011), potable water demand in the state of Selangor is 1,787 MCM in 2020 and expected to increase up to 2,364 MCM in 2050 which is the highest demand among the other states in Malaysia. Meanwhile, projected water demand for irrigation, non-paddy cultivation, livestock, and fisheries in 2050 are 482 MCM, 48 MCM, 27.9 MCM, and 351.5 MCM, respectively. Total consumptive water demand for various sectors in Selangor is projected to increase by about 21% from 2,490 MCM in 2020 to 2,922 MCM in 2050 as listed in Table 3-5. The potable water demand is projected to increase drastically compared to other sectors as illustrated in Figure 3-11. Meanwhile, the irrigated paddy sector is projected to decrease water demand by approximately 24% in 2010.

Table 3-2: Total water demand for various sectors in Selangor (source: NWRS (2011))

| Sectors                               | 2010                      | 2020         | 2030         | 2040         | 2050         | 2010                              | 2020         | 2030         | 2040         | 2050         |
|---------------------------------------|---------------------------|--------------|--------------|--------------|--------------|-----------------------------------|--------------|--------------|--------------|--------------|
|                                       | Water Demand MCM per year |              |              |              |              | Water Demand mm rainfall per year |              |              |              |              |
| Potable Water Demand                  | 1,474                     | 1,787        | 1,960        | 2,181        | 2,364        | 175.5                             | 212.8        | 233.5        | 259.8        | 281.6        |
| Irrigated Paddy                       | 720                       | 655          | 555          | 516          | 482          | 85.8                              | 78.0         | 66.1         | 61.5         | 57.4         |
| Non-Paddy Crops                       | 36.0                      | 37.0         | 39.0         | 43.0         | 48.0         | 4.3                               | 4.4          | 4.6          | 5.1          | 5.7          |
| Livestock                             | 8.6                       | 11.3         | 14.9         | 20.2         | 27.9         | 1.0                               | 1.3          | 1.8          | 2.4          | 3.3          |
| <b>Total consumptive water demand</b> | <b>2,238</b>              | <b>2,490</b> | <b>2,569</b> | <b>2,760</b> | <b>2,922</b> | <b>266.6</b>                      | <b>296.6</b> | <b>306.0</b> | <b>328.7</b> | <b>348.0</b> |
| Fisheries                             | 159.2                     | 194.1        | 236.6        | 288.4        | 351.5        | 19.0                              | 23.1         | 28.2         | 34.3         | 41.9         |
| <b>Total demand</b>                   | <b>2,397</b>              | <b>2,684</b> | <b>2,806</b> | <b>3,048</b> | <b>3,274</b> | <b>285.5</b>                      | <b>319.7</b> | <b>334.2</b> | <b>363.1</b> | <b>389.9</b> |

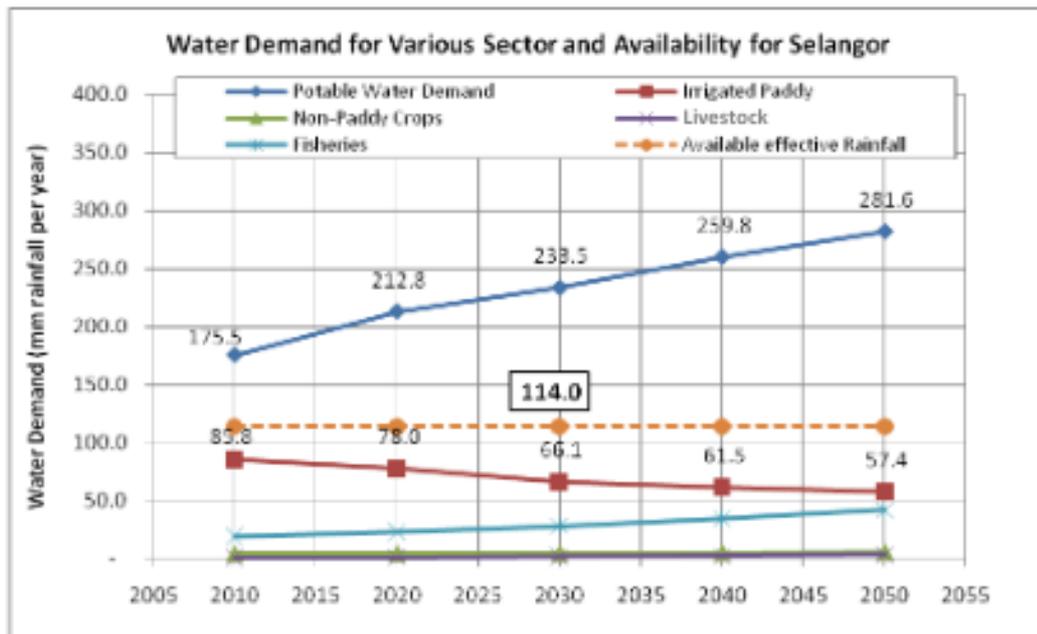


Figure 3-11: Water demand for various sectors in Selangor (source: NWRS (2011))

### 3.6 Water Crisis in Selangor

In 2004, The American Meteorological Society (2004) adopted a four-type approach for classifying droughts based on the type of water shortfall. This categorization defines meteorological, hydrological and agricultural droughts as periods of insufficient precipitation, river flow, groundwater, and soil moisture, respectively. These conditions are referred to as environmental droughts. The socio-economic drought is the fourth form of drought, and it results from the inability of water supply infrastructure to keep up with demand. (Vasiliades et al., 2017). Figure 3-12 depicted the drought classification system derived from (Stahl et al., 2008).

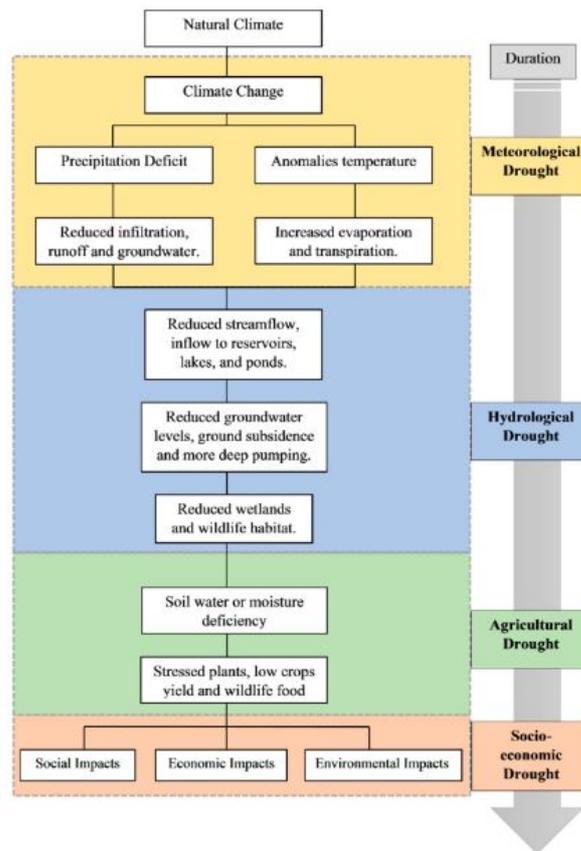


Figure 3-12: Drought classification system. (source: Hasan et al., (2019))

Malaysia suffered a particularly hard year of drought in recent years, and the frequency of drought is increasing over the year (Sanusi et al., 2015). Peninsular Malaysia experiences frequent droughts due to climate change, despite the region receiving an average of about 2500 mm of precipitation annually (Chinnasamy & Ganapathy, 2018). Ahmadi (2019) stated that the most challenging aspect of managing water resources in urban planning is facing the drought phenomenon, which is a major concern among researchers, politicians, and stakeholders (Ahmadi & Zarghami, 2019). The first three months of 2014 were particularly difficult in Selangor as a result of the state's rivers drying up and water levels dropping to dangerously low levels in seven of its dams, including as low as 31% of capacity at the Sg. Selangor Dam, which provides more than 60% of the state's water (Hong & Hong, 2016).

This phenomenon was widespread and affected about four million residents in major ways, as well as the industries in Kuala Lumpur, Petaling, Gombak, and Hulu Selangor (Boelee et al., 2017). This situation forced the Selangor state government to implement the water restriction scheme to address the lack of water resources in Selangor (Hong & Hong, 2016). Furthermore, this scenario became an important factor in increasing demand for water (Boelee et al., 2017). According to NWRS (2011), total consumptive water demand in state of Selangor is expected to increase up to 348 mm/year in 2050 as listed in Table 3-6 with a water deficit of about 234 mm/year.

Table 3-3: Total water demand in state of Selangor (mm rainfall per year) (source: NWRS (2010))

| Year | Runoff | Estimate % available (15%) | Total consumptive water demand | Deficit |
|------|--------|----------------------------|--------------------------------|---------|
| 2010 | 760    | 114                        | 266.6                          | (152.6) |
| 2020 | 760    | 114                        | 296.6                          | (182.6) |
| 2030 | 760    | 114                        | 306.0                          | (192.0) |
| 2040 | 760    | 114                        | 328.7                          | (214.7) |
| 2050 | 760    | 114                        | 348.0                          | (234.0) |

### 3.7 Hydrological data

The department of Irrigation and Drainage Malaysia (DID), a government organisation, administers and maintains hydrological data in Malaysia. As of 2021, Malaysia had 1,458 hydrological stations operating under the JPS National Network, including 977 rainfall stations, 150 water level stations, 118 gauging stations, 118 sediment stations, 75 water quality stations, and 25 evaporation stations throughout the country (DID, 2021). There are 177 hydrological stations in the state of Selangor, including 120 rainfall stations, 8 streamflow stations, and 49 others (DID, 2021). However, for SRB hydrological modelling purposes, only 30 rainfall stations and 3 streamflow stations, as depicted in Figure 3-13, are considered. Rainfall and streamflow data were first operated in 1918 and 1921, respectively. Due to the most readily available information across all hydrological stations, this study used hydrological data from the years 2008 to 2018 as the study period. Information on each hydrological station is listed in Table 3-4.

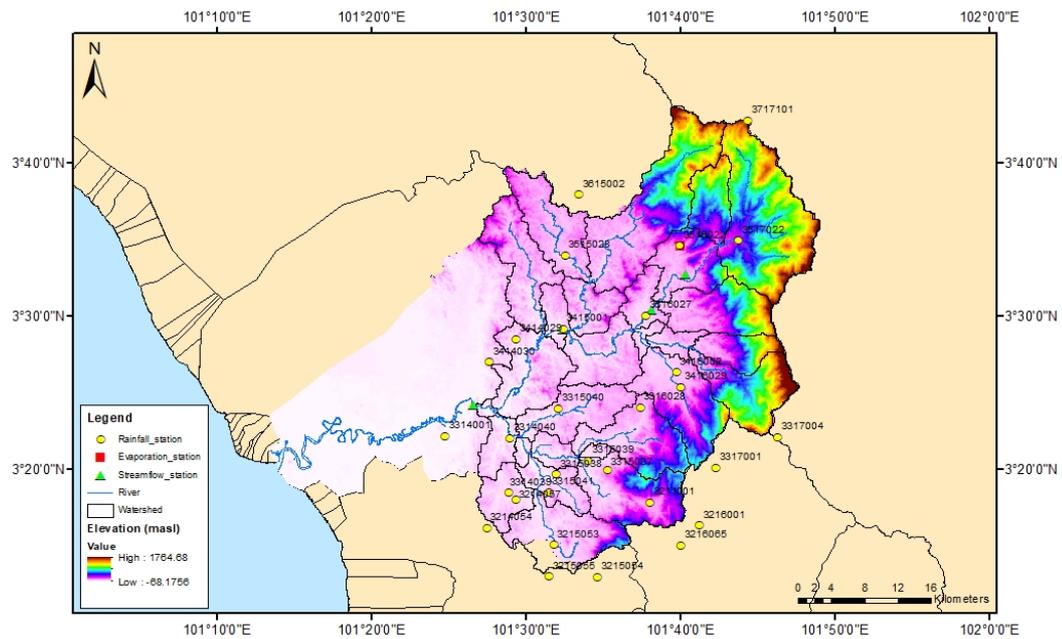


Figure 3-13: Hydrological stations in Sg. Selangor basin (source: DID Malaysia)

Table 3-4: List of hydrological station in Sg. Selangor basin (source: DID Malaysia)

| NO. | STATION_ID | STATION_NA                            | LAT_Y1   | LONG_X1    | STATION_TY  | STATUS   | YEAR_OPEN | DAERAH         | STATE_1                  |
|-----|------------|---------------------------------------|----------|------------|-------------|----------|-----------|----------------|--------------------------|
| 1   | 3211001    | Taman Templer at Selangor             | 3.296944 | 101.632806 | Rainfall    | Active   | 2004      | Gombak         | Selangor                 |
| 2   | 3214054    | Ldg. Tuan Mee at Selangor             | 3.269194 | 101.457111 | Rainfall    | Active   | 1/1/1918  | Kuala Selangor | Selangor / Kuala Lumpur  |
| 3   | 3214057    | Bandar Tasik Puteri at Selangor       | 3.299722 | 101.488333 | Rainfall    | Active   | 2006      | Gombak         | Selangor                 |
| 4   | 3215053    | Kg. Setia Kuang at Selangor           | 3.250833 | 101.529167 | Rainfall    | Active   | 2006      | Gombak         | Selangor                 |
| 5   | 3312042    | Ldg. Bkt Belimbing at Selangor        | 3.395833 | 101.274667 | Rainfall    | Active   | 7/1/1920  | Kuala Selangor | Selangor                 |
| 6   | 3312045    | Km. 45.5 Jln Kelang/ at K.Selangor    | 3.336278 | 101.256194 | Rainfall    | Active   | 1946      | Kuala Selangor | Selangor                 |
| 7   | 3313040    | Ldg. Bkt Talang at Bhg Kg Baru        | 3.395278 | 101.311778 | Rainfall    | Active   | 1/1/1928  | Kuala Selangor | Selangor                 |
| 8   | 3313043    | Ldg. Kuala Selangor at Selangor       | 3.307778 | 101.304972 | Rainfall    | Active   | 1935      | Kuala Selangor | Selangor                 |
| 9   | 3313060    | Ldg. Sg Buloh at Selangor             | 3.308722 | 101.321000 | Rainfall    | Active   | 1930      | Kuala Selangor | Selangor                 |
| 10  | 3314001    | Rmh Pam JPS Jaya Setia at Selangor    | 3.368889 | 101.412194 | Rainfall    | Active   | 1972      | Kuala Selangor | Selangor                 |
| 11  | 3314039    | Batu Arang at Selangor                | 3.307500 | 101.481389 | Rainfall    | Active   | 2004      | Gombak         | Selangor                 |
| 12  | 3314040    | Jln Gombak/K.Selangor at Selangor     | 3.366861 | 101.482444 | Rainfall    | Active   | 2004      | Gombak         | Selangor                 |
| 13  | 3315037    | Taman Bukit Rawang at Selangor        | 3.331944 | 101.586944 | Rainfall    | Active   | 2004      | Gombak         | Selangor                 |
| 14  | 3315038    | Country Home at Selangor              | 3.327222 | 101.531944 | Rainfall    | Active   | 2004      | Gombak         | Selangor                 |
| 15  | 3315039    | Taman Garing Utama at Selangor        | 3.342500 | 101.566111 | Rainfall    | Active   | 1/1/2005  | Hulu Selangor  | Selangor                 |
| 16  | 3315040    | Kg. Sungai Buaya at Selangor          | 3.398889 | 101.533889 | Rainfall    | Active   | 2006      | Hulu Selangor  | Selangor                 |
| 17  | 3315041    | Taman Desa Kundang at Selangor        | 3.308333 | 101.523611 | Rainfall    | Active   | 2007      | Gombak         | Selangor                 |
| 18  | 3316028    | Ldg. Sg Gapi at Selangor              | 3.400278 | 101.622500 | Rainfall    | Active   | 1970      | Hulu Selangor  | Selangor                 |
| 19  | 3412041    | Ldg. Raja Musa at Selangor            | 3.410278 | 101.283889 | Rainfall    | Active   | 1939      | Kuala Selangor | Selangor                 |
| 20  | 3414029    | Ldg. Sg Tinggi at Bhg Mary Selangor   | 3.474444 | 101.488694 | Rainfall    | Active   | 1/1/1927  | Hulu Selangor  | Selangor                 |
| 21  | 3414030    | Ldg. Hopeful at Selangor              | 3.449722 | 101.460083 | Rainfall    | Active   | 1/1/1928  | Kuala Selangor | Selangor                 |
| 22  | 3414031    | Pengorekan Bijih Berju at Selangor    | 3.422222 | 101.456944 | Rainfall    | Inactive | 1970      | Kuala Selangor | Selangor                 |
| 23  | 3414421    | Sg. Selangor at Rantau Panjang        | 3.402778 | 101.443056 | Water Level | Active   | 1921      | Kuala Selangor | Selangor                 |
| 24  | 3415001    | Kampung Timah at Selangor             | 3.485556 | 101.539167 | Rainfall    | Active   | 2005      | Hulu Selangor  | Selangor                 |
| 25  | 3415401    | Sg. Selangor at Kg. Timah             | 3.485556 | 101.539167 | Water Level | Active   | 2010      | Hulu Selangor  | Selangor                 |
| 26  | 3416001    | Ulu Yam (A,b) at Selangor             | 3.460667 | 101.637333 | Rainfall    | Inactive | 1974      | Hulu Selangor  | Selangor                 |
| 27  | 3416002    | Kg. Kalong Tengah (Ab) at Selangor    | 3.438333 | 101.661389 | Rainfall    | Active   | 8/1/1978  | Hulu Selangor  | Selangor                 |
| 28  | 3416026    | Ldg. Rasa at Selangor                 | 3.494444 | 101.602778 | Rainfall    | Inactive | 2003      | Hulu Selangor  | Selangor                 |
| 29  | 3416029    | Taman Desa Kelisa at Selangor         | 3.422222 | 101.666389 | Rainfall    | Active   | 2006      | Hulu Selangor  | Selangor                 |
| 30  | 3515028    | Lembah Beringin at Selangor           | 3.565278 | 101.541389 | Rainfall    | Active   | 2004      | Hulu Selangor  | Selangor                 |
| 31  | 3516022    | Loji Air K.Kubu Bahru at Selangor     | 3.575833 | 101.665556 | Rainfall    | Active   | 6/1/1942  | Hulu Selangor  | Selangor                 |
| 32  | 3516027    | Taman Rasa at Selangor                | 3.500278 | 101.628611 | Rainfall    | Active   | 2005      | Hulu Selangor  | Selangor                 |
| 33  | 3516322    | Loji Air Kuala Kubu Bahru             | 3.575861 | 101.665611 | Evaporation | Active   | 1972-2020 | Hulu Selangor  | Selangor                 |
| 34  | 3516422    | Sg. Selangor at Rasa Selangor         | 3.506944 | 101.634722 | Water Level | Inactive | 1961      | Hulu Selangor  | Selangor                 |
| 35  | 3516424    | Sg. Selangor at Ampang Pecah Selangor | 3.541417 | 101.662333 | Water Level | Active   | 1998      | Hulu Selangor  | Selangor                 |
| 36  | 3517022    | Kampung Pertak at Selangor            | 3.581667 | 101.728056 | Rainfall    | Active   | 2004      | Hulu Selangor  | Selangor                 |
| 37  | 3717101    | Bukit Fraser at Selangor              | 3.711944 | 101.738333 | Rainfall    | Active   | 2006      | Raub           | Pahang / Negeri Sembilan |

The average annual rainfall for Malaysia is about 3,000 mm/year, while the evaporation and surface runoff are 1,700 mm/year and 1,000 mm/year respectively (DID, 2009a). Particularly in the state of Selangor, the average annual rainfall, evaporation, and surface runoff recorded by DID is 2,190 mm/year, 1,280 mm/year, and 760 mm/year respectively (DID, 2011a). According to the SRB rainfall pattern from 1971 to 2015, the basin's average annual rainfall fell between 1,600 and 2500 mm, as illustrated in Figure 3-14. Whereas, based on the interested study area, analysis on 30 rainfall stations listed in Table 3-4 from 2008 to 2018 shows the average annual rainfall in SRB is 2,469 mm, with the highest record of 2,785 mm in 2018 and the lowest recorded of 2,124 mm in 2016, as shown in Figure 3-15. The depth of the rainfall was greater at the end of the year, with 420 mm in November 2015, followed by 411 mm in November 2012, compared to the beginning, with only 6.3 mm recorded in February 2014.

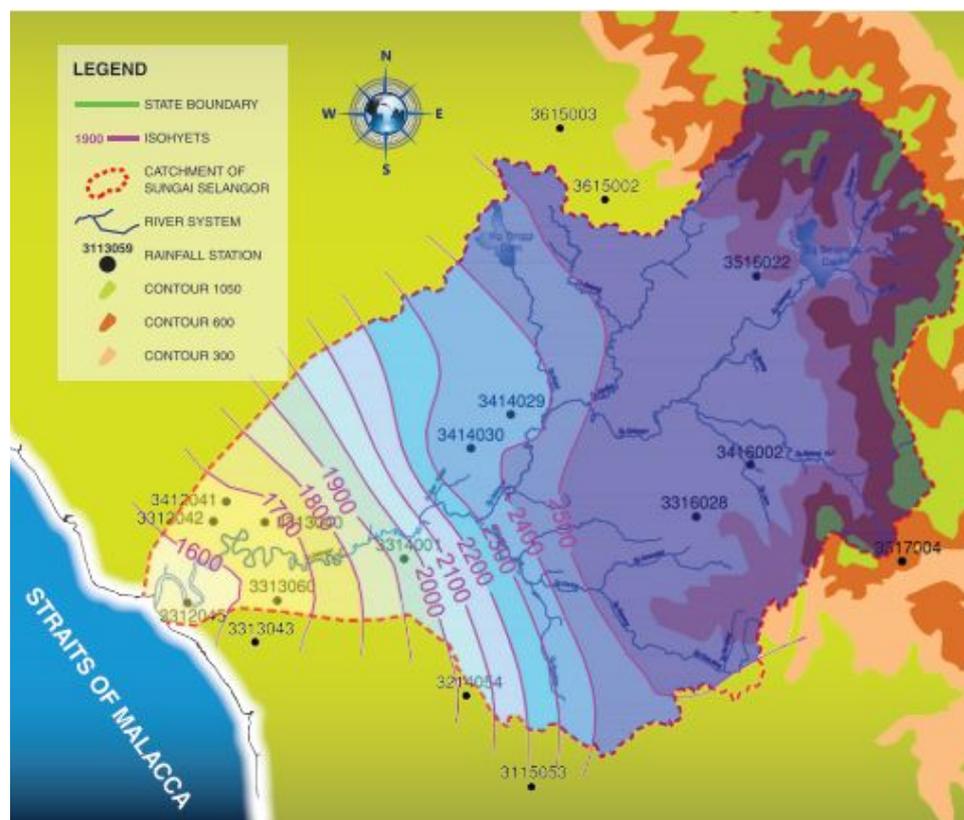


Figure 3-14: Rainfall pattern of Selangor River Basin from 1971–2015 (source: LUAS (2011))

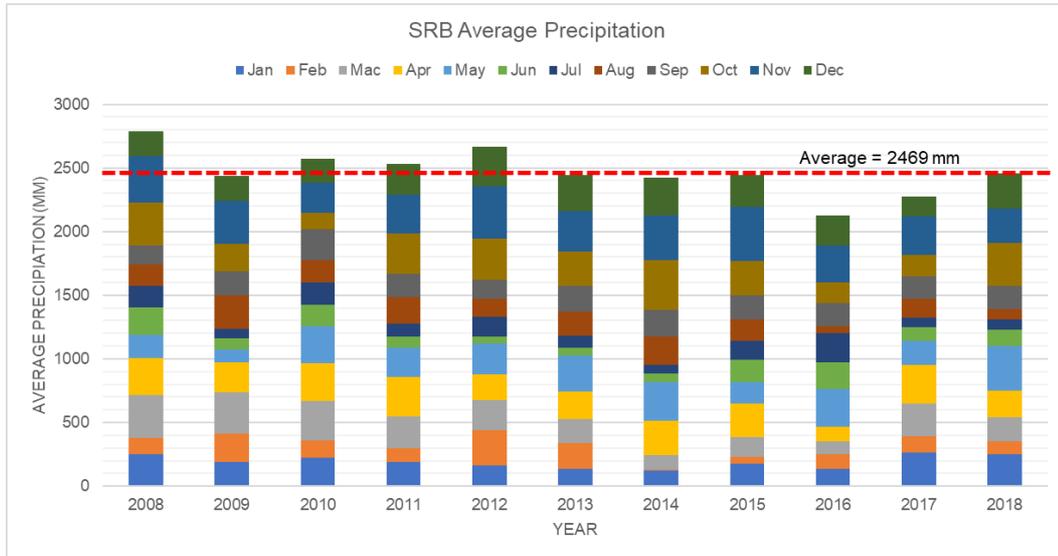


Figure 3-15: SRB annual average rainfall from 2008-2018

### 3.8 Meteorological data

Datasets for meteorological variables including temperature, relative humidity, sun radiation, and wind speed were obtained from MET Malaysia within the same time frames as datasets for hydrological variables. The weather conditions of the SRB were best represented by the meteorological stations from Tennamaran Estate and Pusat Pertanian Batang Kali, according to the location map in Figure 3-16. However, the Tennamaran Estate meteorological station began to operate in 2012 with limited parameters, while the Pusat Pertanian Batang Kali meteorological station only recorded the rainfall parameter. Therefore, the meteorological data provided by these observation stations is limited. The average maximum and minimum temperatures recorded from Tennamaran Estate station from 2012-2019 are 33.1°C and 24.4° as depicted in Figure 3-17.

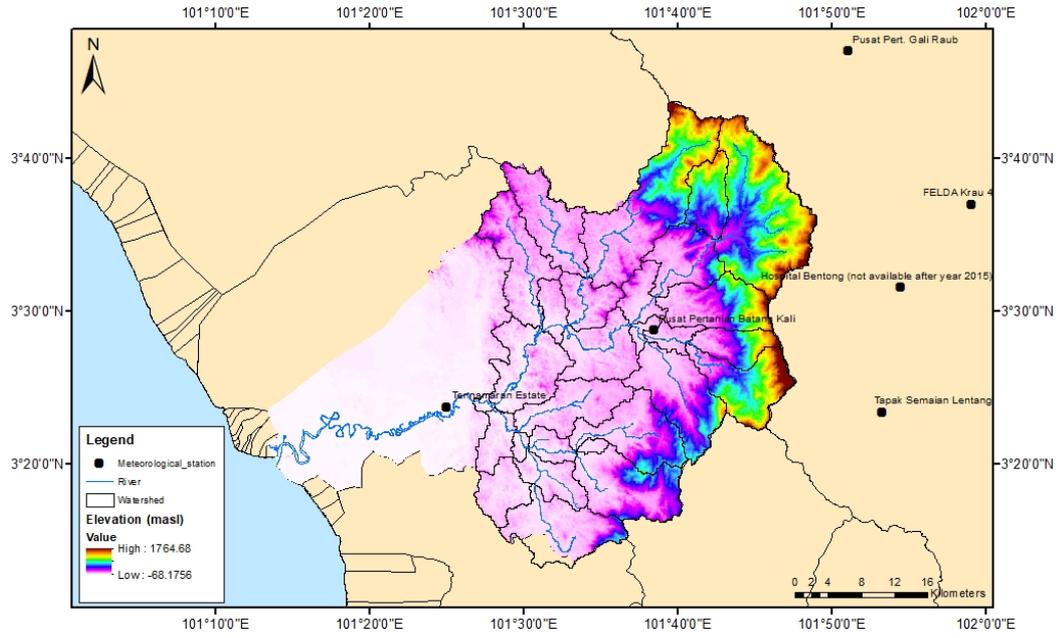


Figure 3-16: Meteorological stations in SRB

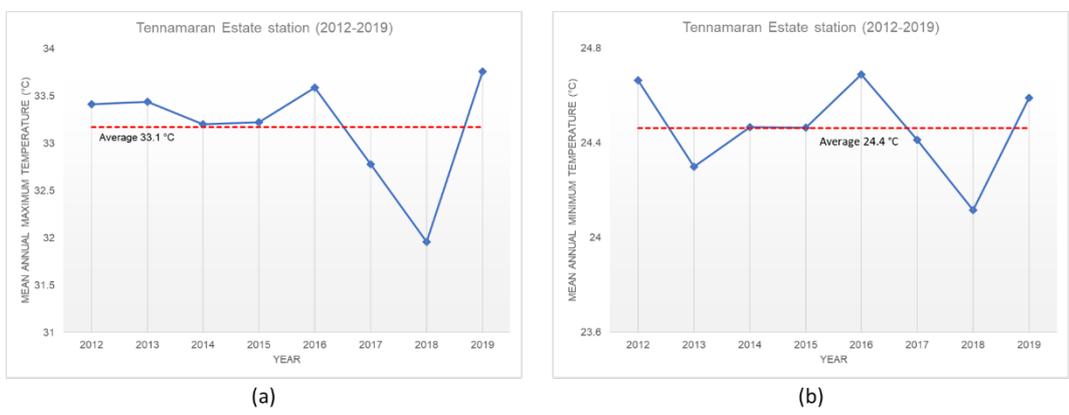


Figure 3-17: Mean annual (a) maximum and (b) minimum temperature at the Tendam Estate meteorological station from 2012-2019

### **3.9 Global satellite/reanalysis data**

Global satellite datasets are becoming more popular among researchers these days since they serve as a substantial supplement to the paucity of gauged data (Worqlul et al., 2014). The best option for displaying meteorological data, according to previous research by Zhang et. al., (2013), is to employ global reanalysis datasets. As a result, this study used two global reanalysis datasets, the Climate Forecast System Reanalysis (CFSR) and China Meteorological Assimilation Driving Datasets (CMADS), as input data for a hydrological model because there has not been much research on the application of these data in Southeast Asia, particularly in Malaysia. These datasets were selected based on (i) the availability of the data as open-source data; (ii) the performances from previous studies; and (iii) their suitability in terms of time and coverage within study area.

#### **3.9.1 Climate Forecast System Reanalysis (CFSR)**

The Climate Forecast System Reanalysis (CFSR) is a third-generation reanalysis product that has been generated since the first introduction of the system in August of 2004 at the National Centers for Environmental Prediction (NCEP) (Saha et al., 2006). CFSR includes three novel features: (1) the first reanalysis technique where the guess fields are taken as a 6-h forecast (analysis hours = 0000, 0600, 1200, and 1800 UT) from an atmosphere-ocean climate system with an interactive portion of the sea ice; (2) a higher horizontal atmospheric resolution at  $\sim 0.31^\circ$  (38 km) than previous atmospheric reanalyses; and (3) the assimilates satellite radiances rather than the temperature and humidity values reanalysed (W. Wang et al., 2011). This includes historically predicted precipitation and temperatures per hour and offers a real-time hydrological forecast of  $0.25^\circ$  (3 km) for any place in the world with the global ocean at the equator, reaching beyond the tropics to a global  $0.5^\circ$ , with 40 levels (Fuka et al., 2014). Precipitation, temperature, wind speed, relative humidity, and solar radiation are the five climatic parameters that are accessible and can be obtained from <https://ncar.ucar.edu/>. Latitude  $0.7^\circ$  N –  $6.8^\circ$  N and longitude  $98.7^\circ$  E –  $105.2^\circ$  E were used to collect datasets for the whole of Peninsular Malaysia.

### **3.9.2 China Meteorological Assimilation Driving Datasets for SWAT Model (CMADS)**

Dr. Xianyong Meng of the China Agricultural University (CAU) created a new collection of publicly available reanalysis climate datasets called the China Meteorological Assimilation SWAT model driving datasets (CMADS) (Xianyong Meng et al., 2018). It was developed using a variety of technological and scientific techniques, including data nesting, resampling, and bilinear interpolation, and was designed specifically for the East Asia region by combining Local Analysis and Prediction System/Space-Time Multiscale Analysis System (LAPS / STMAS) (Xianyong Meng et al., 2018). CMADS assimilates hourly precipitation products from 40,000 regional automated stations and 2,421 national automatic stations in China utilising Climate Prediction Center Morphing Method (CMORPH) satellite product as background (Guo et al., 2018).

CMADS provides daily climate spatial data of 0.25° from 2008 to 2016 between latitudes of 0° N - 65° N and longitudes of 60° E - 160° E consisting of 400 x 260 grid points (total 10,400) in order to provide high-quality and high-resolution meteorological data for the investigation of hydrology and meteorology in the East Asian region (Q. Wang et al., 2020). This dataset includes many variables in comparison to TMPA or PERSIANN, such as the average daily temperature (°C), the maximum daily temperature (°C), the minimum daily temperature (°C), the 24-hour daily cumulative precipitation, the average daily relative humidity (%), the average specific daily humidity (g/kg), average daily solar radiation (MJ/m<sup>2</sup>), the average daily wind speed (m/s), the average daily atmospheric pressure (hPa), the soil temperature (K) and the soil moisture (mm<sup>3</sup>/mm<sup>3</sup>) (Xianyong Meng & Wang, 2017). Since both water resources and non-point source pollution can be recognised with more accuracy, CMADS can significantly reduce the uncertainty of the meteorological input and improve the performance of modelling non-point source pollution (Xianyong Meng & Wang, 2017). This dataset can be downloaded at <http://www.cmads.org/>.

### 3.10 Future climate data

General Circulation Models (GCMs) are the most sophisticated instruments for illustrating physical processes in the cryosphere, atmosphere, ocean, and land surface (Mahmood, 2013). In combination with layered regional models, GCMs may provide estimates of regional climate models that are physically and spatially consistent (IPCC, 2022a). This research employed an ensemble technique of different future predictions from various GCMs based on diverse scenarios to address the uncertainty in climate forecasts as outlined by (Hawkins & Sutton, 2009). Three distinct GCMs (ECHAM5, CCSM3, and MRI), which are shown in Table 3-5, and four future greenhouse emission scenarios (B1, A1FI, A1B, and A2) from the Special Report on Emission Scenarios (Nakicenovic et al., 2000) were used to account for model uncertainty in climate change simulations.

Table 3-5: Special Report on Emission Scenarios (SRES) of 15 different future climate realisations. (source: (IPCC, 2014b))

| SRES  | Scenario  | Realization | Model  |
|-------|---|-------------|--|
| A1B   | The most plausible scenario, describes a future world of rapid economic growth, a global population that peaks in mid-century and declines thereafter, and increase cultural and social interactions.<br><br>The technological emphasis of this scenario is on a balance across all energy sources, not relying too heavily on any particular energy source.  | ccsm3a1b    | CCSM3: The Community Climate System Model (CCSM) is a coupled model for simulating past, present and future climates. The Community Climate System Model version 3 (CCSM3) is a coupled climate model with components representing the atmosphere, ocean, sea ice, and land surface connected by a flux coupler.<br><br>ECHM5: The atmospheric GCM ECHM5 has a spectral dynamical core where vorticity, divergence, temperature and surface pressure are represented in the horizontal by a truncated series of harmonics. A semi-Lanrangian scheme is used for the transport of water components (water vapour, cloud liquid water and cloud ice).<br><br>MRI: The AGCM component of MRI-CGCM2.3.2 is based on a version of the operational weather forecasting model of the Japan Meteorological Agency (JMA). It has a spectral dynamical core where vorticity, divergence, temperature, specific humidity and surface pressure represented in the horizontal by a truncated series of spherical harmonics. |
|       |   | echam5a1b1  |  |
|       |   | echam5a1b2  |  |
|       |   | echam5a1b3  |  |
|       |   | mria1b      |  |
| A2    | Very heterogeneous world. The underlying theme is self-reliance and preservation of local identities.<br><br>Fertility patterns across regions of the world converge very slowly, resulting in continuously increasing world population. Economic development is primary regionally oriented and slower than in other scenarios.  | ccsm3a2     |  |
|       |   | echam5a2_1  |  |
|       |   | echam5a2_2  |  |
|       |   | echam5a2_3  |  |
| B1    | A convergent world with the global population peaking around mid-century and declining thereafter.<br><br>Rapid changes in economic structures toward a service and information economy, with reductions in materials intensity, and the introduction of clean and resource-efficient technologies.   | ccsm3b1     |  |
|       |   | echam5b1_1  |  |
|       |   | echam5b1_2  |  |
|       |   | echam5b1_3  |  |
| mrib1 |   |             |  |
| A1F1  | A future world of very rapid economy growth, low population growth and rapid introduction of new and more efficient technology.<br><br>Major underlying themes are economic and cultural convergence and capacity building, with a substantial reduction in regional differences in per capita income.<br><br>People pursue personal wealth rather than environmental quality. The worst case scenario among all scenarios. | ccsm3a1fi   |  |

### 3.11 Landscape features

In order to set up a hydrological model, several landscape features such as land use, soil type, and elevation are required. Local and global landscape feature data were obtained from various sources, and for some of them, two sources were combined.

#### 3.11.1 Land use

This study used the 2013 land use dataset obtained from PLAN Malaysia, the government agency that oversaw the land use information in Malaysia. The four distinct land use zones in SRB included forest, agriculture, built-up areas, and water bodies comprising rivers, lakes, and ponds. As shown in Figure 3-18, natural forest covers 57% of the SRB region, with agriculture activities accounting for 22%, built-up areas accounting for 17%, and water areas accounting for 4% (LUAS, 2015). Table 3-6 identified the 22 categories of land use code based on the SWAT database that were specifically designated in the SRB area.

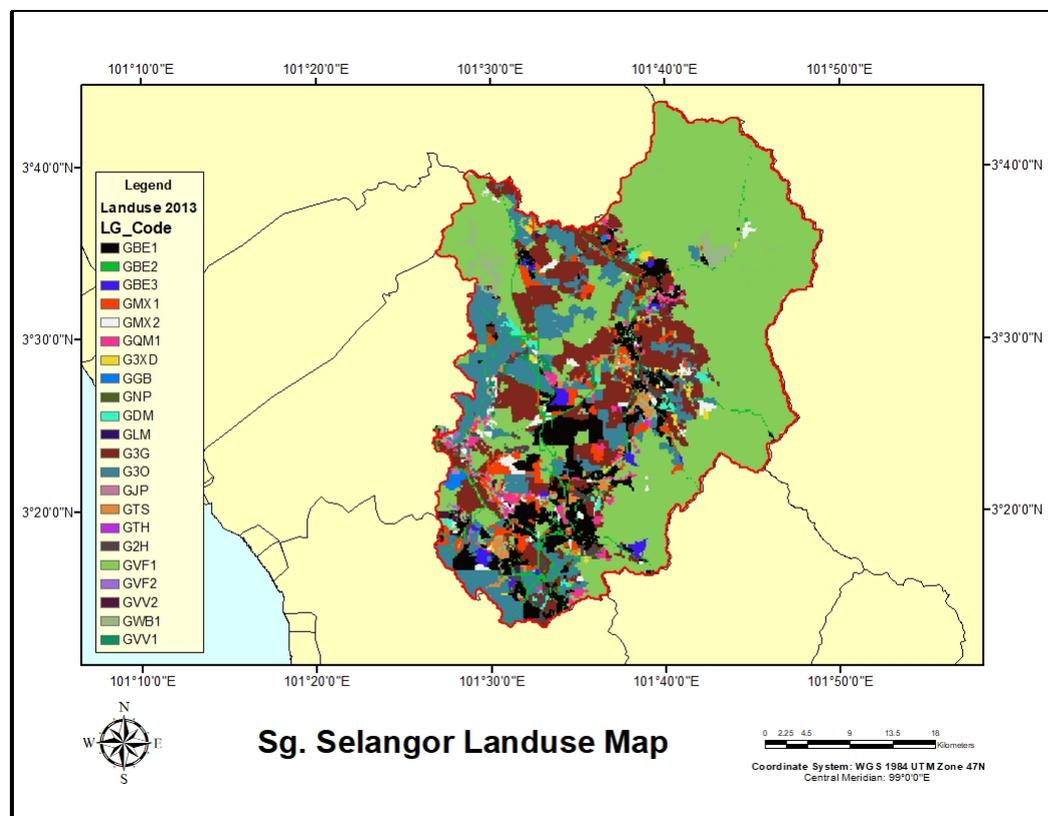


Figure 3-18: Land use (2013) in the Sg. Selangor watershed

Table 3-6: Sg. Selangor land use information

| LG_Code      | Type                               | Area (sqkm) | LG_Code | Type                               | Area (sqkm)    |
|--------------|------------------------------------|-------------|---------|------------------------------------|----------------|
| G2H          | Agricultural Land-Generic          | 21.71       | GMX1    | Range-Brush                        | 42.16          |
| G3G          | Rubber Trees                       | 198.21      | GMX2    | Range-Grasses                      | 21.25          |
| G3O          | Oil Palm                           | 164.72      | GNP     | Orchard - Jackfruit                | 0.38           |
| G3XD         | Orchard - Durian                   | 7.21        | GQM1    | Industrial                         | 20.78          |
| GBE1         | Residential                        | 117.98      | GTH     | Agricultural Land-Generic (Flower) | 0.78           |
| GBE2         | Transportation                     | 38.13       | GTS     | Agricultural Land-Generic          | 14.23          |
| GBE3         | Residential-Low Density            | 12.62       | GVF1    | Forest-Evergreen                   | 713.55         |
| GDM          | Orchard - Mix                      | 10.90       | GVF2    | Wetlands-Forested                  | 1.84           |
| GGB          | Agricultural Land-Row Crops (Mix)  | 3.47        | GVV1    | Water                              | 6.49           |
| GJP          | Bananas                            | 0.64        | GVV2    | Pasture                            | 3.97           |
| GLM          | Agricultural Land-Row Crops (Lime) | 0.15        | GWB1    | Water - Lake                       | 37.29          |
| <b>TOTAL</b> |                                    |             |         |                                    | <b>1438.46</b> |

### 3.11.2 Soil Type

The soil dataset was acquired from the Ministry of Agriculture, Malaysia (MOA) and the information was listed in Table 3-7. There are ten types of soil, within SRB, with Steepland accounting for 47.8% of the total, followed by Serdang-Bungor-Munchong for 26%, and water bodies accounting for only 0.05 percent. Steepland covered almost 48 percent of the SRB soil, followed by Serdang-Bungor-Munchong with 26 percent. The lowest is the urban land soil type, which accounts for less than 1% of the area in this catchment, as shown in Figure 3-19. This research merged the soil characteristics based on commonalities from both local and worldwide soil data. This was done since local soil factors have constraints that prevent them from being used alone. In this particular instance, the Harmonized World Soil Database (HWSD) maintained by the Food and Agricultural Organization (FAO) of the United Nations was used alongside the data collected locally on the soil. Downloadable versions of the soil maps and databases may be found at the following web address: <https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/>.

Table 3-7: Sg. Selangor soil series area

| Soil Type                    | Area (ha)       | %Watershed |
|------------------------------|-----------------|------------|
| Steepland                    | 69461.43        | 47.75      |
| Telemong-Akob-Local Alluvium | 6968.9          | 4.79       |
| Mined Land                   | 13667.68        | 9.4        |
| Munchong-Seremban            | 2284.75         | 1.57       |
| Serdang-Kedah                | 9786.26         | 6.73       |
| Rengam-Jerangau              | 2467.26         | 1.7        |
| Serdang-Bungor-Munchong      | 38098.14        | 26.19      |
| Peat                         | 1474.64         | 1.01       |
| URBAN LAND                   | 1191.54         | 0.82       |
| WATER-6997                   | 74.08           | 0.05       |
| <b>TOTAL</b>                 | <b>145474.7</b> | <b>100</b> |

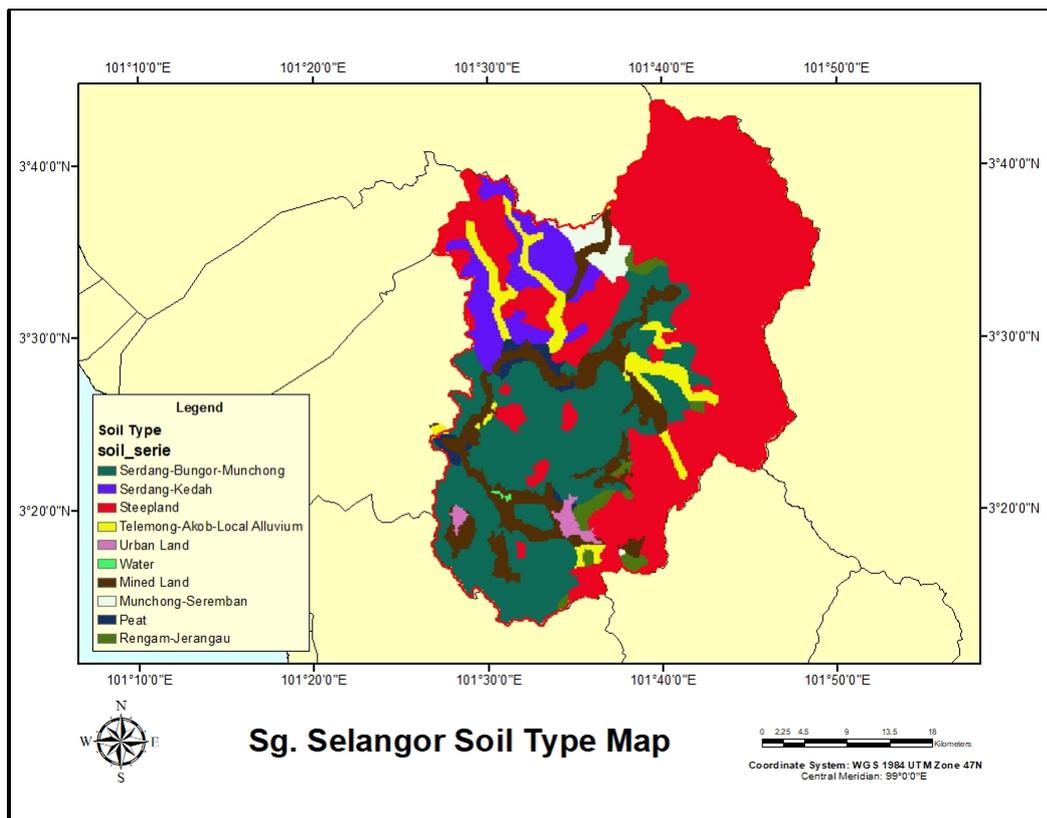


Figure 3-19: Local soil series in the Sg. Selangor watershed

### 3.11.3 Digital Elevation Model (DEM)

DEM is one of the important datasets required for setting up a hydrological model. This study used the Multi-Error-Removed Improved-Terrain DEM (MERIT DEM) dataset as a basis in the hydrological model to generate the SRB watershed. MERIT-DEM is the most recent version of DEM, and it was produced from the previously available spaceborne DEMs (SRTM3 v2.1 and AW3D-30 m v1) by deleting many error components (absolute bias, stripe noise, speckle noise, and tree height bias). It represents topography elevations with a resolution of 3 seconds (90 m at the equator), and it contains land regions between 90N-60S, which are referred to as the EGM96 geoid. This dataset may be obtained by downloading it from the following address: . Figure 3-20 shows the SRB elevation from mean annual sea level (masl), with the highest elevation of SRB at 1,906.53 km and the lowest at -76.79 km [http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT\\_DEM/](http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_DEM/). Figure 3-20 shows the SRB elevation from mean annual sea level (masl), with the highest elevation of SRB at 1,906.53 km and the lowest at -76.79 km.

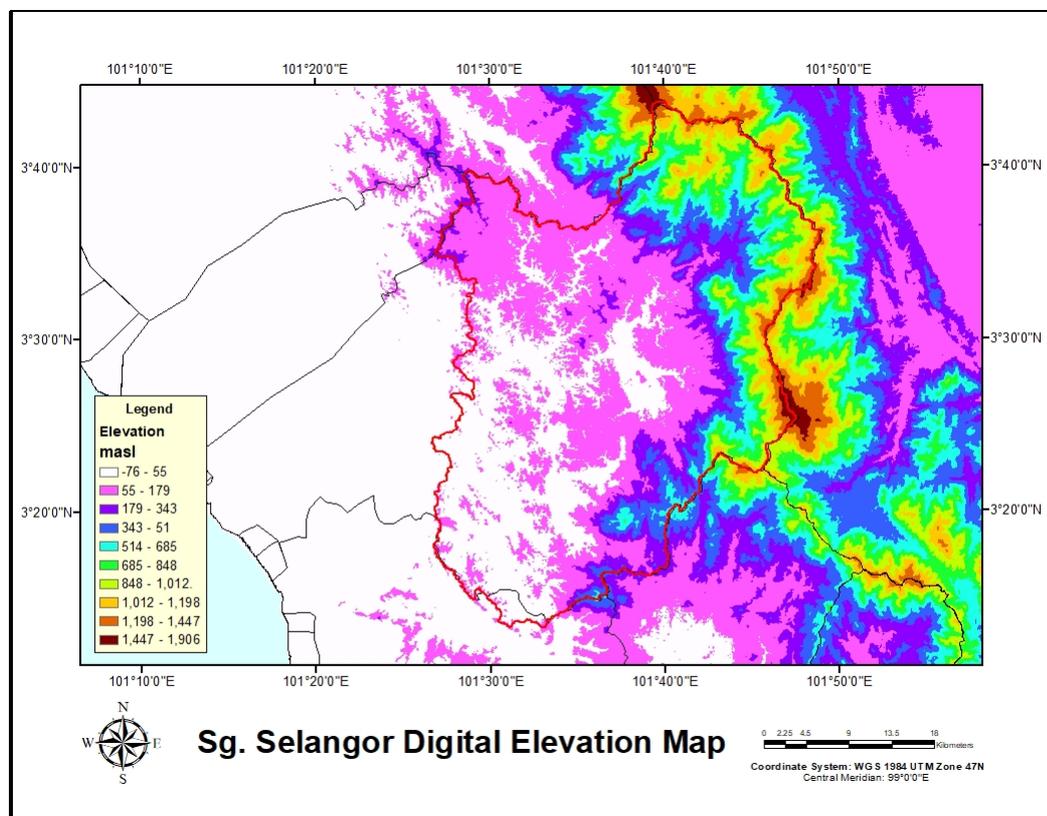


Figure 3-20: Sg. Selangor elevation by MERIT DEM

### 3.12 Conclusion

- This study selected the SRB as a research area because it is an incredibly significant watershed that supplies 67% of the water resources for the state of Selangor and Kuala Lumpur Federal Territory.
- This study focused on the effective and semi-effective zones, which had a catchment area of about 1,454.74 km<sup>2</sup>.
- Hydro-meteorological data from various sources were obtained for the hydrological model's input data.
- In this work, an ensemble technique based on several future estimates from different GCMs and based on multiple scenarios was utilised to analyse climate change.

## **CHAPTER 4**

### **METHODOLOGY**

#### **4.1 Overall methodology**

This research was carried out in the Selangor River Basin (SRB), which is situated in the upper part of the Selangor state and is discussed in Chapter 3. SRB was selected as a study area due to its significant potential for future development, particularly in terms of water resources (DID, 2011c). Through the Sungai Selangor Water Supply Scheme, SRB supplies about 67% (3,018 MLD) of the total water supply requirement for Kuala Lumpur and Selangor (LUAS, 2015). The sequence procedure is depicted in Figure 4-1, as it achieved all objectives in this study. Details of the methodology for each specific objective are explained in each chapter separately.

Hydro-meteorological and landscape features, including DEM (Digital Elevation Model), soil type, and land use, were collected from various sources, as described in Chapter 3. In the first objective described in Chapter 1, the performance of global satellite/reanalysis datasets was evaluated at an observed station for the Selangor River Basin (SRB). The hydrological model also used the cross-combined datasets as input data in addition to the individual data. This approach is designed to achieve the second goal of this research as outlined in Chapter 5. The third objective consisted of evaluating the performance of the multi-site calibration (MSC) method and the techniques used in MSC were analysed for objective number four outlined in Chapter 6. The optimum hydrological model was employed to estimate future runoff in the SRB using validated future climate datasets in the last objective described in Chapter 7.

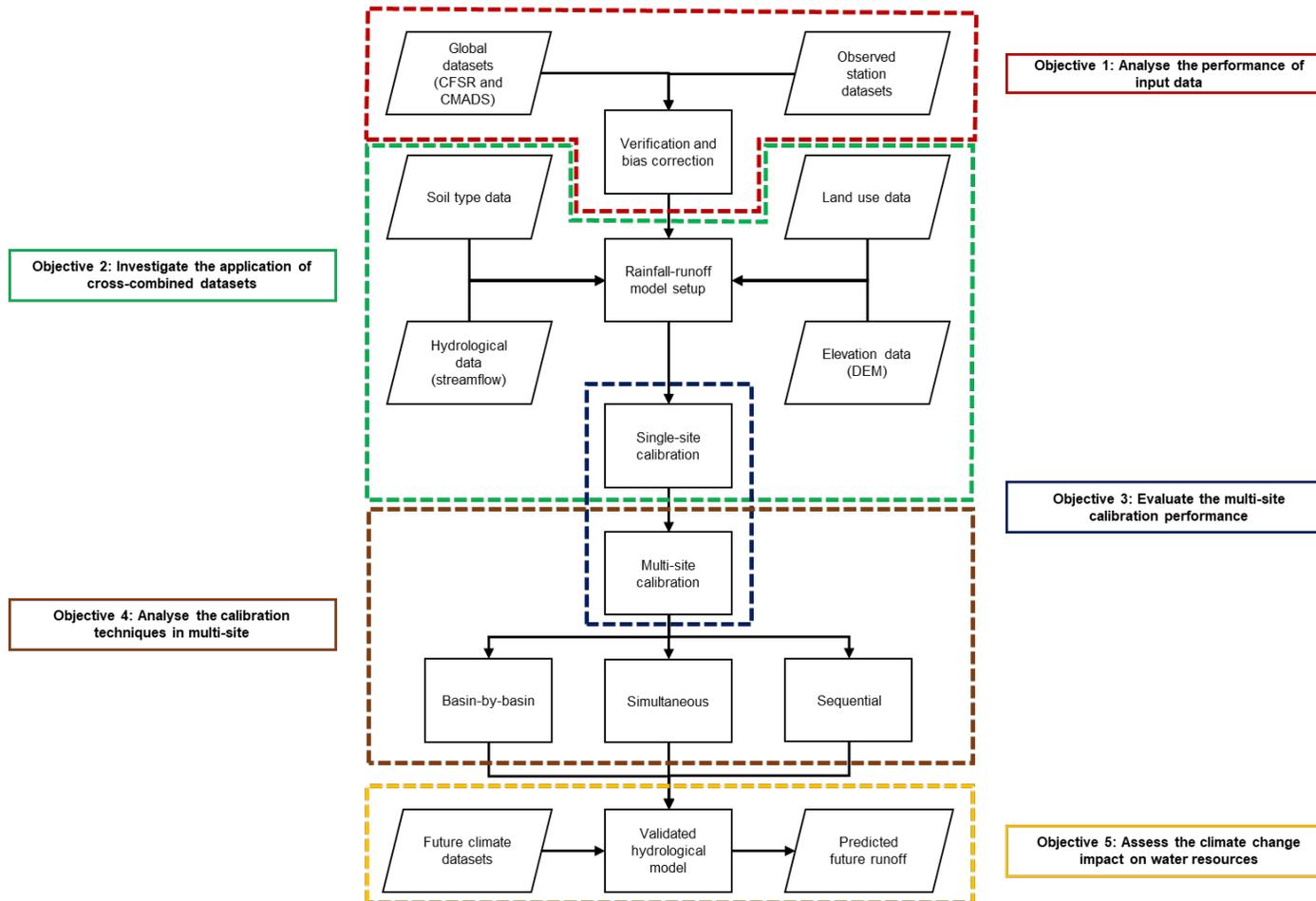


Figure 4-1: Flowchart of an overall procedure

## **4.2 Analysis input data performance**

The accuracy of a model that predicts rainfall-runoff is contingent on the precision of the input data. A major problem that most hydrology and water resources management encounter is the ability to predict hydrologic reactions without the proper input data (Farmer, 2016). Observation station data is typically used as the basis for input data for hydro-climate models. However, the global satellite/reanalysis dataset, is the best alternative to be used as an input dataset because of the sparseness and unevenness of the observation station in a particular location (Worqlul et al., 2014). Furthermore, the accuracy and consistency of the input data were verified before applying it to a rainfall-runoff model. It would be advantageous to examine the temporal and geographical characteristics of these supplemental data in order to comprehend their differences, similarities, and validity (Nasrabadi et al., 2013). Details of this analysis process are described in each sub-topic.

### **4.2.1 Data selection**

#### *Hydro-meteorological data*

The observation data for this study were obtained from the Department of Irrigation and Drainage (DID) Malaysia, and the Malaysian Meteorological Department (Met Malaysia), respectively, via their official websites at <http://h2o.water.gov.my/> and <https://m.met.gov.my/>. Daily data from 2008 to 2016 was retrieved from both sources based on their availability. This involved thirty rainfall stations, three streamflow stations, and two meteorological stations, as illustrated in Figure 4-2. Information on these hydro-meteorological stations is described in Chapter 3.

#### *Global satellite/reanalysis data*

Two global reanalysis datasets, the Climate Forecast System Reanalysis (CFSR) and China Meteorological Assimilation Driving Datasets (CMADS) were downloaded at <https://ncar.ucar.edu/> and <http://www.cmads.org/>. From 2008 to 2016, five meteorological metrics, including precipitation, temperature, sun radiation, relative humidity, and wind speed, were obtained.

As seen in Figure 4-2, these datasets were extracted into a grid-based dataset.

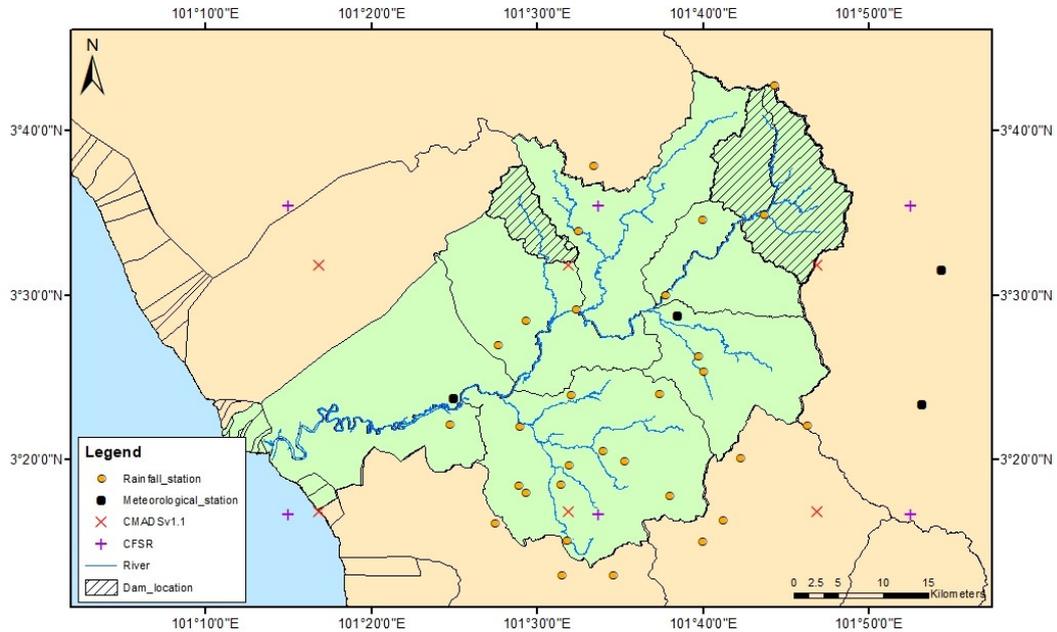


Figure 4-2: Location map of hydro-meteorological stations and gridded point of global reanalysis datasets in Selangor River Basin

#### 4.2.2 Evaluation methods

A time series of hydrological data may exhibit spikes or jumps and missing during the recording time. These usually arise when using different instruments and methods of observation. Even natural and man-made conditions can cause these errors. Therefore, at this stage, observed data was examined, verified, and continued with an error correction, if possible, before being included in a hydrological model. To examine the consistency of the observed data over time and to identify patterns through changes in slope, the double mass curve method was employed (Albert, 2004). Global reanalysis datasets were retrieved in a form of gridded data. To analyse these datasets, point-to-point analysis was used by comparing these datasets to observed data at gauge points. Monthly data from 2008 to 2016 was used to evaluate the global reanalysis datasets.

### *Double mass curve analysis*

The Double Mass Curve (DMC) method was used in order to validate the precision and reliability of the observed data, in particular the data pertaining to rainfall and streamflow. According to Searcy and Hardison (1960), a lengthy record is more likely to contain inconsistent data because of changes to the data collection method or to the physical environment, such as a catchment region. To check the consistency of the precipitation records, the steps below are as follows:

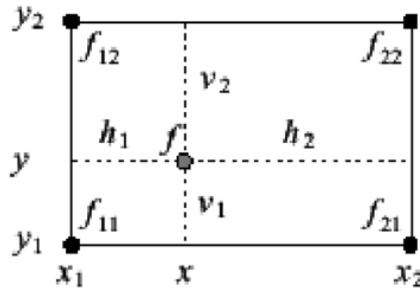
1. Annual precipitation data are tabulated and accumulated in chronological order for each year.
2. The average of the cumulative precipitation is calculated, and the pattern is used to test individual station records.
3. The graph was plotted between the cumulative precipitation for each station against the cumulative precipitation of the pattern.

The double mass curve with an unbroken straight line indicates that the record is consistent. While the broken straight line does not necessarily indicate inconsistent records, it needs further examination. In some cases, this happens due to changes in gauge location (Searcy & Hardison, 1960).

These methods also apply to streamflow analysis; however, the data must first be transformed to a similar basis, such as runoff in inches, cubic feet per second per square mile, or percentage of mean flow. Otherwise, a huge stream flow in a group would have a greater impact on the pattern than numerous little streams. Details on this method can be found in (Searcy & Hardison, 1960).

### *Point-to-point analysis*

Gridded datasets of CFSR and CMADS were extracted at a point location using the bilinear interpolation equation. In this study, the point location is based on the observation stations determined in Chapter 2, as illustrated in Figure 4-2. The calculation for the bilinear interpolation approach is as follows:



$$f(x_1, y) = \frac{v_1 f(x_1, y_2) + v_2 f(x_1, y_1)}{v_1 + v_2} \quad \text{Equation 1}$$

$$f(x_2, y) = \frac{v_1 f(x_2, y_2) + v_2 f(x_2, y_1)}{v_1 + v_2} \quad \text{Equation 2}$$

$$f(x, y) = \frac{h_1 f(x_2, y) + h_2 f(x_1, y)}{h_1 + h_2}$$

$$= \frac{h_1 v_1 f_{22} + h_1 v_2 f_{21} + h_2 v_1 f_{12} + h_2 v_2 f_{11}}{(h_1 + h_2)(v_1 + v_2)} \quad \text{Equation 3}$$

In bilinear interpolation method,  $f_{11}$ ,  $f_{12}$ ,  $f_{21}$  and  $f_{22}$  represent gridded data from CFSR or CMADS. While  $f$  represents a station to which the climate data is interpolated. Example calculation using bilinear interpolation method is described below.

### Example

*Location and value of CFSR gridded data*

$$(x_1, y_1) = (101.56, 2.96); f(x_1, y_1) = 1.99 \text{ mm}$$

$$(x_1, y_2) = (101.56, 3.27); f(x_1, y_2) = 1.25 \text{ mm}$$

$$(x_2, y_1) = (101.87, 2.96); f(x_2, y_1) = 4.85 \text{ mm}$$

$$(x_2, y_2) = (101.87, 3.27); f(x_2, y_2) = 3.48 \text{ mm}$$

Point location  $(x,y) = (101.57,3.21)$ .

From equation 1:

$$f(x_1,y) = [(3.21-2.96)(1.25) + (3.27-3.21)(1.99)] / (1.25 + 1.99) = 0.133$$

From equation 2

$$f(x_2,y) = [(3.21-2.96)(3.48) + (3.27-3.21)(4.85)] / (3.48 + 4.85) = 0.139$$

From equation 3

$$f(x,y) = [(101.57-101.56)(3.21-2.96)(3.48) + (101.57-101.56)(3.27-3.21)(4.85) + (101.87-101.57)(3.21-2.96)(1.25) + (101.87-101.57)(3.27-3.21)(1.99)] / [((101.57-101.56)+(101.87-101.57)) \times ((3.21-2.96)+(3.27-3.21))] = \mathbf{1.47 \text{ mm}}$$

### 4.2.3 Verification method

The coefficient of determination ( $R^2$ ) was employed to confirm the performance data in the double mass curve approach. This coefficient measures how strongly two variables are linearly correlated. The more closely an estimate agrees with observable data, the closer an  $R^2$  value is to 1.0 (Santhi et al., 2002). The coefficient of determination ( $R^2$ ) equation is given by:

$$R^2 = \left[ \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \right]^2 \quad \text{Equation 4}$$

Where,  $n$  is the number in the observations in the dataset,  $\sum x$  is the sum of the first variable,  $\sum y$  is the sum of the second variable,  $\sum xy$  is the sum of the product of the first and second variable,  $\sum x^2$  is the sum of the squares of the first variable, and  $\sum y^2$  is the sum of the squares of the second variable. In this analysis, the first variable is referred to as the cumulative precipitation/streamflow for the pattern and the second variable is referred to as the cumulative precipitation/streamflow for individual stations or streams.

The correlation coefficient ( $r$ ), mean absolute error (MAE), and root-mean square error (RMSE) are typical statistical verification methods that have been used to gridded datasets. These are the common verification methods that compare the gridded datasets with the observed datasets, as described by Ebert et. al. (2007) and Ghulami (2018). Using the continuous verification statistics, the following was done to figure out how accurate a continuous variable was:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad \text{Equation 5}$$

Where,  $r$  is the correlation coefficient,  $n$  is the number of observations,  $x$  is the first variable in the context, and  $y$  is the second variable. In this statistic, the  $r$  value indicates of how well the points fit a straight line. According to Taylor (1939), numbers between -1 and +1, represent the greatest conceivable agreement, while 0 represents the highest possible disagreement.

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x| \quad \text{Equation 6}$$

In MAE,  $n$  is the number of errors,  $x_i$  is the measurement value, and  $x$  is the observed value. The range of MAE is 0 to  $\infty$  and to interpret it, the closer value of MAE to 0 is the better. However, the interpretation of MAE depends on the range of the values and the acceptability of error (Datagy, 2022).

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(x_i - y_i)^2}{n}} \quad \text{Equation 7}$$

In RMSE,  $x_i$  is the observed values, and  $y_i$  is the modelled values at time  $i$ , while  $n$  is the number of samples. In this statistic, the lowest values of RMSE indicate the better fit between the modelled and observed values.

### 4.3 Investigate of cross-combine dataset

In this sub-topic, the development and assessment of the cross-combined datasets in the hydrological model were covered.

#### 4.3.1 Data preparation

Observed data from gauge stations are often utilised as inputs for hydrological models. Due to the sparseness and inconsistency of gauge stations, however, global satellite datasets have become an alternate source to supplement the absence of observable data (Worqlul et al., 2014). Instead of using a single dataset as input data to the hydrological model, this study employed cross-combined datasets from gauged stations and global satellite datasets to enhance model performance. A cross-combined technique included the use of two elements (precipitation and temperature) from the gauged stations and five elements (precipitation, temperature, solar radiation, relative humidity, and wind speed) from global satellite datasets. Table 4-1 lists the seven scenarios that were created using the three sources' datasets (observed station, CFSR, and CMADS). Scenario 1 to scenario 3 represents the individual data from each source. Scenario 4 to scenario 7 are the combinations of observed and global satellite data. All datasets were produced in Microsoft Access database format using the Microsoft Excel macro-enabled provided by <https://swat.tamu.edu/> for use with the SWAT rainfall-runoff model, as shown in Figure 4-3.

Table 4-1: Cross-combine input data

| Type   | Local gauge | CMADS          | CFSR           |
|--------|-------------|----------------|----------------|
| Sc - 1 | P & T       |                |                |
| Sc - 2 |             | P, T, R, S & W |                |
| Sc - 3 |             |                | P, T, R, S & W |
| Sc - 4 | P & T       | R, S & W       |                |
| Sc - 5 | P           | T, R, S & W    |                |
| Sc - 6 | P & T       |                | R, S & W       |
| Sc - 7 | P           |                | T, R, S & W    |

Sc – Scenario; P – Precipitation; T – Temperature; R – Relative humidity; S – Solar radiation; W – Wind.

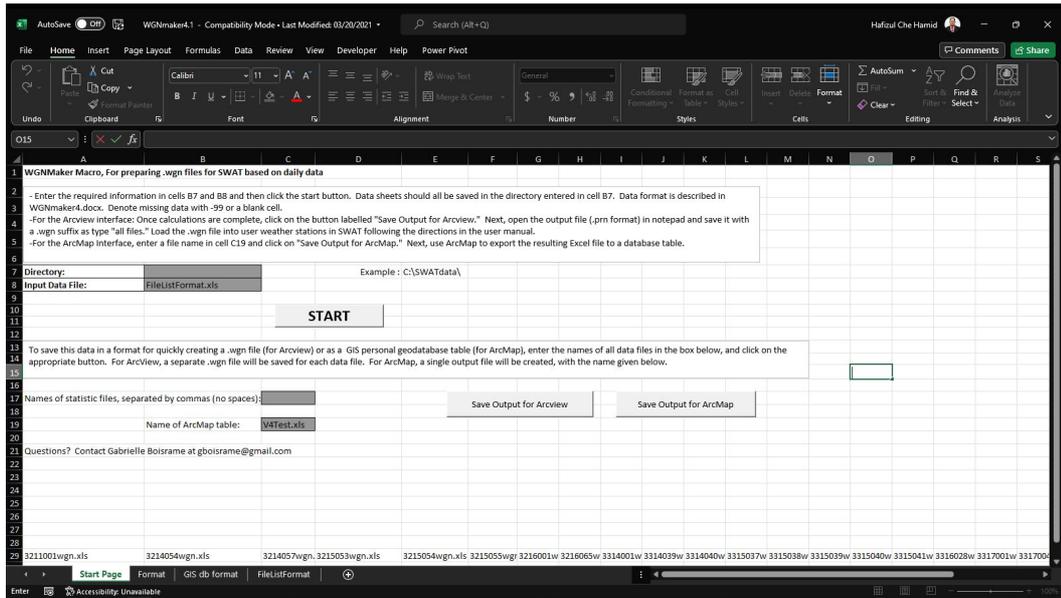


Figure 4-3: Microsoft excel macro-enable platform to generate the input database in Microsoft Access database.

#### 4.3.2 Hydrological model setup

The Soil Water Assessment Tool, often known as SWAT, was employed in this study as a tool for creating a rainfall-runoff model. SWAT is a semi-distributed, continuous-time, process-based river basin model that is connected to ArcGIS/QGIS to examine the effects of various management strategies on water resources and nonpoint-source pollution in substantial river basins (Arnold et al. 2012). SWAT has been widely used among researchers and is the most popular modelling tool compared to other models (Mannschatz et al., 2016). According to Ang and Oeurng (2018), SWAT has the potential to help increase the accuracy of the findings of streamflow simulations that are based on rainfall and the physical parameters of the basin.

## Soil Water Assessment Tool

SWAT is a semi-distributed hydrological model created by the Agricultural Research Service (ARS) of the USDA (United States Department of Agriculture) during the last three decades. It is an operational and conceptual model that may be utilised in a broad variety of environmental settings (Arnold & Fohrer, 2005) and is commonly used for analysing hydrologic consequences (Dlamini et. al., 2017). Additionally, SWAT has acquired recognition as an efficient instrument and earned acceptance on a global scale as a reliable transdisciplinary watershed modelling tool (Gassman, Reyes, Green, Arnold, et al., 2007). SWAT may also model all significant elements (hydrology, sediment, and chemicals) relevant to catchments at the watershed size (Golmohammadi et al., 2014). The water balance equation of soil serves as the foundation for the hydrologic component of SWAT (M. et al., 2003):

$$SW_t = SW_0 + \sum_{i=1}^t (R_{day} - Q_{surf} - ET_a - W_{seep} - Q_{gw}) \quad \text{Equation 8}$$

Where  $SW_t$  is the final soil water content (mm),  $SW_0$  is the initial soil water content (mm),  $t$  is time in days,  $R_{day}$  is the amount of precipitation (mm),  $Q_{surf}$  indicates the amount of surface runoff (mm),  $ET_a$  is the amount of evapotranspiration (mm),  $W_{seep}$  is the amount of water entering the vadose zone from the soil profile (mm), and  $Q_{gw}$  is the amount of return flow (mm). Furthermore, SWAT used the Soil Conservation Service (SCS) curve number method to estimate the surface runoff at each sub-basin. The SCS curve number equation is:

$$Q_{surf} = \frac{(R_{day} - 0.2 S)^2}{(R_{day} + 0.8 S)} \quad \text{Equation 9}$$

Where  $Q_{surf}$  is the accumulated runoff or rainfall excess (mm),  $R_{day}$  is the rainfall depth for the day (mm), and  $S$  is the retention parameter (mm).

In this study, QSWAT3 version 1.1.1 (Integration of QGIS version 3.10.12 and SWAT) was used as the SRB hydrological model. The SWAT model was set up using the basic map inputs, including Digital Elevation Model (DEM), land use, soil type, and climate. Monthly precipitation, maximum and minimum temperature, solar radiation, relative humidity, and wind speed data were used as climatic inputs. Figure 4-4 depicts the flowchart used to develop the SRB SWAT model.

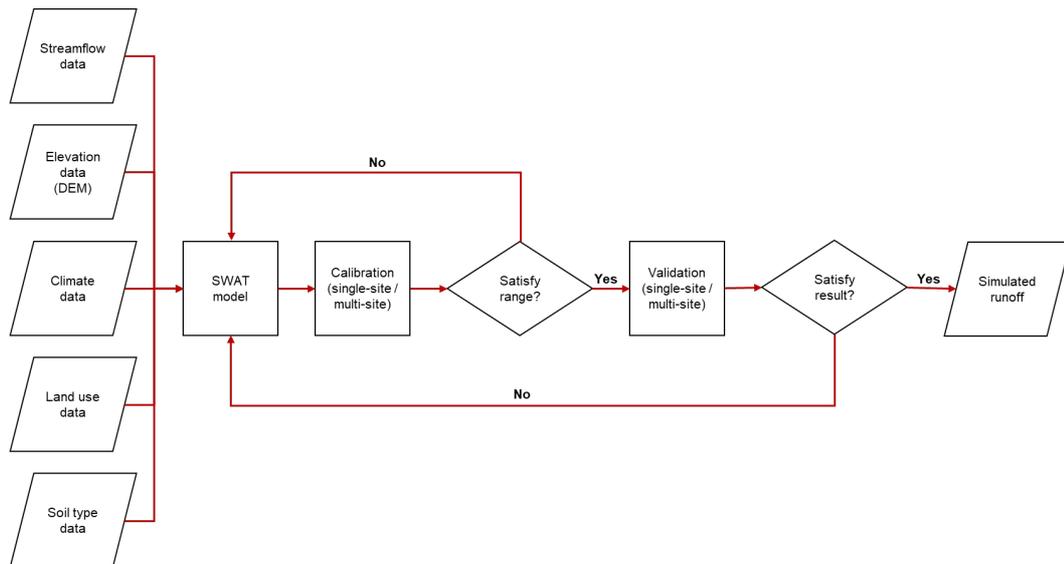


Figure 4-4: Flowchart of SRB SWAT model

### *Watershed Delineation*

In the SWAT model, watersheds need to be discretised as it is a physical based semi-distributed model. The delineation process starts with setting up the threshold of the catchment area based on input data from the Digital Elevation Model (DEM) and SWAT will automatically create a sub-basin. In this study, raster data from MERIT-DEM was used as a base layer in the SWAT model and 30 km<sup>2</sup> was defined as a threshold area to generate sub-basins of SRB.

### *Hydrological Response Units (HRUs)*

The hydrological response units, often known as HRUs, are the smallest spatial units that the model uses to reflect the geographic diversity of the basin. Flugel (1995) employed a Geographical Information System (GIS) to examine the terrain, soils, geology, rainfall, and how the land is managed in order to determine the HRUs. The typical HRUs definition approach includes all comparable land uses, soil types, and slopes in a sub-basin based on user-defined thresholds (Gassman, Reyes, Green, & Arnold, 2007). HRUs were identified in this research using the geographical distribution data on topography (MERIT-DEM), land use, and soil type.

### *Selection of Parameter (Parameterisation)*

Prior to the calibration process, the first step is to identify a relevant parameter significant to the SRB SWAT model. In SWAT, not all parameters are relevant to be used and not all should be used simultaneously in the calibration process (Karim C. Abbaspour et al., 2017). As far as this study is concerned and reviewed in the Web of Science database, there is no previous study specifically using SWAT in SRB. Therefore, a set of parameters were determined by referring to a similar previous study that was conducted in Malaysia and globally. At least 10 prior studies (1: (Brighenti et al., 2019), 2: (Deng et al., 2019), 3: (Odusanya et al., 2019), 4: (Rivas-Tabares et al., 2019), 5: (Pandey et al., 2019), 6: (Mou Leong Tan et al., 2018), 7: (Mou Leong Tan, Ibrahim, et al., 2017), 8: (N. S. Dlamini et al., 2017), 9: (Rusli et al., 2017), 10: (Raffar et al., 2022), as listed in Table 4-2, were used as reference in selecting parameters for the SRB SWAT model.

Table 4-2: SWAT parameters used by other researchers.

| No. | Parameters      | Global |   |   |   |   | Malaysia |   |   |   |    | Description  |
|-----|-----------------|--------|---|---|---|---|----------|---|---|---|----|--|
|     |                 | 1      | 2 | 3 | 4 | 5 | 6        | 7 | 8 | 9 | 10 |  |
| 1   | r_CN2.mgt       | x      | x | x | x | x | x        | x | x | x | x  | Initial SCS runoff curve number for moisture condition II                                    |
| 2   | r_SOL_AWC.sol   |        | x | x | x |   | x        | x |   | x | x  | Available water capacity of the soil layer.  |
| 3   | r_SOL_K.sol     |        | x | x |   | x |          |   |   | x | x  | Saturated hydraulic conductivity.  |
| 4   | r_SOL_BD.sol    | x      | x | x |   |   |          |   | x |   |    | Moist bulk density   |
| 5   | v_GWQMN.gw      | x      |   |   | x |   | x        | x | x | x | x  | Threshold depth of water in the shallow aquifer required for return flow to occur            |
| 6   | v_GW_REVAP.gw   | x      | x |   | x |   | x        | x | x | x | x  | Groundwater "revap" coefficient  |
| 7   | v_REVAMPMN.gw   |        | x |   | x |   | x        | x |   | x | x  | Threshold depth of water in the shallow aquifer for "revap" to occur (mm).                   |
| 8   | v_ESCO.hru      | x      | x | x | x | x | x        | x | x | x | x  | Soil evaporation compensation factor   |
| 9   | r_HRU_SLP.hru   |        |   |   | x |   |          |   |   |   |    | Average slope steepness  |
| 10  | r_OV_N.hru      | x      |   |   | x |   |          |   |   |   |    | Manning's "n" value for overland flow.   |
| 11  | r_SLSUBBSN.hru  | x      |   |   | x |   |          |   |   |   |    | Average slope length.  |
| 12  | v_ALPHA_BF.gw   |        | x | x | x | x | x        | x | x | x | x  | Baseflow alpha factor (days).  |
| 13  | v_GW_DELAY.gw   | x      | x |   | x | x | x        | x | x | x | x  | Groundwater delay time (days).   |
| 14  | v_CH_K2.rte     | x      | x |   | x | x | x        | x | x | x | x  | Effective hydraulic conductivity in main channel alluvium                                    |
| 15  | v_CH_N2.rte     | x      | x |   | x | x | x        | x | x | x | x  | Manning's "n" value for the main channel.  |
| 16  | v_CANMX.hru     |        |   | x | x | x |          |   |   | x |    | Maximum canopy storage.  |
| 17  | v_RCHRG_DP.gw   | x      | x |   | x |   | x        | x |   |   |    | Deep aquifer percolation fraction.   |
| 18  | v_SURLAG.bsn    | x      |   |   | x |   | x        | x |   | x | x  | Surface runoff lag time  |
| 19  | v_EPCO.hru      |        |   | x | x |   |          |   |   | x | x  | Plant uptake compensation factor.  |
| 20  | r_CH_W2.rte     | x      |   |   |   |   |          |   |   |   |    | Average width of main channel  |
| 21  | r_CH_S2.rte     | x      |   |   |   |   |          |   |   |   |    | Average slope of main channel  |
| 22  | r_CH_L2.rte     | x      |   |   |   |   |          |   |   |   |    | Length of main channel   |
| 23  | v_LAT_TTIME.hru | x      |   |   | x | x |          |   |   |   |    | Lateral flow travel time   |
| 24  | r__MSK_CO1.bsn  | x      |   |   |   |   |          |   |   |   |    | Calibration coefficient used to control impact of the storage time constant for normal flow. |
| 25  | r__MSK_CO2.bsn  | x      |   |   |   |   |          |   |   |   |    | Calibration coefficient used to control impact of the storage time constant for low flow.    |
| 26  | v_EVRSV.res     |        |   | x |   |   |          |   |   |   |    | Lake evaporation coefficient   |
| 27  | v_GSI.plant.dat |        |   | x |   |   |          |   |   |   |    | Max stomatal conductance (in drought)  |
| 28  | v_FFCB.bsn      |        |   | x |   |   |          |   |   |   |    | Initial soil water storage expressed as a fraction of field capacity water content.          |
| 29  | SHALLST.gw      |        |   |   | x |   |          |   |   |   |    | Initial depth of water in the shallow aquifer  |
| 30  | SLSOIL.hru      |        |   |   | x |   |          |   |   |   |    | Slope length for lateral subsurface flow   |
| 31  | SOL_Z.sol       |        |   |   | x | x |          |   |   |   | x  | Depth from soil surface to bottom of layer.  |
| 32  | CH_K1.sub       |        |   |   | x |   |          |   |   |   |    | Effective hydraulic conductivity in tributary channel alluvium                               |
| 33  | CH_N1.sub       |        |   |   | x | x |          |   |   |   |    | Manning's "n" value for the tributary  |
| 34  | EVRCH.bsn       |        |   |   | x |   |          |   |   |   |    | Reach evaporation adjustment factor.   |
| 35  | PLAPS.sub       |        |   |   | x |   |          |   |   |   |    | Precipitation lapse rate.  |
| 36  | TLAPS.sub       |        |   |   | x |   |          |   |   |   |    | Temperature lapse rate.  |

### Sensitivity Analysis

Sensitivity analysis is a method to identify the model's key affecting factors, where it offers details on the most important procedure in the study domain, and to get rid of any parameters that are not sensitive (Karim C. Abbaspour et al., 2017). The first step in the SWAT calibration process is to identify the parameters that are most sensitive for a certain watershed or sub-watershed. This phase is essential for identifying significant parameters and the precision needed for calibration (Ma et al., 2000). Local sensitivity (also known as *one-at-a-time*) and global sensitivity (also known as *all-at-a-time*) are the two types of sensitivity analysis that are typically used (Karim C. Abbaspour et al., 2017). Local sensitivity is a straightforward and efficient technique that involves simply modifying one parameter at a time to determine its impact on a particular model result. However, because some of the factors commonly

depend on other values, this technique is less precise. While it is more accurate to determine the impact of each parameter on the objective function for a global sensitivity than a local sensitivity when all the parameters are changing across many runs (typically more than 500 runs). The disadvantage of global sensitivity analysis is that many simulations are required (Karim C. Abbaspour et al., 2017). Both approaches are essential for calibrating the model since they provide insight about the sensitivity of the parameters. Global sensitivity determines the sensitivity of each parameter using a multiple regression model as follows:

$$g = \alpha + \sum_{i=1}^n \beta_i b_i \quad \text{Equation 10}$$

Where  $g$  is the objective function value,  $\alpha$  is the regression constant, and  $\beta$  is the coefficient of parameters. The value of the  $t$ -stat and the  $p$ -value were used to identify which parameter was the most sensitive. A more sensitive parameter is indicated by a bigger absolute value of the  $t$ -stat and a lower value for the  $p$ -value.

In this study, the sensitivity analysis was computed using SWAT-CUP (Calibration and Uncertainty Programs) with 2000 iterations.

### **4.3.3 Model Calibration**

#### **4.3.3.1 SWAT-CUP**

Calibration is a crucial step in establishing a set of parameters for a hydrological model by using observational streamflow data. A successful model relies on how well it has been calibrated, which is a time-consuming and laborious procedure (Yang et al., 2008). The SWAT model's calibration and validation processes were carried out in this work using the SWAT-CUP software, which includes the Sequential Uncertainty Fitting (SUFI-2) approach. SUFI-2 aims to capture the majority of the observed data inside the 95% prediction uncertainties (95PPU) of the model by using an iterative mapping method that represents all uncertainties (parameter, conceptual model, input, etc.) on the parameters (represented as uniform distributions or ranges) (K. C. Abbaspour et al., 2015).

SUFI-2 was chosen for the calibration phase over other algorithms (PSO, MCMC, ParaSol, and GLUE) because it integrates optimization with uncertainty analysis and can handle a lot of parameters (K. C. Abbaspour et al., 2004). SUFI-2 is also quite efficient, particularly for time-consuming large-scale models (K. C. Abbaspour et al., 2015; Yang et al., 2008).

#### 4.3.3.2 Goodness-of-fit and uncertainty.

The uncertainty in the SUFI-2 algorithm's input parameters is described by uniform distributions. The 95 percent prediction uncertainty (95PPU) is then used to gauge the amount of uncertainty in the model's output. It is calculated using the cumulative distribution of output variables produced by Latin hypercube sampling until the 95PPU between the 2.5 and 97.5 percentiles bracket more than 90% and the average distance between the 2.5 and 97.5 percentiles is smaller than the standard deviation of measured data levels (K. C. Abbaspour et al., 2015). The anticipated variables' distribution is not always Gaussian and may be significantly skewed. For each variable  $q$ , the goodness of fit is calculated using the proportion of measured data falling within the 95PPU region (prediction uncertainty), the correlation coefficient  $R^2$  between the optimised and observed data, and the average distances  $\bar{d}$  between the upper and lower 95PPU as calculated from

$$\bar{d} = \frac{1}{K} \sum_{l=1}^K (q_v - q_L) l \quad \text{Equation 11}$$

in which  $l$  is a counter, and  $K$  is the total number of observations for variable  $q$ . The ideal result is when all measurements fall inside the 95PPU,  $R^2$  is close to 1, and  $\bar{d}$  is nearly zero. Unfortunately, in most cases this won't be the case due to measurement errors and model uncertainty. 90% or more of the measured data must fall within the 95PPU,  $\bar{d}$  be smaller than the measured data's standard deviation, and have  $R^2 > 0.8$  for the model to be considered calibrated (K. C. Abbaspour et al., 2004).

When approximately 100% of the data are contained within the 95PPU in the first sampling round, the value of  $\bar{d}$  tends to be rather high. Further sampling rounds are therefore required with new parameter ranges derived from:

$$b'_{j,min} = b_{j,lower} - \max\left(\frac{b_{j,lower} - b_{j,min}}{2}, \frac{b_{j,max} - b_{j,upper}}{2}\right) \quad \text{Equation 12}$$

$$b'_{j,max} = b_{j,upper} + \max\left(\frac{b_{j,lower} - b_{j,min}}{2}, \frac{b_{j,max} - b_{j,upper}}{2}\right) \quad \text{Equation 13}$$

Where  $b'$  indicate updated value. The top  $p$  solutions are used to calculated  $b_{j,lower}$  and  $b_{j,upper}$ , and the largest  $(b'_{j,max} - b'_{j,min})$  is used for the updated parameter range.

The above criteria ensure that the updated parameter ranges are constantly centred on the most recent best estimates while also producing reduced parameter ranges. If the best estimates are on the verge of their limits, parameter ranges are widened without going over the absolute limits. The final phase involves ranking the parameters according to their sensitivities and identifying the parameters that are highly linked. The strongly correlated parameters should be fixed to their best estimates and excluded from further sampling rounds for those with smaller sensitivities. This process is repeated until the two-model prediction uncertainty stopping rules are met. The generated parameter ranges are regarded as the ones that fit the subject under investigation the best (K. C. et al., 2004).

The conventional  $R^2$  and NSE statistics cannot be used to compare the 95PPU with observational signals. As a result, Abbaspour et al., (2004) advise using the  $p$ -factor and  $r$ -factor. The  $p$ -factor and  $r$ -factor, which express how well the simulation reproduces the observed data, often equal 100% and 0, respectively. An indicator of a model's capacity to account for uncertainty is the  $p$ -factor. Whereas the 95PPU's thickness is indicated by the  $r$ -factor, which measures the accuracy of the calibration. The  $r$ -factor is the average thickness of the 95PPU band divided by the standard deviation of the observed data, and the  $p$ -factor is the percentage of the observed data that is contained inside the 95PPU band (Kouchi et al., 2017). Since model calibration and uncertainty assessment are closely related, the  $p$ -factor and  $r$ -factor jointly show the strength of the model (Arnold, Moriasi, et al., 2012).

#### 4.3.3.3 *Parallel processing*

The SRB SWAT model was calibrated by fitting it to the observed streamflow data, which required a total of 2000 iterations to accomplish. This study applied parallel processing to speed up the calibration process instead of using single processing to run the 2000 iterations in SWAT-CUP. By utilising more than one central processing unit (CPU) or processor core in order to carry out the execution of a programme, this approach is capable of doing several operations or tasks all at the same time. For the purpose of carrying out hydrologic model calibration and uncertainty analysis, parallel processing is a method that was developed by making use of the best aspects of the many systems that were already accessible (Rouholahnejad et al., 2012).

#### 4.3.3.4 *Calibration*

Most of the researcher focused on the influence of precipitation element on hydrological model (Li et al., 2018), but insufficient literatures focus on the influence of the others meteorological elements on hydrological models (Gu et al., 2020). As mention in *section 4.3.1*, this study applied 5 hydro-meteorological elements in the calibration process, whether the sources are from single source or combination with a global reanalysis dataset. Seven scenarios were test in order to get the optimum result and minimise the uncertainties in the hydrological model.

Researchers most frequently compare model simulation to actual streamflow recorded at the outlet basin using the single-site calibration (SSC) method. SSC is commonly used in a lumped model such as an empirical or conceptual model where it is designed to simulate overall runoff and streamflow at the catchment's exist, rather than discrete flows (Sitterson et al., 2018). As shown in Figure 4-5(a), the model parameters used in this study were uniformly set across each SRB sub-basin and simulated for each dataset scenario at the catchment outlet. Three years from 2011-2013 were defined as calibration periods after 2008-2010 which were used as model warm-up periods.

#### 4.3.4 Model Validation

Validation is the process of establishing confidence in calibrated parameters to be applied without modification to an independent measured dataset, and it should be evaluated statistically and visually. Five (5) hydro-meteorological elements, as shown in Table 4-1, were utilised to assess the precision and application potential in streamflow simulation. Based on the effectiveness of the objective functions and the best dataset used in the model to reduce uncertainty, the best single model was identified. The period of 2014-2016 was used for the validation process in this study.

Five objective functions, including Kling Gupta Efficiency (KGE), Nash-Sutcliffe Efficiency (NSE), coefficient of determination ( $R^2$ ), Percent Bias (Pbias), and the RMSE-observation standard deviation ratio (RSR), were utilised to establish the confidence level and to acquire a better understanding of the model's performance. The following are the calculated model evaluation metrics:

KGE is the decomposition of NSE in three components of alpha, beta, and  $r$  with range from  $-\infty$  to 1. The King-Gupta efficiency (KGE) equation is given by:

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad \text{Equation 14}$$

Where  $r$  is the linear correlation between observations and simulation,  $\alpha$  a measure of the flow variability error, and  $\beta$  a bias term. KGE values range from  $-\infty$  to 1, with KGE=1 representing a perfect fit between datasets, KGE=0 representing a good fit on average values, and KGE<0 representing worse fitting than using the mean as a predictor (Castaneda-Gonzalez et al., 2018).

The NSE is utilised to evaluate the predictive capacity of the model and determines the degree to which the observed and simulated flows are comparable (Nash & Sutcliffe, 1970). The Nash-Sutcliffe efficiency (NSE) equation is given by:

$$NSE = 1 - \left[ \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \quad \text{Equation 15}$$

Where  $n$  is the total number of time steps,  $x_i$  is the simulated discharge at time  $t$ ,  $y_i$  is the observed flow at time  $t$ , and the  $\bar{x}$  mean observed discharge. Performance levels between 0 and 1 are considered to be acceptable, with  $NSE=1$  indicating full agreement between simulations and observations.  $NSE=0$  means that the model simulation has the same level of explanatory power as the observational mean. Meanwhile,  $NSE<0$  indicates that the model predicts worse than the observational mean (Knoben et al., 2019).

The average tendency of the simulated values to be greater or smaller than their observed counterparts is what the Pbias or percent bias metric attempts to quantify. The Pbias equation is given by:

$$Pbias = \left[ \frac{\sum_{i=1}^n (y_i - o_i)}{\sum_{i=1}^n o_i} \right] \times 100 \quad \text{Equation 16}$$

Where,  $y_i$  is observed flow, and  $o_i$  is simulated flow. The value of zero is considered to be optimum, with positive values indicating an underestimation bias and negative values indicating an overestimation bias (Gupta et al., 1999). This coefficient of determination ( $R^2$ ) evaluates the strength of the linear correlation between two variables and has been explained in subtopic 4.2.3. The coefficient of determination ( $R^2$ ) equation is given in equation 5. The RMSE-observations standard deviation ratio (RSR) is computed by dividing the RMSE by the standard deviation of measured data. RSR incorporates the benefits of error index statistics and a scaling or normalisation factor, allowing the statistic to be applied to a range of elements (D. N. Moriasi et al., 2007). The RSR equation is given by:

$$RSR = \frac{RMSE}{STDEV_{obs}} = \left[ \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \right] \quad \text{Equation 17}$$

The value of RSR might range anywhere from the ideal of 0 to a significant number that is positive. The lower the RSR, the lower the RMSE, and the better the performance of the model simulation (D. N. Moriasi et al., 2007).

#### **4.4 Multi-site calibration evaluation**

MSC is a new approach to enhancing the calibration process and it helps address the impact of human activities on water flow (Bai et al., 2017). MSC has emerged as a key development direction in contemporary hydrological simulation research, with the goal of increasing the model's simulation accuracy (Desai et al., 2021; Shrestha et al., 2016; Song et al., 2021). This is due to the limitation of SSC in its effectiveness, particularly on the numerous parameters and expressions of the spatial variability of a large watershed (Anderton et al., 2002; Bai et al., 2017). In addition, MSC allows a greater degree of parameter flexibility than SSC (Shrestha et al., 2016). According to Leta et. al., (2017), to improve the performance of the hydrological model, MSC was utilised to reflect spatial variability with varying parameter values.

To enhance the SSC performance results, this study applied the MSC approach in the SRB SWAT model. As shown in Figure 4-5(b), this is accomplished by dividing the SRB into three watersheds based on the calibration and validation points. To evaluate the MSC performances, three MSC techniques were used in this study, which are basin-by-basin, simultaneous, and sequential. Similar to the SSC method, the basin-by-basin (BB) calibration technique determines a smaller watershed than the SSC. Instead of lumping model parameters and calibrating them as a single large watershed, the BB technique is more realistic in that the parameters are determined and calibrated independently for each (B1, B2 and B3) independently. The simultaneous (SM) calibration technique is the second MSC method employed in this study. In the simultaneous technique, parameters were calibrated at once in all basins (B1, B2 and B3). Meanwhile, the sequential technique is the third MSC technique used in this study. In the sequential technique, a set of parameters were first determined on the upper stream basin (B3), followed by the downstream basins of B2 and B1. Because streamflow data in basin B2 is limited, this study used the regionalisation technique to determine the optimum parameters. The model was simulated in this study up to the model's maximum performance at both the calibration

stations and the five objective functions covered in subtopic 4.3.4, in order to assess the effectiveness of the MSC approach, were utilised.

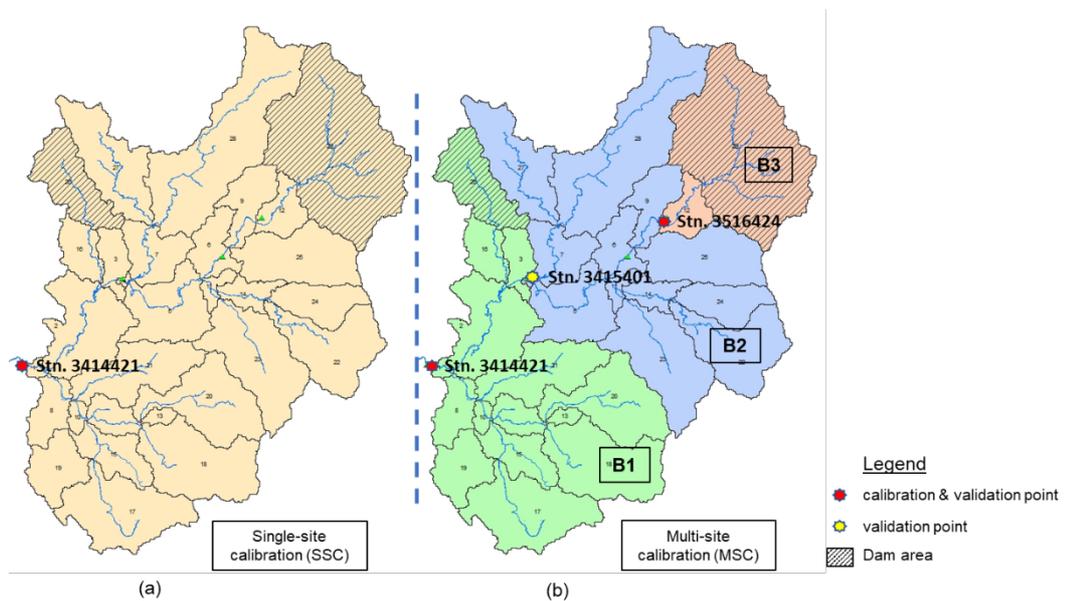


Figure 4-5: Calibration techniques implemented in SRB SWAT model for (a) single-site and (b) multi-site.

#### 4.5 Analysis multi-site calibration technique

Three calibration techniques of MSC, which are basin-by-basin, simultaneous, and sequential, were applied in the SRB SWAT model in order to enhance the model performance. This study compares the effectiveness of the three techniques using a validation method from the objective functions discussed in subtopic 4.3.4 in order to determine the best hydrological model for SRB. The basin-by-basin technique was illustrated in Figure 4-6(a), while the simultaneous and sequential techniques were illustrated in Figure 4-6(b).

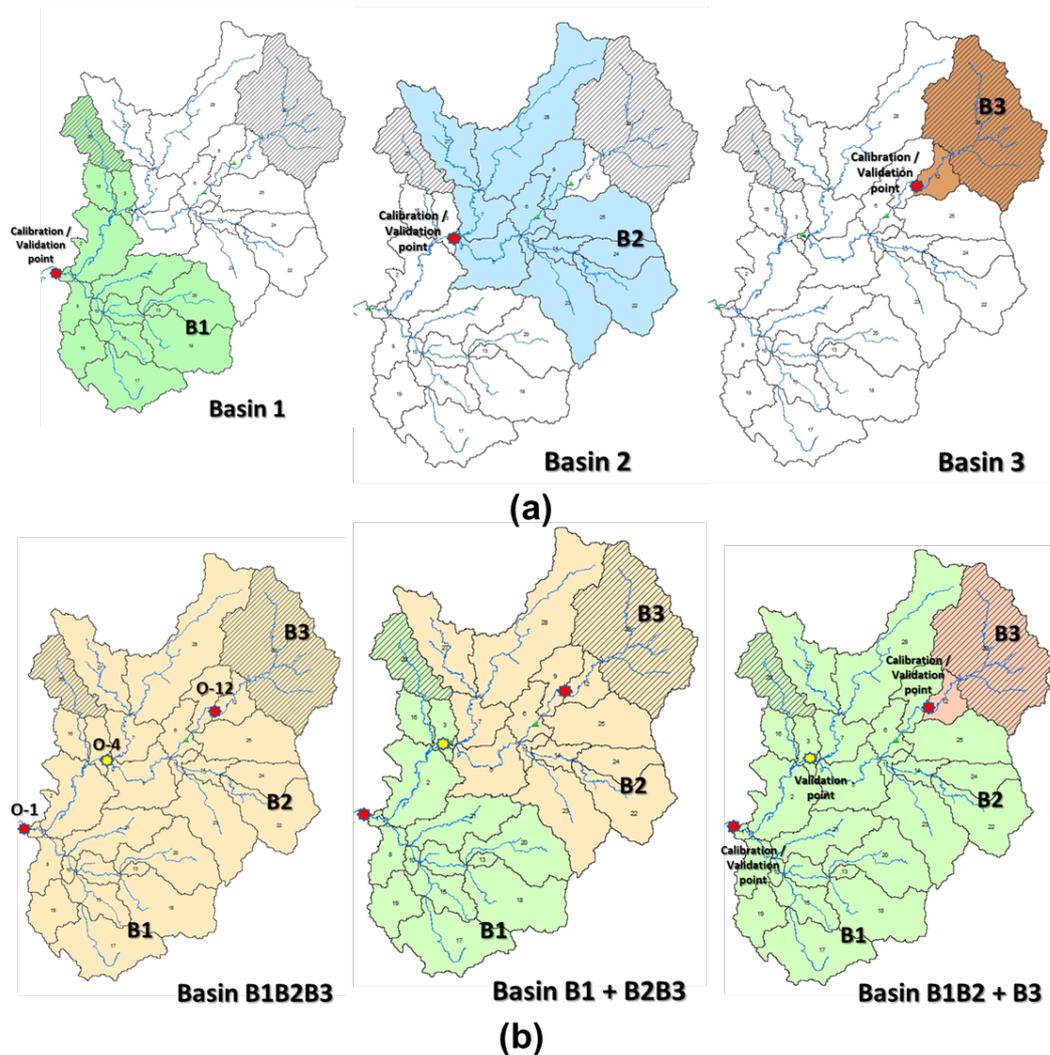


Figure 4-6: Multi-site calibration method used on SRB SWAT model.

#### 4.6 Climate change impact assessment

The General Climate Model (GCM), the most advanced and trustworthy instrument available for analysing the impacts of greenhouse gases on the atmosphere and predicting how climate variables are changing (Gebremeskel et al., 2005), was used to evaluate the impact of climate change on SRB water resources. A downscaling strategy is usually required for a better translation of climate data because GCMs are typically provided at coarse resolution. Dynamic downscaling and statistical downscaling are the two main downscaling techniques (Nkululeko Simeon Dlamini et al., 2017).

In this study, the coarse spatial resolution of GCMs was downscaled to a smaller gridded resolution using the Regional Hydroclimate Model of Peninsular Malaysia (RegHCM-PM). The coupling of land surface hydrologic activities with atmospheric processes is necessary for proper assessment of land surface fluxes (Kavvas et al., 2013). The RegHCM-PM and the SWAT model were linked in this work to reflect the interactions better accurately between the atmosphere and land surface hydrologic processes. Figure 4-7 illustrates the general procedures of the assessment of the effects of climate change on the water resources of the SRB.

Amin et al. (2017) advocates the ensemble approach of numerous GCMs to assess uncertainties from the inherent unpredictability of the climate system, model uncertainty, and scenario uncertainty, which are outlined by Hawkins and Sutton (2009). The maximum, lowest, and average of 15 different future climatic realisations produced from the coarse resolution of the General Climate Model (GCM) forecasts were used to create future precipitation and temperature estimates for 2030, 2040, and 2050. The climate change projection datasets were acquired from the National Water Research Institute of Malaysia (NAHRIM), the government agency in charge of managing and producing climate change projection data for Malaysia (<https://www.nahrim.gov.my/en.html>). In addition, these climate projection datasets were published by NAHRIM in the vulnerability and adaptation to climate change in Malaysia report that was compiled as part of the Third National Communication (TNC) report (NAHRIM, 2016) .

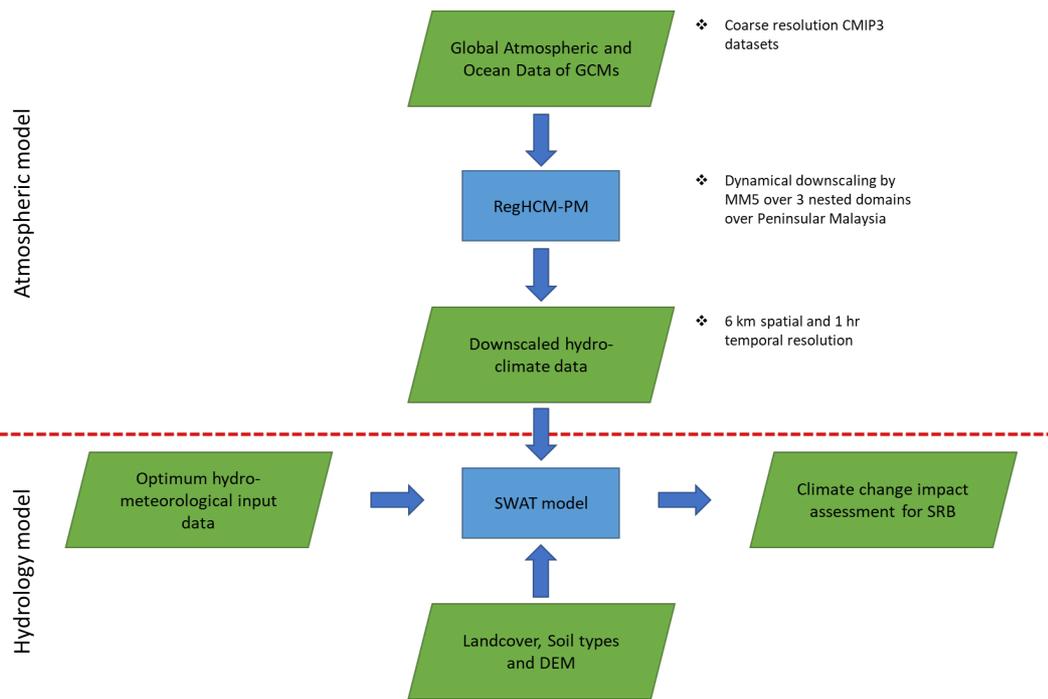


Figure 4-7: Schematic description of the climate modelling approach

#### 4.6.1 Regional Hydro Climate Model for Peninsular Malaysia

In this study, three general circulation models (GCMs) covering four emission scenarios (SRES B1, A1FI, A1B, and A2) based on Coupled Model Intercomparison Project phase 3 (CMIP3) datasets were dynamically downscaled to a reduced gridded resolution of 6 kilometres using the RegHCM-PM. This was done in order to analyse the impact that climate change will have on Peninsular Malaysia's hydrology and water resources. RegHCM-PM was utilised because it generated plausible simulations of hydrological hydroclimatic trends over Peninsular Malaysia, and the simulated hydroclimatic variables (air temperature, precipitation, and river flow) have comparable magnitudes and seasonal trends to their observed counterparts at the scales of watersheds and subregions (Ir. Mohd Zaki bin Mat Amin et al., 2019). Furthermore, the integrated hydroclimate simulations of the RegHCM-PM have been validated by comparing them to historically observed counterparts over selected river stations and subregions (Chen et al., 2006)

#### **4.6.2 Ensemble data of GCMs model**

Globally, General Circulation Models (GCMs) have produced many climate change projection datasets, and they are available at very coarse resolutions. In order to account for model uncertainty in the simulation of climate change, three different coupled land-atmosphere-ocean Global Circulation Models (GCMs) were used in this study. These GCMs were based on the Fourth Assessment Report (AR4) of the United Nations Intergovernmental Panel on Climate Change (IPCC) (IPCC, 2007a). The applied models included the Coupled Ocean-Atmosphere General Circulation Model version 2.3.2 (Yukimoto et al., 2001), the fifth-generation global climate model ECHAM5 (Roeckner et al., 2006) from the German Max-Planck Institute (MPI), and the third-generation community climate model CCSM3 (Collins et al., 2006) from the University Corporation for Atmospheric Research (UCAR) in the United States. The whole range of future greenhouse emission scenarios from the Special Report on Emission Scenarios (SRES) were included in order to account for the uncertainty in future greenhouse gas emissions (Nakicenovic et al., 2000). The ensemble data utilised by Amin et. al., (2017), to explore the assessment of the potential effects of climate change in Malaysia's Muda River Basin and Dungun River Basin are equivalent to the ensemble data used in this study.

#### **4.6.3 Climate simulation**

The best SWAT model, which was created in the earlier section 4.5 (objective three), was employed since the goal of this study is to ascertain how the consequences of climate change would manifest in the water supply. The climate projection datasets were prepared in a SWAT database as described in subtopic 4.3.1, then the SWAT model simulated the climate change projections on climatic variables for three future periods (2030s, 2040s, and 2050s). In the end, the impacts of climate change on SRB water resources are explored.

### Rainfall and Temperature analysis trend

This study applied the non-parametric Mann-Kendall test with Sen's slope as this method was widely used by researchers. The trend analysis was carried out in annually data series to check the trend whether decreases, increases or no trends. While the Sens' slope estimator is used to estimate the magnitude of trend and regression model was developed for the simulation datasets. The statistics of the Mann-Kendall test ( $S$ ) are given below:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sign}(x_j - x_i) \quad \text{Equation 18}$$

Where,  $n$  is number of data points,  $x_j$  and  $x_i$  are annual values in years  $j$  and  $i$ ,  $j > i$  and  $\text{sign}(x_j - x_i)$  calculated using the equation:

$$\text{sign}(x_j - x_i) = \begin{cases} -1 & \text{for } (x_j - x_i) < 0 \\ 0 & \text{for } (x_j - x_i) = 0 \\ +1 & \text{for } (x_j - x_i) > 0 \end{cases} \quad \text{Equation 19}$$

A positive  $S$  value indicates an ever-increasing trend, and a negative value indicates a downward trend. The variance ( $S$ ) is calculated with the following equation:

$$\text{Var}(S) = \frac{1}{18} \left[ n(n-1)(2n+5) - \sum_{p^{th}}^g t_p(t_p-1)(2t_p+5) \right] \quad \text{Equation 20}$$

Where,  $n$  is the data point number,  $g$  is the zero difference between compared values number,  $t_p$  is the number of data points in the  $p^{th}$  group. A standardised measure of test statistics ( $Z_{mk}$ ), determined using the following equation:

$$Z_{mk} = \begin{cases} \frac{s-1}{\sqrt{\text{Var}(s)}}, & \text{if } s < 0 \\ 0, & \text{if } s = 0 \\ \frac{s+1}{\sqrt{\text{Var}(s)}}, & \text{if } s > 0 \end{cases} \quad \text{Equation 21}$$

The determined standardised  $Z_{mk}$  values follow distribution normal with variances normal "0" and "1", it is utilised a measure of trend significance.

## CHAPTER 5

### DATASETS PERFORMANCE ANALYSIS

This chapter discusses the performance of global satellite/reanalysis datasets on observed data and their application in the SWAT (Soil and Water Assessment Tool) model. To set up an optimum hydrological model, 35 hydro-meteorological stations and two global reanalysis datasets, the Climate Forecast System Reanalysis (CFSR) and the China Meteorological Assimilation Driving Datasets (CMADS), were employed and assessed in this study.

#### **5.1 Evaluation of global reanalysis datasets.**

The consistency and reliability of input data are important in determining a realistic hydrological model result. A good rainfall-runoff model relies on the quality of the input data to represent the area. In this study, input data from gauged stations was examined for any errors or missing data before being used in the model whereas the bias correction is applied to check the difference between observed and global reanalysis data. To examine the consistency of precipitation and streamflow data from 33 hydrological stations, the Double Mass Curve method is used and verified using statistical analysis of the coefficient of determination ( $R^2$ ). The global reanalysis datasets (CMADS and CFSR) were evaluated using point-to-point analysis and verified using three statistics analysis, which are the correlation coefficient ( $r$ ), mean absolute error (MAE), and the root-mean-square error (RMSE). Meanwhile, in the SWAT model, input data from these three sources (gauged station, CMADS, and CFSR) were evaluated using five statistical analyses, including Kling-Gupta Efficiency (KGE), Nash-Sutcliffe Efficiency (NSE), Coefficient of determination ( $R^2$ ), Percent Bias (Pbias) and the RMSE-observations standard deviation ratio (RSR).

### 5.1.1 Data screening

The entire 2008 to 2019 observed data was screened to check the reliability and consistency. However, the exact timing of the screening depended on the data's availability because some of it had a limited period, such as CFSR, only available up to 2014. It is good practice to ensure that any errors or missing data are fixed before being incorporated into the hydrological model.

#### *Errors and missing data analysis*

Six climate variables (precipitation, temperature, relative humidity, solar radiation, and wind speed) together with streamflow were considered as input data to the hydrological model. However, due to the limited data from gauged stations, only two climate variables (precipitation and temperature) and streamflow were available for errors or missing data examination. No missing data detected on the streamflow and precipitation data recorded from gauged stations between 2008 and 2019. However, there are approximately 47% and 38% of the daily maximum and minimum temperatures, respectively, missing data as illustrated in Figure 5-1. Most of the missing data is from August 2013 to April 2015 and April 2017 to March 2019.

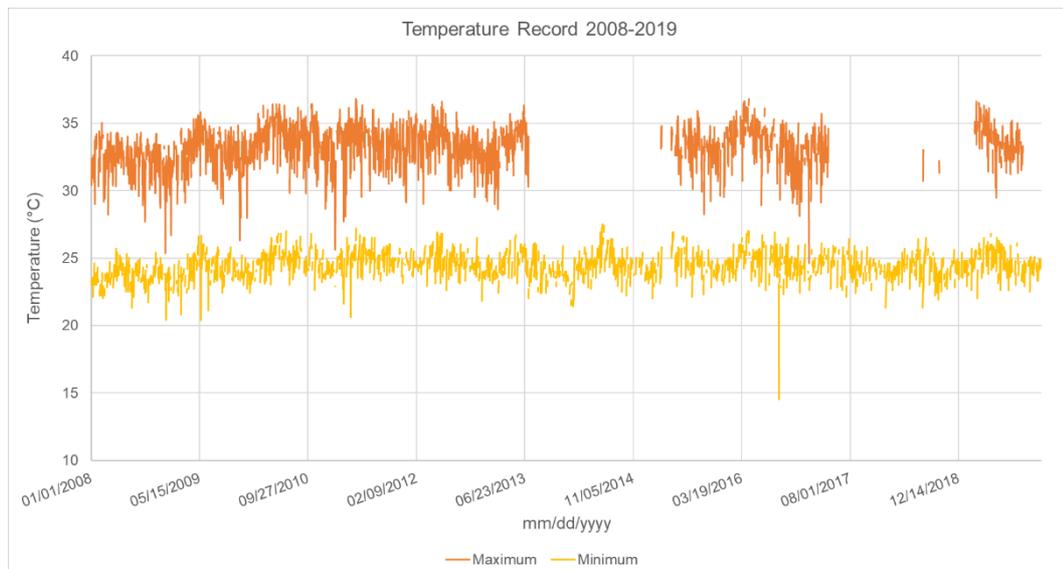


Figure 5-1: Missing data analysis for temperature data record from 2008-2019

A linear interpolation method was applied to fill up the missing data in maximum and minimum temperatures. However, the linear interpolation method is not applicable for a long duration as it will determine a straight line, such as in January-December 2014, as depicted in Figure 5-2.

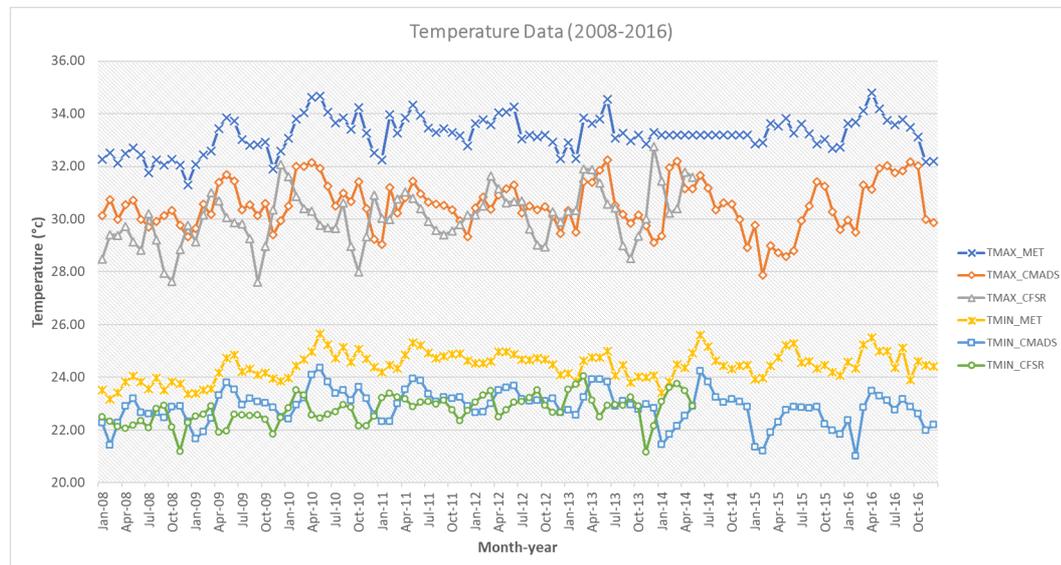


Figure 5-2: Monthly temperature data (2011-2016) after using the linear interpolation method to fill in the missing data

From the plotted graph in Figure 5-2, it shows the differences in range between gauged stations and global reanalysis data. The maximum and minimum temperatures recorded from the gauged stations showed a higher range compared to CMADS and CFSR. Meanwhile, the maximum and minimum temperatures from CMADS and CFSR were recorded at the same ranges. The average maximum and minimum temperatures of the gauged station is 33.2°C and 24.4°C, respectively. Whereas the average maximum and minimum temperatures data of CMADS and CFSR are about 30°C and 22°C, respectively. CMADS shows a similar temperature pattern to gauged stations, particularly from January 2011 to July 2013 for the maximum temperature, and January 2011 to December 2014 for the minimum temperature. The bias correction between observed and global reanalysis datasets were continued in the point-to-point analysis.

### 5.1.2 Double mass curve

The earliest precipitation and streamflow data from the observed station was in 1935 and 1960, respectively. However, it is not applicable to other stations, and the precipitation data from meteorological stations was exempted due to data availability. The most available data is in 2008 onwards, referring to all hydrological stations except for one streamflow station which was only available in 2014 onwards. As a result, double mass curve analysis examined the precipitation and streamflow data from 2008-2019. The correlation of determination ( $r^2$ ) is calculated for 30 rainfall stations and 3 streamflow stations in the Selangor River Basin (SRB). The statistical analysis results on precipitation and streamflow data were illustrated in Figure 5-3 and Figure 5-4, respectively. Meanwhile, Table 5-1 shows the calculated values of  $r^2$  on precipitation data. From the analysis, it shows most of the results were nearly 1.0, with average values of precipitation and streamflow data are 0.99 and 0.98, respectively. These results indicate that the precipitation and streamflow data are consistent and reliable to be used as input data in the SRB hydrological model.

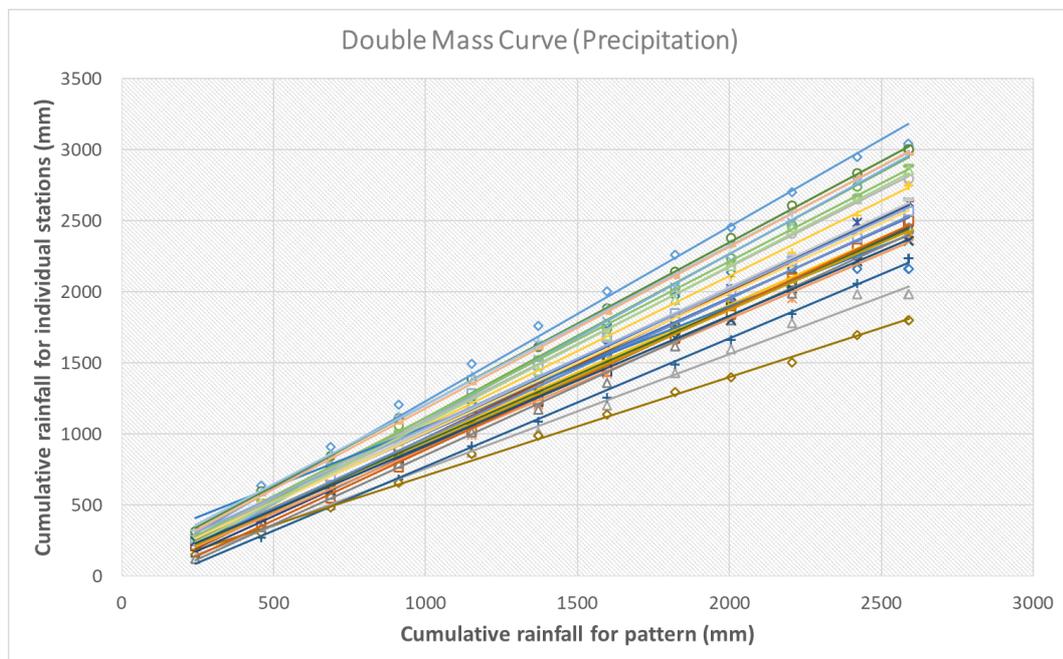


Figure 5-3: Double Mass Curve analysis on 2008-2019 precipitation data

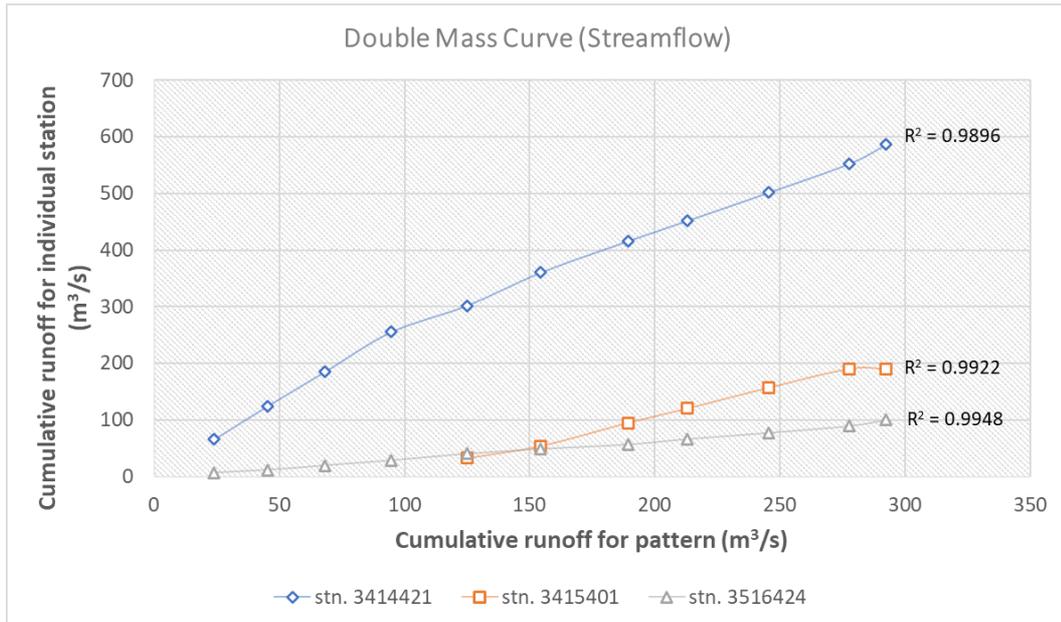


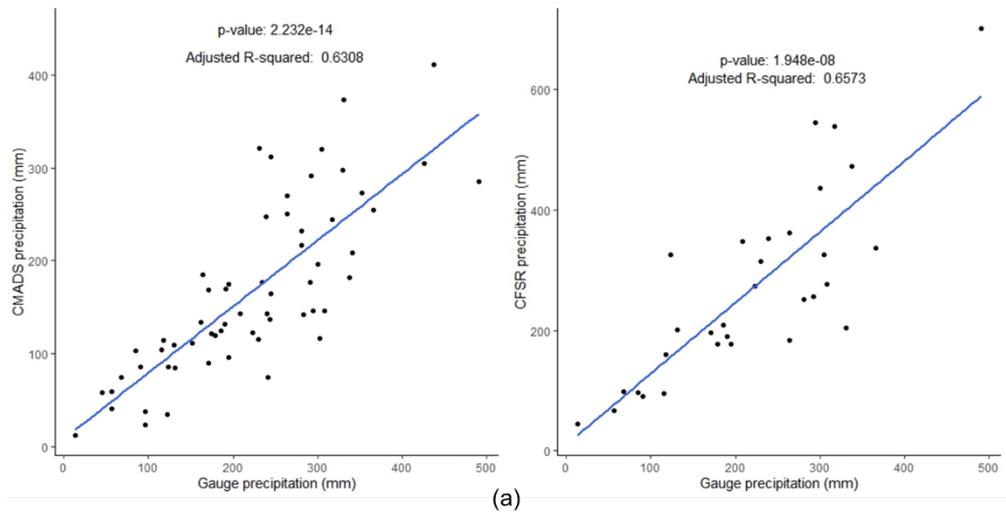
Figure 5-4: Double Mass Curve analysis on 2008-2019 streamflow data

Table 5-1: Double Mass Curve on yearly rainfall data from 2008-2019

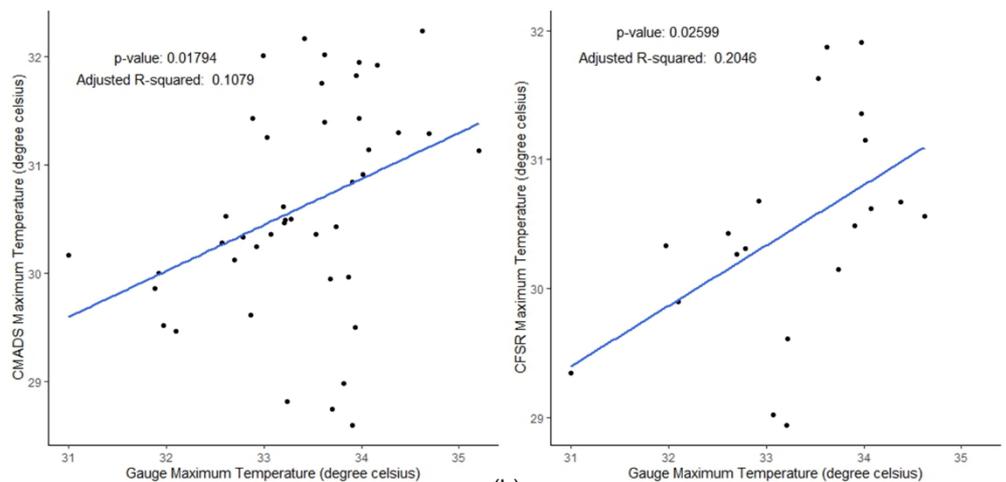
| Year / Statio | 2008  | 2009  | 2010  | 2011   | 2012   | 2013   | 2014   | 2015   | 2016   | 2017   | 2018   | 2019   | r     | r <sup>2</sup> |
|---------------|-------|-------|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------|----------------|
| 3211001       | 310.0 | 635.9 | 908.8 | 1207.3 | 1492.1 | 1760.2 | 2005.8 | 2262.3 | 2455.4 | 2701.7 | 2947.4 | 3044.4 | 0.999 | 0.998          |
| 3214054       | 169.8 | 331.7 | 499.0 | 674.2  | 871.8  | 1010.5 | 1201.7 | 1428.1 | 1597.7 | 1779.6 | 1985.3 | 1985.3 | 0.998 | 0.996          |
| 3214057       | 230.5 | 442.3 | 656.3 | 853.9  | 1109.2 | 1332.2 | 1548.0 | 1753.0 | 1926.1 | 2091.4 | 2268.3 | 2460.8 | 1.000 | 0.999          |
| 3215053       | 282.0 | 478.5 | 704.5 | 915.2  | 1144.1 | 1341.0 | 1554.3 | 1817.2 | 2025.0 | 2158.1 | 2329.8 | 2512.3 | 0.999 | 0.999          |
| 3215054       | 253.5 | 492.2 | 775.7 | 1057.4 | 1284.3 | 1516.0 | 1742.1 | 2029.7 | 2238.3 | 2469.7 | 2742.8 | 3004.0 | 0.999 | 0.999          |
| 3215055       | 138.5 | 269.8 | 489.0 | 680.8  | 915.3  | 1085.5 | 1255.5 | 1490.0 | 1661.4 | 1846.3 | 2060.0 | 2237.0 | 0.999 | 0.999          |
| 3216001       | 206.4 | 432.5 | 668.4 | 890.2  | 1113.3 | 1320.1 | 1559.9 | 1805.2 | 1987.4 | 2218.8 | 2427.7 | 2643.6 | 1.000 | 0.999          |
| 3216065       | 206.4 | 432.5 | 668.4 | 890.2  | 1149.8 | 1348.9 | 1571.0 | 1752.1 | 1913.5 | 2131.5 | 2339.6 | 2559.4 | 1.000 | 0.999          |
| 3314001       | 152.1 | 303.6 | 484.3 | 658.0  | 861.8  | 987.3  | 1138.9 | 1293.4 | 1399.3 | 1505.7 | 1697.4 | 1799.9 | 0.999 | 0.998          |
| 3314039       | 212.7 | 411.2 | 619.6 | 813.3  | 1025.7 | 1227.7 | 1436.6 | 1703.5 | 1892.5 | 2057.1 | 2298.0 | 2490.2 | 0.999 | 0.998          |
| 3314040       | 227.9 | 427.3 | 646.4 | 849.1  | 1090.1 | 1321.1 | 1547.7 | 1741.0 | 1893.9 | 2065.5 | 2257.9 | 2438.1 | 1.000 | 0.999          |
| 3315037       | 231.8 | 481.7 | 760.6 | 964.3  | 1220.4 | 1522.9 | 1789.6 | 2034.3 | 2223.7 | 2488.8 | 2770.3 | 2997.3 | 0.999 | 0.999          |
| 3315038       | 237.8 | 431.2 | 633.1 | 804.5  | 995.2  | 1206.4 | 1423.9 | 1617.1 | 1797.0 | 1955.3 | 2211.4 | 2390.7 | 0.999 | 0.998          |
| 3315039       | 280.4 | 508.9 | 772.2 | 971.3  | 1196.3 | 1490.0 | 1759.6 | 1985.8 | 2174.6 | 2411.8 | 2661.0 | 2793.0 | 1.000 | 0.999          |
| 3315040       | 260.8 | 460.1 | 717.4 | 981.2  | 1210.2 | 1464.5 | 1710.5 | 1931.2 | 2107.9 | 2275.3 | 2536.8 | 2748.8 | 1.000 | 0.999          |
| 3315041       | 245.6 | 448.3 | 689.3 | 875.4  | 1089.5 | 1303.0 | 1546.0 | 1746.0 | 1941.2 | 2121.7 | 2342.4 | 2580.3 | 0.999 | 0.999          |
| 3316028       | 291.6 | 533.1 | 764.7 | 1017.7 | 1264.8 | 1534.0 | 1756.6 | 1985.6 | 2203.1 | 2448.8 | 2675.1 | 2884.9 | 1.000 | 1.000          |
| 3317001       | 259.4 | 489.9 | 708.4 | 959.9  | 1222.0 | 1486.4 | 1775.4 | 1969.2 | 2139.0 | 2161.6 | 2161.6 | 2161.6 | 0.976 | 0.952          |
| 3317004       | 218.0 | 373.6 | 545.5 | 762.9  | 1012.0 | 1243.2 | 1471.9 | 1671.8 | 1839.0 | 2095.7 | 2309.4 | 2500.5 | 0.999 | 0.998          |
| 3414029       | 123.1 | 343.2 | 570.4 | 792.6  | 1014.8 | 1171.4 | 1360.7 | 1616.3 | 1802.5 | 1990.2 | 2264.7 | 2451.1 | 0.999 | 0.998          |
| 3414030       | 256.0 | 403.9 | 612.8 | 838.1  | 1062.9 | 1260.5 | 1502.9 | 1729.9 | 1903.7 | 2086.0 | 2277.0 | 2410.6 | 1.000 | 0.999          |
| 3415001       | 254.4 | 474.8 | 739.4 | 948.4  | 1151.1 | 1390.0 | 1621.8 | 1801.3 | 1974.9 | 2203.0 | 2488.9 | 2629.1 | 0.999 | 0.999          |
| 3416002       | 316.8 | 596.9 | 844.2 | 1112.6 | 1384.2 | 1613.9 | 1882.8 | 2139.9 | 2382.0 | 2607.0 | 2833.7 | 3002.5 | 1.000 | 1.000          |
| 3416029       | 318.2 | 593.8 | 851.9 | 1125.3 | 1410.2 | 1652.7 | 1883.6 | 2117.0 | 2335.6 | 2558.3 | 2778.7 | 2970.2 | 1.000 | 0.999          |
| 3515028       | 311.2 | 570.7 | 838.5 | 1077.9 | 1353.7 | 1601.3 | 1850.5 | 2101.6 | 2323.7 | 2587.5 | 2811.2 | 2969.0 | 1.000 | 1.000          |
| 3516022       | 274.2 | 506.3 | 709.8 | 922.5  | 1186.7 | 1390.5 | 1597.6 | 1812.1 | 2022.0 | 2256.0 | 2448.6 | 2649.7 | 1.000 | 0.999          |
| 3516027       | 265.9 | 510.0 | 724.0 | 905.6  | 1130.0 | 1348.3 | 1577.5 | 1805.1 | 1983.0 | 2194.3 | 2428.8 | 2588.1 | 1.000 | 0.999          |
| 3517022       | 246.1 | 475.4 | 703.1 | 976.8  | 1283.5 | 1483.2 | 1676.0 | 1850.6 | 2010.7 | 2217.3 | 2380.2 | 2566.8 | 0.998 | 0.996          |
| 3615002       | 272.8 | 486.8 | 746.0 | 971.3  | 1204.3 | 1445.3 | 1690.7 | 1943.3 | 2179.5 | 2452.5 | 2650.0 | 2853.3 | 0.999 | 0.999          |
| 3717101       | 243.9 | 405.7 | 601.2 | 831.0  | 1081.6 | 1291.4 | 1501.9 | 1677.3 | 1799.2 | 2022.6 | 2214.9 | 2361.7 | 0.999 | 0.999          |
| Average       | 243.3 | 458.4 | 688.4 | 911.0  | 1151.0 | 1371.6 | 1598.0 | 1820.4 | 2004.4 | 2205.3 | 2420.0 | 2589.5 | 0.999 | 0.997          |



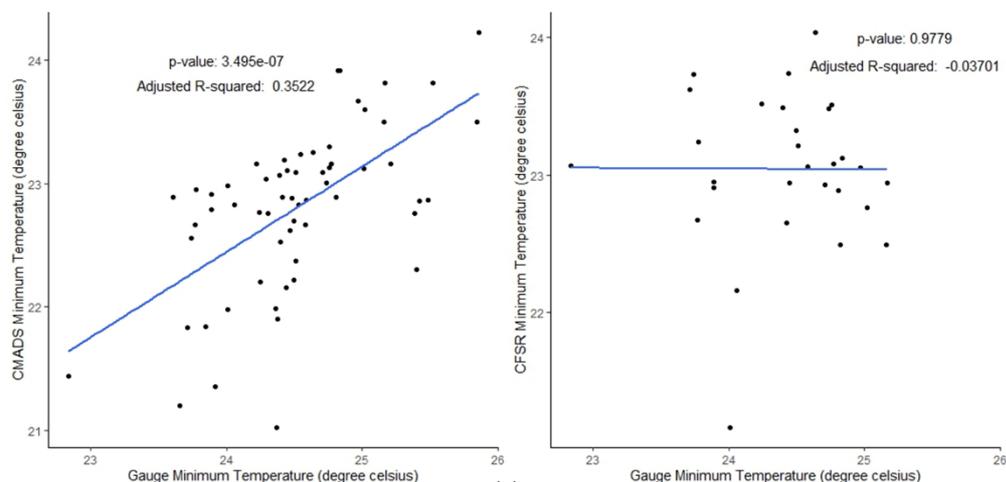
Validation results of global reanalysis datasets (CMADS and CFSR) to observed data from meteorological gauge station (Tennemaran Estate) on a monthly scale are depicted in Figure 5-6. On precipitation data, CMADS has a Pearson's R value of 0.63, and CFSR has a value of 0.65. Both the CMADS and CFSR precipitation data had highly significant *p-values* of less than 0.01 for their respective datasets as depicted in Figure 5-6(a). The *r-value* of maximum temperature for CMADS and CFSR shows in Figure 5-6(b), are 0.1 and 0.2, respectively. Both datasets are significant, with *p-value* of 0.02 and 0.03, respectively. Meanwhile the *r-value* for minimum temperature of CMADS is 0.4 while for CFSR is -0.04 as illustrated in Figure 5-6(c). CMADS's minimum temperature outperformed CFSR with a *p-value* of less than 0.01 and CFSR's minimum temperature of 0.97 more than 0.01. The Mean-Absolute-Error (MAE) of CMADS and CFSR precipitation is 64.9 and 75.8, respectively. While the Root-Mean-Square Error (RMSE) for CMADS is 84.1 and CFSR is 99.5, CMADS precipitation data is more significant compared to CFSR, where the MAE and RMSE results are both smaller than CFSR. The MAE of CMADS maximum and minimum temperatures are 2.82, and 1.65, respectively Whereas the MAE of CFSR maximum and minimum temperature are 2.99 and 1.42, respectively. From the results, temperature data from CMADS is better than CFSR, with MAE and RMSE values that are lower than CFSR. A summary of the statistical analyses is shown in Table 5-2. Overall, in point-to-point analysis, it shows that CMADS had superior performance estimation for observed data on a monthly scale than CFSR.



(a)



(b)



(c)

Figure 5-6: Scatter plots on monthly CMADS and CFSR to gauged stations: (a) Precipitation; (b) Max. temperature; (c) Min. temperature

Table 5-2: the point-to-point statistical analysis of CMADS and CFSR

|       |                  | <i>r</i> | MAE  | RMSE |
|-------|------------------|----------|------|------|
| CMADS | Precipitation    | 0.63     | 64.9 | 84.1 |
|       | Max. temperature | 0.11     | 2.82 | 14.9 |
|       | Min. Temperature | 0.4      | 1.65 | 1.7  |
| CFSR  | Precipitation    | 0.65     | 75.8 | 99.5 |
|       | Max. temperature | 0.2      | 2.99 | 29   |
|       | Min. Temperature | -0.04    | 1.42 | 21.5 |

*Bias correction*

Bias correction is a technique used to adjust climate model outputs or observational data to make them more accurate and reliable. It is particularly useful for rainfall data because precipitation measurements can be affected by various sources of bias, including gauge undercatch, wind effects, or systematic errors in measurements. Based on the point-to-point analysis, it shows an approximately  $\pm 30\%$  difference in the precipitation data between observed and global reanalysis data. Overall, CFSR precipitation data is higher than observed, while CMADS is lower than observed data, as shown in Figure 5-7. The CMADS precipitation value was increased by 30%, while the CFSR precipitation data was decreased by 30%, in order to adjust for bias in both global reanalysis datasets. The bias correction of CMADS and CFSR is shown in Figure 5-8 and Figure 5-9. From the results, the precipitation pattern from CMADS is more significant to observed data compared to CFSR. The magnitude of CMADS was nearly identical to the observed data after the correction had been made, with an average error of about 15 mm. However, the magnitude of CFSR was inaccurate, and the average error is about 40 mm.

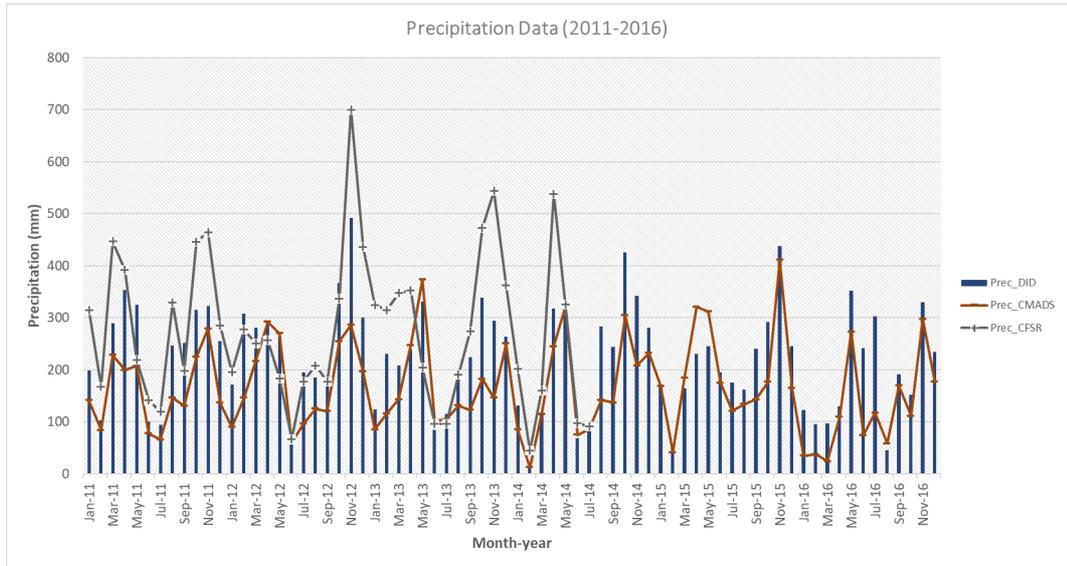


Figure 5-7: Monthly precipitation data analysis on CMADS and CFSR to gauged station.

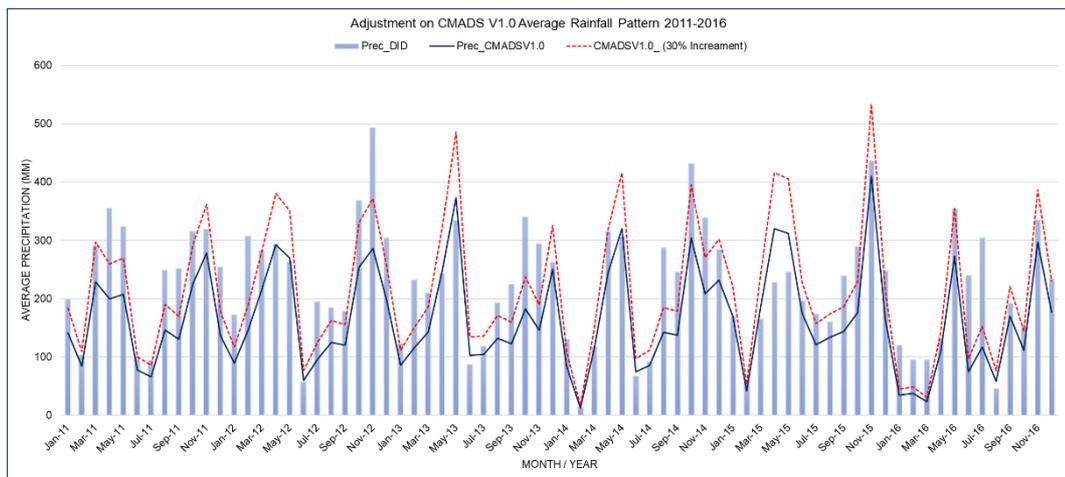


Figure 5-8: Bias correction on CMADS precipitation data

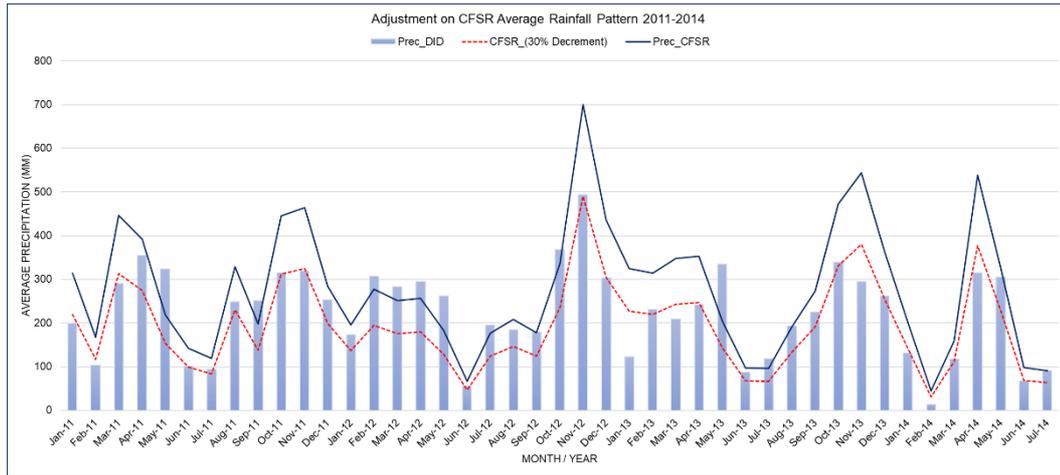


Figure 5-9: Bias correction on CFSR precipitation data

Analyses on maximum and minimum temperature data show the average bias correction for CMADS is about 0.6°C and 0.4°C, respectively, as depicted in Figure 5-10. As depicted in Figure 5-11, the maximum average bias correction for CFSR is about 0.7°C and the minimum is about 0.5°C. The findings indicate that the CMADS temperature pattern and the observed temperature pattern are identical.

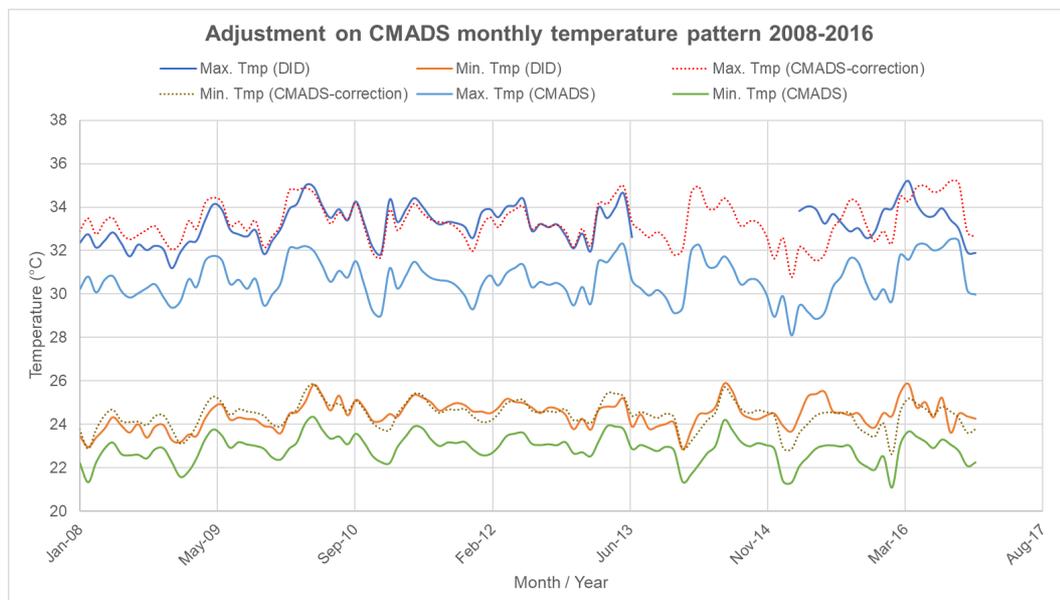


Figure 5-10: Bias correction on CMADS temperature data

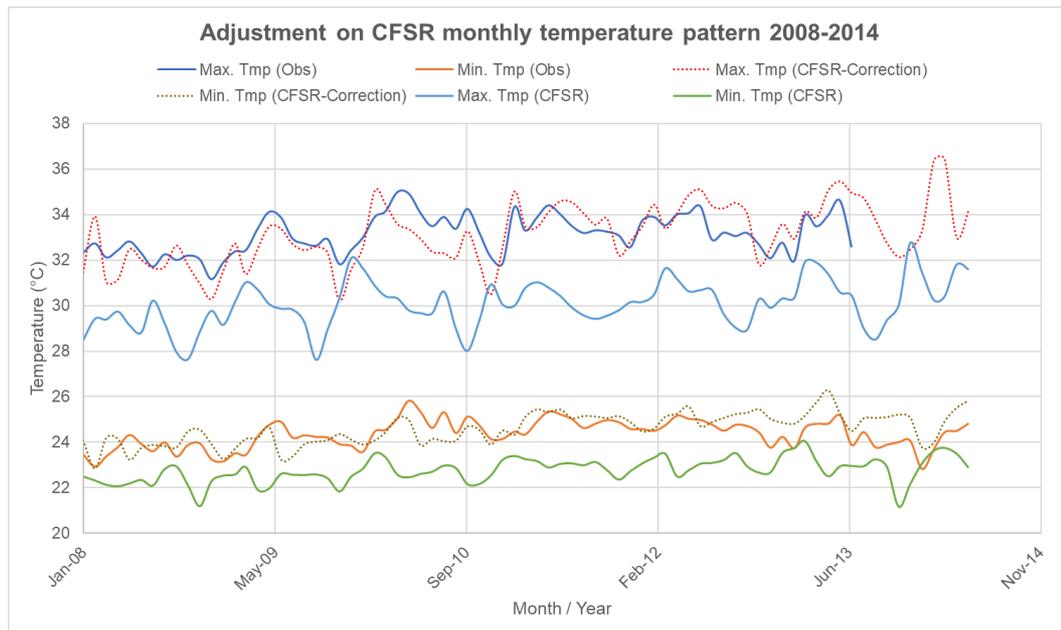


Figure 5-11: Bias correction on CFSR temperature data

As a result, the precipitation and temperature data from CMADS is more significant compared to CFSR and is reliable to be used in order to fill in the missing data from the gauged station.

## 5.2 Analysis on cross-combined datasets

The application of the observed and global reanalysis data as input data in hydrological models was assessed in this chapter. Input data was prepared based on 7 scenarios as described in Chapter 4, where Scenario 1 to Scenario 3 are the input data from individual sources. While Scenario 4 to 7 are the cross-combined datasets from multiple sources. All datasets were prepared in a SWAT model database format using Microsoft Access, as illustrated in Figure 5-12.

| COLLECTED | STATE   | STATION    | WATERSHIP  | ID | WATERSHIP    | WBSN          | RAIN_YES | IMPACT1 | IMPACT2 | IMPACT3 | IMPACT4 | IMPACT5 | IMPACT6 | IMPACT7 | IMPACT8 |
|-----------|---------|------------|------------|----|--------------|---------------|----------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1         | SLANGOR | 3213001wgn | 3213001wgn | 1  | 3.2909444444 | 101.612000556 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 2         | SLANGOR | 3214054wgn | 3214054wgn | 2  | 3.2691544444 | 101.457111111 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 3         | SLANGOR | 3214057wgn | 3214057wgn | 3  | 3.2972222222 | 101.400333333 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 4         | SLANGOR | 3215053wgn | 3215053wgn | 4  | 3.2508333333 | 101.529166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 5         | SLANGOR | 3215055wgn | 3215055wgn | 5  | 3.2161111111 | 101.376111111 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 6         | SLANGOR | 3215055wgn | 3215055wgn | 6  | 3.2161111111 | 101.376111111 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 7         | SLANGOR | 3216001wgn | 3216001wgn | 7  | 3.2722222222 | 101.466111111 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 8         | SLANGOR | 3216005wgn | 3216005wgn | 8  | 3.2508333333 | 101.466111111 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 9         | SLANGOR | 3216001wgn | 3216001wgn | 9  | 3.2698888889 | 101.412166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 10        | SLANGOR | 3216001wgn | 3216001wgn | 10 | 3.2698888889 | 101.412166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 11        | SLANGOR | 3216004wgn | 3216004wgn | 11 | 3.2698888889 | 101.412166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 12        | SLANGOR | 3216001wgn | 3216001wgn | 12 | 3.2698888889 | 101.412166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 13        | SLANGOR | 3216001wgn | 3216001wgn | 13 | 3.2722222222 | 101.511666667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 14        | SLANGOR | 3216001wgn | 3216001wgn | 14 | 3.2698888889 | 101.412166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 15        | SLANGOR | 3216001wgn | 3216001wgn | 15 | 3.2698888889 | 101.412166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 16        | SLANGOR | 3216001wgn | 3216001wgn | 16 | 3.2698888889 | 101.412166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 17        | SLANGOR | 3216001wgn | 3216001wgn | 17 | 3.2698888889 | 101.412166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 18        | SLANGOR | 3217001wgn | 3217001wgn | 18 | 3.3347222222 | 101.704166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 19        | SLANGOR | 3217001wgn | 3217001wgn | 19 | 3.3347222222 | 101.704166667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 20        | SLANGOR | 3418029wgn | 3418029wgn | 20 | 3.4754444444 | 101.488666667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 21        | SLANGOR | 3418030wgn | 3418030wgn | 21 | 3.4497222222 | 101.460833333 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 22        | SLANGOR | 3418030wgn | 3418030wgn | 22 | 3.4497222222 | 101.460833333 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 23        | SLANGOR | 3418030wgn | 3418030wgn | 23 | 3.4497222222 | 101.460833333 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 24        | SLANGOR | 3418029wgn | 3418029wgn | 24 | 3.4222222222 | 101.466666667 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 25        | SLANGOR | 3513001wgn | 3513001wgn | 25 | 3.3007777778 | 101.341000000 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 26        | SLANGOR | 3514002wgn | 3514002wgn | 26 | 3.2758333333 | 101.465555556 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 27        | SLANGOR | 3513002wgn | 3513002wgn | 27 | 3.3007777778 | 101.341000000 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 28        | SLANGOR | 3513002wgn | 3513002wgn | 28 | 3.3007777778 | 101.341000000 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 29        | SLANGOR | 3513001wgn | 3513001wgn | 29 | 3.4111111111 | 101.501111111 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |
| 30        | SLANGOR | 3717011wgn | 3717011wgn | 30 | 3.7114444444 | 101.718333333 | -9999    | 12      | 17.44   | 17.40   | 20.42   | 20.45   | 20.50   | 20.34   | 19.01   |

Figure 5-12: SWAT database prepared using Microsoft Access for 7 scenarios.

In the SWAT model, SRB was discretised into 29 sub-basins based on a 30 km<sup>2</sup> threshold area, as illustrated in Figure 5-13. Sub-basin 29 is the largest basin with 199.57 km<sup>2</sup>, whereas sub-basin 1 is the smallest basin with 3.81 km<sup>2</sup>. The dam areas were designated as sub-basins 26 and 29, and the watershed outlet was designated as sub-basin 1. The input and output data on dam operation are excluded from this study due to data restrictions and limitations.

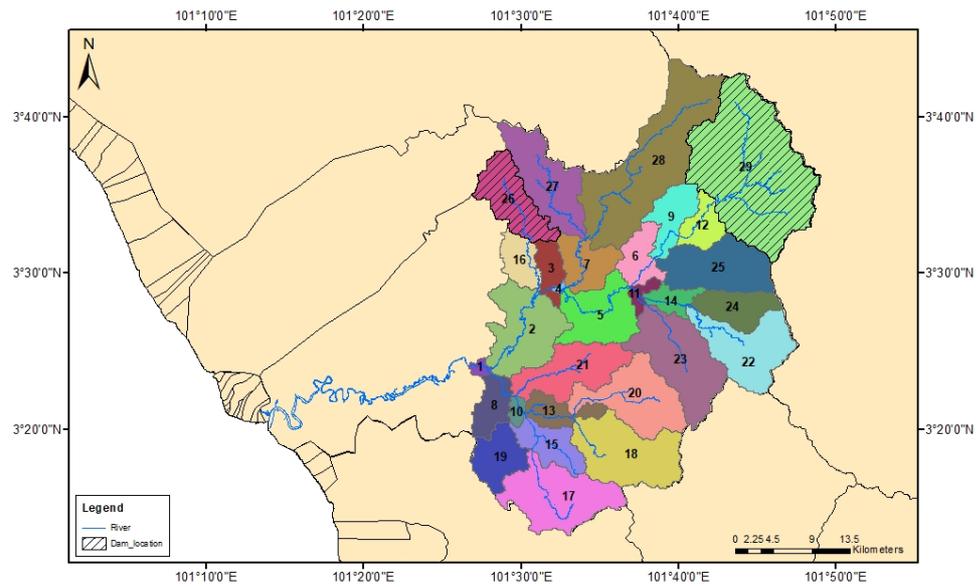


Figure 5-13: SRB sub-basins delineated using SWAT model.

Based on 29 sub-basins and a threshold of 10% of basin area, 171 hydrological response units (HRUs) were generated using topography, soil, and land use data, as illustrated in Figure 5-14.

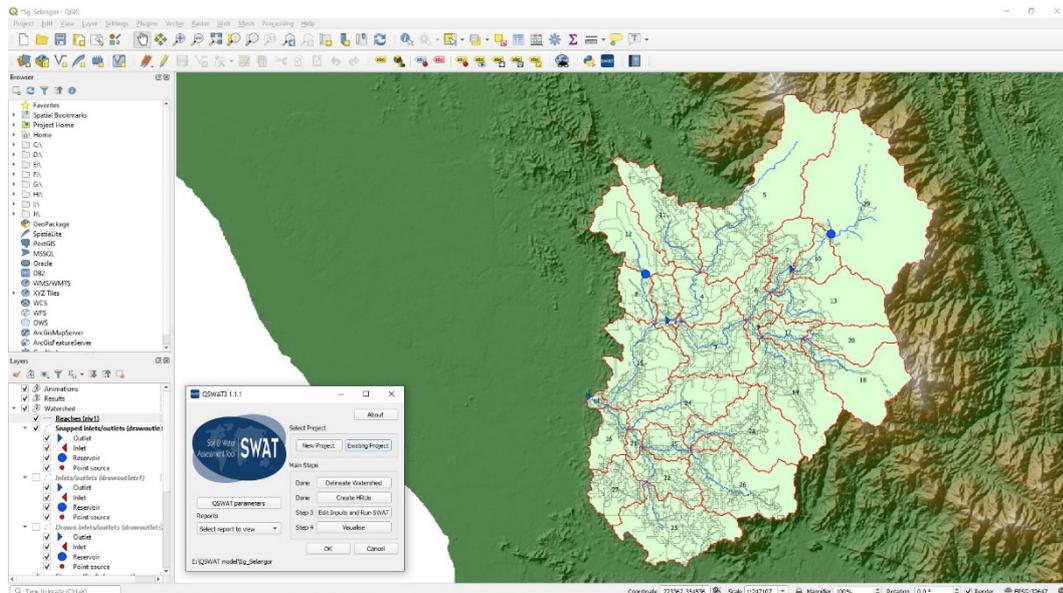


Figure 5-14: 171 Hydrological response units (HRUs)

All these datasets were applied in the SRB SWAT model, which was calibrated and validated using five objective functions, including Kling-Gupta Efficiency (KGE), Nash-Sutcliffe Efficiency (NSE), Coefficient of determination ( $R^2$ ), Percent bias (Pbias), and RMSE-observation standard deviation ratio (RSR). In the calibration and validation procedure, the single-site calibration (SSC) method was employed to analyse the performance of these input data. The model's performance was assessed using the streamflow station (stn. 3414421) at the SRB's outlet, as illustrated in Figure 5-15.



Figure 5-15: SRB single-site calibration and validation method

### 5.2.1 Model simulation

In the SSC method, a set of parameters defined in Table 5-3 were modelled in the SWAT rainfall-runoff model as one unit for the entire catchment of the SRB and the information was lumped as a single homogenous unit as shown in Figure 5-15. The 20 parameters were selected by referring to a similar study by other researchers in a global and local area, as described in chapter 4.

Table 5-3: List of parameters used in the calibration process

| No. | Parameters      | Unit              | Description   |
|-----|-----------------|-------------------|---|
| 1   | r_CN2.mgt       | %                 | Curve number for moisture condition II  |
| 2   | r_SOL_AWC.sol   | %                 | Available water capacity of the soil layer  |
| 3   | v_ESCO.hru      | -                 | Soil evaporation compensation factor  |
| 4   | v_GW_DELAY.gw   | days              | Groundwater delay   |
| 5   | v_GWQMN.gw      | mm                | Threshold depth of water in the shallow aquifer required for return flow to occur |
| 6   | v_GW_REVAP.gw   | -                 | Groundwater 'revap' coefficient   |
| 7   | r_HRU_SLP.hru   | m/m               | Average slope steepness   |
| 8   | r_OV_N.hru      | -                 | Manning's 'n' value for overland flow   |
| 9   | r_SLSUBBSN.hru  | m                 | Average slope length  |
| 10  | r_SOL_K.sol     | mm/h              | Saturated hydraulic conductivity  |
| 11  | r_SOL_BD.sol    | mg/m <sup>3</sup> | Moist bulk density  |
| 12  | v_CANMX.hru     | mm                | Maximum canopy storage  |
| 13  | v_CH_K2.rte     | mm/hr             | Effective hydraulic conductivity in main channel alluvium                         |
| 14  | v_ALPHA_BF.gw   | days              | Baseflow alpha factor   |
| 15  | v_CH_N2.rte     | -                 | Manning's 'n' value for the main channel  |
| 16  | v_LAT_TTIME.hru | days              | Lateral flow travel time  |
| 17  | v_RCHRG_DP.gw   | -                 | Deep aquifer percolation fraction   |
| 18  | v_SURLAG.bsn    | days              | Surface runoff lag time   |
| 19  | v_EPCO.hru      | -                 | Plant uptake compensation factor  |
| 20  | v_REVAMPMN.gw   | mm                | Threshold depth of water in the shallow aquifer for 'revap' to occur              |

SWAT Calibration and Uncertainty Program (SWAT-CUP) software is used to run the calibration and validation processes, and to speed up the process, parallel processing is used. This study used 2000 iterations in defining the best SWAT model parameters. The most influential parameters in the model were determined using sensitivity analysis.

### 5.2.2 Sensitivity analysis

The SWAT-CUP was used to determine which variables in the SRB SWAT model were the most sensitive. This study made use of the Sequential Uncertainty Fitting method (SUFI-2), a popular SWAT-CUP software module (Abbaspour 2008). In this program, 20 selected parameters in Table 5-3 were altered at the same time in a global sensitivity analysis technique for all scenarios of simulation. *P-value* and *t-stat* are the two indicators to determine the most sensitivity parameter. The significance of sensitivity results is determined using *P-value*, with values closer to zero being more statistically significant. While *t-stat* represents sensitivity, with the larger values in absolute terms being the more sensitive parameters in the model.

As a result, the baseflow alpha factor (ALPHA\_BF) and the curve number from moisture condition II (CN2) displayed in Figure 5-16 are the most sensitive parameters in the SRB SWAT model while the plants uptake compensation factor parameter (EPCO) is the least sensitive.

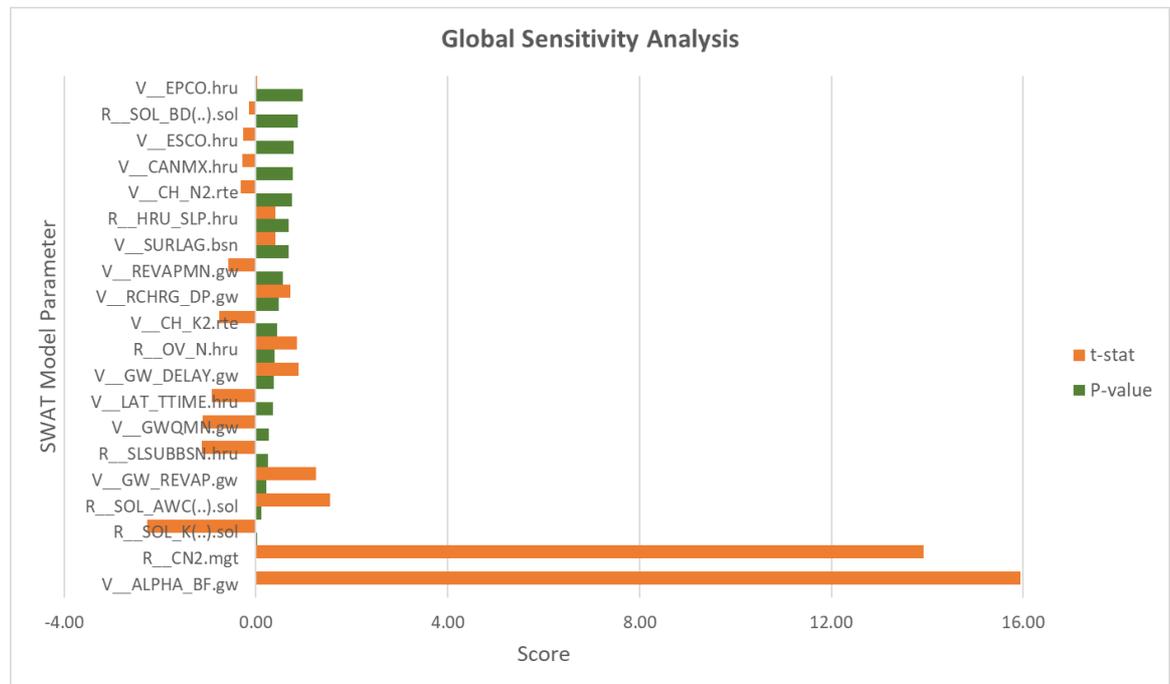


Figure 5-16: Global sensitivity analysis for SRB SWAT model parameters

### 5.2.3 Model evaluation

The SRB SWAT model was developed based on seven datasets, including single-source and multiple-source as described in Chapter 4. The calibration process was set up to 2000 iterations, and the process was repeated with the new set of parameters until the simulated streamflow achieved the best fit to the observed streamflow. To complete the process, calibrated parameters were validated with a new observed dataset over different time periods without making any further modifications to the calibrated parameters. In this study, data from 2008 to 2010 was used for a model warmup session. 2011 to 2013 was utilised as the calibration period, and 2014 to 2016 was used to validate the model. Table 5-4 lists the calibrated and validated results for the 7 scenarios, and Figure 5-17 and 5-18 show the streamflow simulation results.

Table 5-4: Calibrated and validated results on single-site method (SSC)

| SWAT Model  | SC-1           | SC-2           | SC-3           | SC-4           | SC-5           | SC-6           | SC-7           |                |
|-------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Calibration | R <sup>2</sup> | 0.87           | 0.53           | 0.17           | 0.84           | 0.82           | 0.86           | 0.86           |
|             | Range          | Very Good      | Satisfactory   | Unsatisfactory | Very Good      | Very Good      | Very Good      | Very Good      |
|             | NSE            | 0.84           | 0.44           | -0.29          | 0.83           | 0.82           | 0.85           | 0.84           |
|             | Range          | Very Good      | Satisfactory   | Unsatisfactory | Very Good      | Very Good      | Very Good      | Very Good      |
|             | Pbias          | 0.00           | -0.10          | -26.10         | 0.00           | 0.00           | -0.10          | 0.00           |
|             | Range          | Very Good      | Very Good      | Unsatisfactory | Very Good      | Very Good      | Very Good      | Very Good      |
|             | KGE            | 0.92           | 0.69           | 0.31           | 0.89           | 0.88           | 0.90           | 0.89           |
|             | Range          | Very Good      | Satisfactory   | Unsatisfactory | Good           | Good           | Good           | Good           |
|             | RSR            | 0.39           | 0.75           | 1.14           | 0.42           | 0.43           | 0.39           | 0.40           |
|             | Range          | Very Good      | Unsatisfactory | Unsatisfactory | Very Good      | Very Good      | Very Good      | Very Good      |
| Validation  | R <sup>2</sup> | 0.53           | 0.35           |                | 0.53           | 0.51           | 0.55           | 0.53           |
|             | Range          | Satisfactory   | Unsatisfactory |                | Satisfactory   | Satisfactory   | Satisfactory   | Satisfactory   |
|             | NSE            | 0.45           | 0.16           |                | 0.46           | 0.46           | 0.51           | 0.51           |
|             | Range          | Satisfactory   | Unsatisfactory |                | Satisfactory   | Satisfactory   | Satisfactory   | Satisfactory   |
|             | Pbias          | 0.00           | 0.00           |                | 0.00           | 0.10           | 0.20           | 0.00           |
|             | Range          | Very Good      | Very Good      |                | Very Good      | Very Good      | Very Good      | Very Good      |
|             | KGE            | 0.72           | 0.50           |                | 0.66           | 0.68           | 0.71           | 0.66           |
|             | Range          | Satisfactory   | Satisfactory   |                | Satisfactory   | Satisfactory   | Satisfactory   | Satisfactory   |
|             | RSR            | 0.74           | 0.92           |                | 0.73           | 0.74           | 0.71           | 0.70           |
|             | Range          | Unsatisfactory | Unsatisfactory |                | Unsatisfactory | Unsatisfactory | Unsatisfactory | Unsatisfactory |

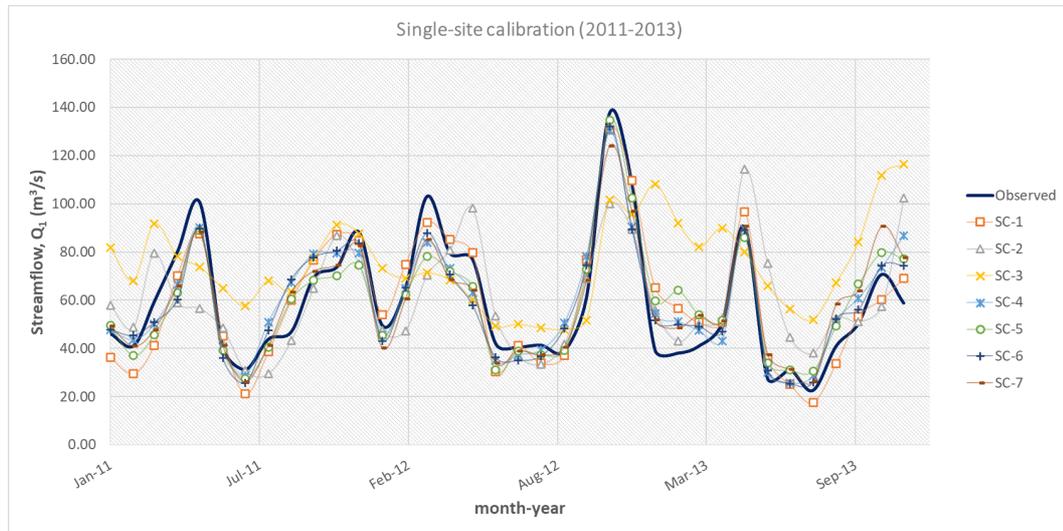


Figure 5-17: Single-site calibration results on seven scenario datasets

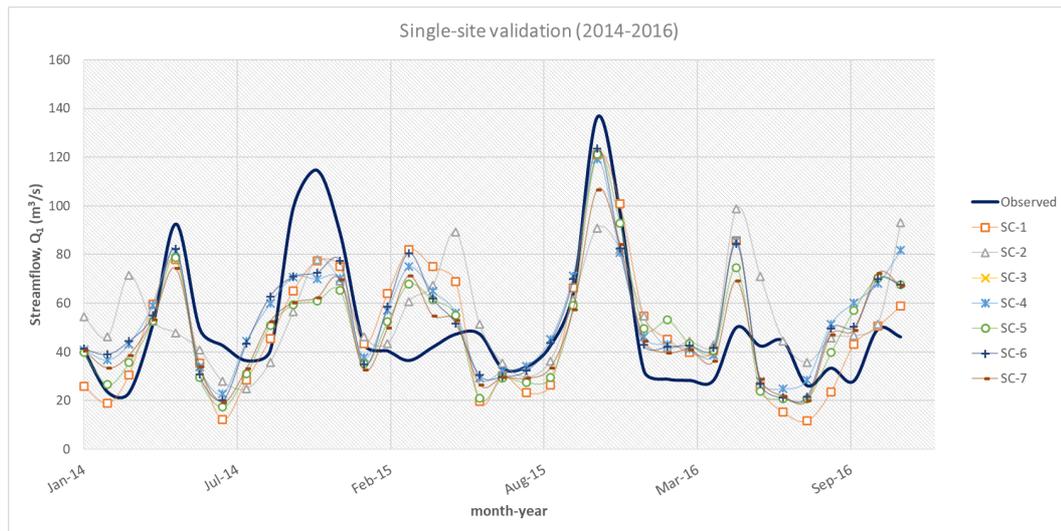


Figure 5-18: Single-site validation results on seven scenario datasets

Calibrated results shown in Table 5-4 indicate most of the scenarios achieved a good range in five performance ratings, except for scenario-3 (SC-3), where all performance ratings show an unsatisfactory range. Referring to the calibrated result in Figure 5-17, the SC-3 simulated graph does not match the observed data and does not reflect high and low flow conditions. The SC-3 result was removed from the validation process because it was outside the permitted range and lacked sufficient data. The calibrated findings for SC-2 reveal a mixed range of outcomes, with  $R^2$ , NSE and KGE being in the satisfactory range while Pbias and RSR were the good and unsatisfactory ranges, respectively. According to the calibrated graph, the SC-2 outcomes for the first half of 2011 did not perform well, but they improved from July 2011 to April 2013. From May 2013, the simulated graph accurately depicted the observed flow pattern but at a higher level. For a single dataset, SC-1 exceeded SC-2 and SC-3, with all five performance ratings indicating the calibrated model was in a very good range whereas for the cross-combine dataset (SC-4 to SC-7), all datasets attained a good or very good range in performance ratings. The simulated graph of these datasets was significantly reflected in the observed graph. Overall, the calibrated parameters from SC-1, SC-4, SC-5, SC-6, and SC-7 show a strong correlation with measured streamflow.

In the validation process, most of the scenarios in Table 5-4 had comparable validated results except for SC-2 which had objective function of  $R^2$ , and NSE recorded an unsatisfactory range. For all scenarios, the objective function of Pbias and RSR reveals a very good and an unsatisfactory range, respectively. Pbias reveals a very good range, even though the validated graph in Figure 5-18 indicates that the simulated results do not perfectly reflect the observed graph. A similar thing happened to RSR, where the results were unsatisfactory even though the graph displayed simulated results that captured some of the high and low flows especially in January to June 2014 and July 2015 to March 2016. Validated result in Table 5-4 shows the observed data from gauged station (SC-1) outperformed global reanalysis datasets (SC-2 and SC-3) in SWAT model. While CMADS (SC-2) performed better than CFSR (SC-3) in the calibration and validation process. According to the results, cross-combined datasets (SC-4 to SC-7) outperformed individual data. For instance, the  $R^2$  performance rating for SC-2 was 0.35. However, after combining datasets,  $R^2$  performance rating was enhanced to 0.53. The simulated graph using the cross-combined datasets (SC-4 and SC-5) in Figure 5-19 is well fitted to the observed streamflow pattern, with a good range of low flow and high flow. The individual dataset, however, was out of range and irregular.

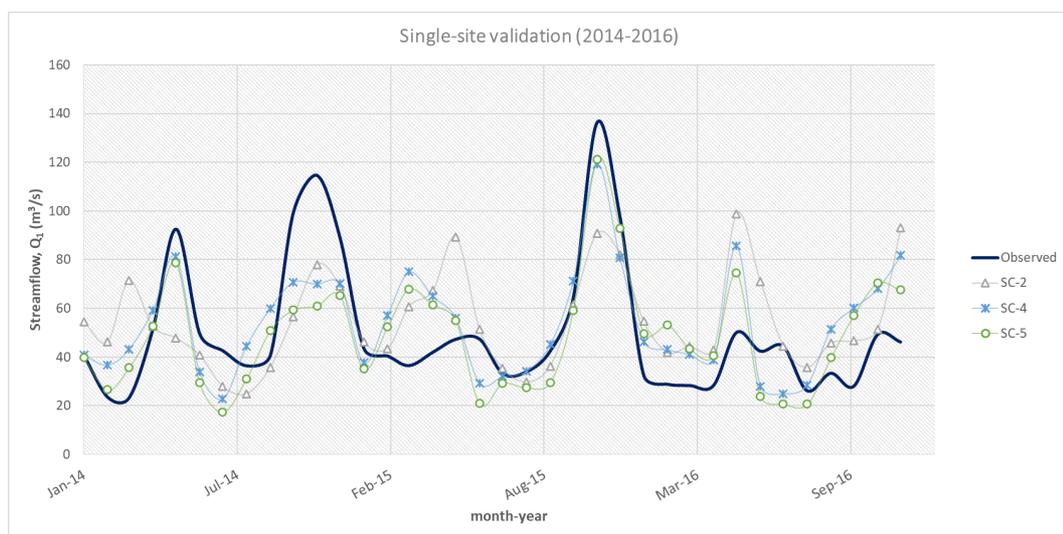


Figure 5-19: Simulated streamflow on CMADS individual and cross-combined datasets

Overall, referred to as the objective function of  $R^2$ , NSE and KGE, the validated results from the six scenarios reveal a consistent streamflow pattern, except for SC-2, which differs marginally. The most significant variation in simulated flow pattern between validated and measured data occurred in 2014, when the dry season led the dam level to fall below critical levels, resulting in a water crisis in Selangor (Anang et al., 2017). As a result of dam/WTP operation in the catchment, the SSC model is unable to effectively depict the flow pattern, and MSC (Multi Site Calibration) is a good approach to describe the catchment's streamflow condition.

### **5.3 Summary of the results**

Hydro-meteorological data plays a key role in determining the rainfall-runoff model to represent the watershed. As an important river basin, developing a rainfall-runoff model for SRB should be at the optimum level to raise the confidence level of the water resources manager. In order to compensate for the inadequate spatial and temporal sampling from gauged stations, global reanalysis datasets, CMADS, and CFSR provide reliable performance data. Point-to-point analysis was used to evaluate the global reanalysis datasets at the daily and monthly scales using data from 2008 to 2016. To ensure that the data was at the optimum level, several techniques were used, including the double mass curve method, error and missing data analysis, and bias. Evaluation of precipitation data from CMADS and CFSR shows both were reliable. However, CMADS performed better on temperature data compared to CFSR. The multiple-source datasets were applied in the rainfall-runoff model to enhance the model's results. For individual data analysis, gauged data outperformed global reanalysis datasets. However, for gridded datasets, CMADS performed better in the rainfall-runoff model compared to CFSR. The application of cross-combined datasets in rainfall-runoff models shows good performance ratings ( $R^2$ , NSE and KGE). Compared to the individual results, the cross-combined datasets performed better. Cross-combined datasets are an excellent approach to enhance the model's performance. The two SRB parameters that were most sensitive in the sensitivity analysis were CN2 and ALPHA\_BF. The curve number, or CN2, is used to calculate the amount of

runoff that will result from a given rainfall in a given watershed or piece of land. While the ALPHA\_BF is the baseflow alpha factor, it is utilised in modelling to determine how much groundwater contributes to streamflow. Based on elements including soil type, land use, and slope, the CN2 watershed was segmented into homogeneous areas. The runoff volume is then calculated using the total rainfall depth and the curve number values for the various locations. CN2 was assigned to each area based on its features. The ALPHA\_BF is frequently calculated by examining a stream's recession curve, which is the stream's slow decline in flow following a rainfall event. The recession curve, which depends on both baseflow and surface flow, can be used to calculate the rate at which water is evaporating from a stream. The alpha factor can be calculated by modelling the recession curve and splitting out the baseflow component.

## CHAPTER 6

### MULTI-SITE CALIBRATION ASSESSMENT

This chapter presents and discusses the calibration and validation methods applied to the SRB SWAT rainfall runoff model in order to determine the best fit model parameters and to enhance model performance. The single-site calibration (SSC) and the multi-site calibration (MSC) are two calibration methods that can be used to produce the best simulated streamflow that best suits the actual streamflow from the best calibrated parameters. SSC is commonly used in a lumped model such as an empirical or conceptual model where it is designed to simulate overall runoff and streamflow at the catchment's exit, rather than discrete flows (Sitterson et al., 2018). While MSC is a new approach in enhancing the calibration process, it helps to address the impact of human activities on water flow (Bai et al., 2017). According to Bai et al., (2017), Nkiaka et al., (2018), and Wang et al., (2012), MSC outperformed SSC in reaching a compromise in terms of model performance and reasonable parameter values. As a result, this study evaluated and compared those two methods to determine the optimal rainfall-runoff model for SRB. The streamflow stations were used as a calibration point by the SSC and MSC methods in the SRB, as illustrated in Figure 6-1. The SRB SWAT model evaluated the performance of the models on the seven scenario datasets described in the previous chapter using the Kling-Gupta Efficiency (KGE), Nash-Sutcliffe Efficiency (NSE), Coefficient of determination ( $R^2$ ), Percent bias (Pbias), and RMSE-observations standard deviation ratio (RSR).

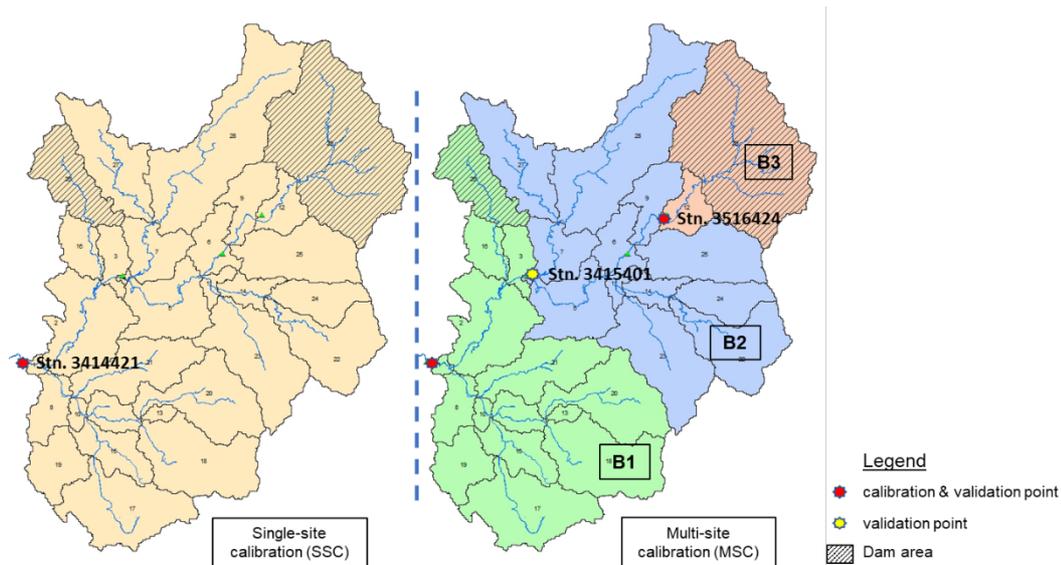


Figure 6-1: SRB calibration and validation method

## 6.1 Single-site calibration (SSC) method

The SSC method is used in the SWAT rainfall-runoff model to simulate a set of parameters described in Chapter 5 as a single unit for the whole catchment of the SRB. Using the SWAT-CUP software, the model was calibrated at the streamflow station 3414421 at the watershed's outlet. Eawag (2009) created SWAT-CUP, a semi-automatic calibration and uncertainty analysis tool, to expedite processing and offer improved calibration methods (Arnold, Moriasi, et al., 2012; Ozdemir & Leloglu, 2019). Parallel processing is used to run the 2000 iterations of the SRB SWAT model.

### 6.1.1 SSC model performance

The SSC results using 7 scenario datasets are demonstrated in Chapter 5, where most of the models achieve a 'good' range of performance rating, except for the model with input data from CFSR (SC-3). For an individual dataset, a model with gauged input data (SC-1) outperformed models CMADS (SC-2), and CFSR (SC-3), with all five performance ratings indicating the calibrated model was in a 'very good' range. While models using the cross-combine dataset (SC-4 to SC-7) attained a 'good' and a 'very good' range in performance ratings.

Most of the scenarios had comparable validated results except for SC-2, which had an objective function of  $R^2$ , and NSE recorded an '*unsatisfactory*' range. Due to inadequate data to run for the validation periods and the fact that the best calibrated result was outside of an acceptable range, SC-3 was not reviewed during the validation phase. Overall, the SSC validated results reveal a consistent streamflow pattern, except for SC-2, which differs marginally.

## **6.2 Multi-site calibration (MSC) method**

The multi-site calibration (MSC) method is a technique for calibrating a large watershed by dividing it into smaller watersheds and establishing parameters in each watershed (Bai et al., 2017). MSC was performed using three calibration techniques: basin-by-basin (BB), simultaneous (SM), and sequential (SE). The SWAT-CUP software is employed to run the calibration and validation processes. These techniques were applied to enhance and minimise uncertainty in the SRB rainfall-runoff model while defining the best calibrated parameters. Similar to the SSC method, the MSC method uses parallel processing and runs through 2000 iterations.

### **6.2.1 MSC model performance**

In this study, the MSC was analysed based on the best validated results from the SSC method and met the selection criteria in order to reduce uncertainties during model development. There are five elements of data: precipitation (P), temperature (T), relative humidity (R), solar radiation (S), and wind speed (W), that can be used as weather input data in the SWAT model. Incorporating more data elements reduces the uncertainty of model development. SC-1 used two elements of data, precipitation (P) and temperature (T) as input data. Therefore, SC-1 is less of a priority to be used in MSC analysis. SC-2 and SC-3 employed five data elements in the SWAT model, but they are not applicable since the calibrated and validated results were outside of the acceptable range. The SC-6 and SC-7 were also not applicable because they were merged with CFSR data, which was only available until 2014 due to data limitations.

The ideal option for MSC analysis is the SC-4 and SC-5, where they employed five elements' data and both validated models achieved a 'satisfactory' range. SC-4 uses temperature data from a gauged station where there is a missing data issue. Whereas SC-5 used temperature data from CMADS, where there is no missing data recorded. As a result, SC-5 was chosen to be employed in MSC analysis instead of SC-4 in order to reduce uncertainty during model development. A summary of the MSC selection criteria is shown in Table 6-1.

Table 6-1: Selection criteria for MSC analysis

| Scenario | Overall calibrated result | Overall validated result | Data availability |   |   |   |   |
|----------|---------------------------|--------------------------|-------------------|---|---|---|---|
|          |                           |                          | P                 | T | R | S | W |
| SC-1     | Very good                 | Satisfactory             | ■                 | ■ | - | - | - |
| SC-2     | Satisfactory              | Unsatisfactory           | ■                 | ■ | ■ | ■ | ■ |
| SC-3     | Unsatisfactory            | NA                       | ■                 | ■ | ■ | ■ | ■ |
| SC-4     | Very good                 | Satisfactory             | ■                 | ■ | ■ | ■ | ■ |
| SC-5     | Very good                 | Satisfactory             | ■                 | ■ | ■ | ■ | ■ |
| SC-6     | Very good                 | Satisfactory             | ■                 | ■ | ■ | ■ | ■ |
| SC-7     | Very good                 | Satisfactory             | ■                 | ■ | ■ | ■ | ■ |

- Available from 2008 to 2016
- Available from 2008 to 2014
- Incomplete data

The same data from SC-5 that had been calibrated by the SSC method was employed in order to analyse the performance of the MSC method. Three calibration techniques, basin-by-basin (BB), simultaneous (SM), and sequential (SE), are used in the MSC method to simulate the streamflow in the SWAT model. The reliable parameters for each watershed were identified and the output of the simulated model using these techniques was improved.

### *Basin-by-basin calibration*

Basin-by-basin calibration (BB) is a technique where the large watershed is divided into smaller watersheds based on the calibration and validation point. This technique is similar to the SSC method except the watershed size is different. Therefore, instead of lumping all the parameters into one large watershed as in the SSC approach, each watershed in the BB technique has its own set of parameters. In BB technique, SRB was divided into 3 basins which is shown in Figure 6-2. Basin-3 (B3) encompasses the upper site of SRB, which includes the Sg. Selangor dam, and basin-2 (B2) encompasses the middle area of SRB, which includes 13 sub-basins. Meanwhile, basin-1 (B1) refers to 14 sub-basins in the SRB downstream area. These clusters are also applied in sequential and simultaneous techniques. In the BB technique, basin-1 (B1) and 3 (B3) were typically simulated for both the calibration and validation processes, as illustrated in Figure 6-2. The parameters were established based on each individual basin and B1 and B3 were separately calibrated and validated. In contrast to sequential technique, BB technique can be applied using either a top-down or bottom-up approach because they are unrelated to one another. However, basin-2 (B2) utilised the regionalisation technique in order to determine the appropriate parameters that best matched the B2 condition due to the scarcity of streamflow data. In addition, B2 was calibrated throughout the same validation period to ensure model consistency.

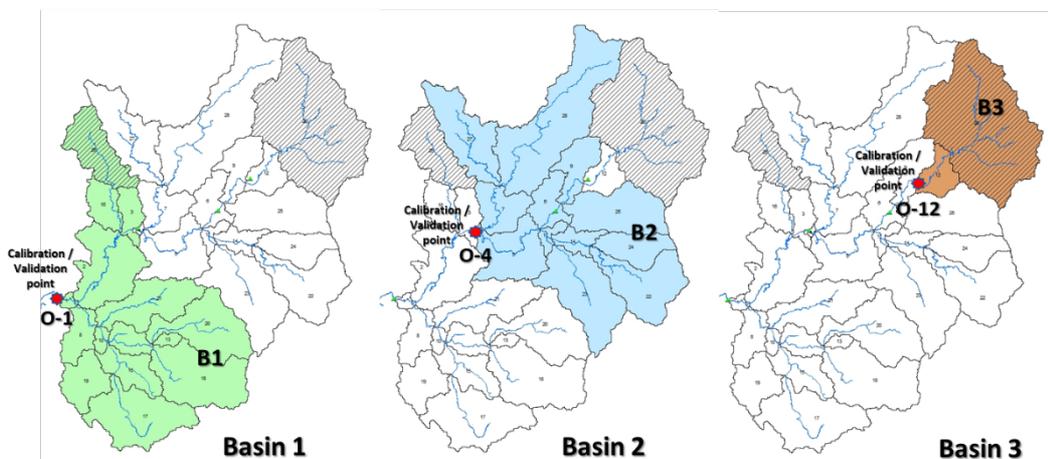


Figure 6-2: multi-site calibration for basin-by-basin technique

*Regionalisation parameter*

Regionalization is the process of transferring characteristics from a nearby basin to the basin of interest, particularly to the basin lacking streamflow data. In ungauged catchments, this method was utilised to decrease uncertainty and boost forecast confidence (Hrachowitz et al., 2013). The presence of similarity indicators, such as spatial proximity and catchment parameters, such as land use, soil type, and topographic features, in a basin is necessary for this method to be effective (Merz & Blöschl, 2004). The regionalisation technique from basin-1 (BB:B2-R1) and 3 (BB:B2-R3) was applied in basin-2 because both basins had similar types of soil and land use.

Table 6-2: multi-site calibrated results using the basin-by-basin technique.

| Model                  |          | Calibration    |       |       |      |      |
|------------------------|----------|----------------|-------|-------|------|------|
|                        |          | R <sup>2</sup> | NSE   | PBias | KGE  | RSR  |
| basin-by-basin<br>(BB) | BB:B1    | 0.71           | 0.64  | 0.00  | 0.82 | 0.60 |
|                        | BB:B3    | 0.17           | -0.08 | 0.00  | 0.40 | 1.04 |
|                        | BB:B2    | 0.36           | 0.25  | -0.10 | 0.58 | 0.86 |
|                        | BB:B2-R1 |                |       |       |      |      |
|                        | BB:B2-R3 |                |       |       |      |      |

Table 6-3: multi-site validated results using the basin-by-basin technique.

| Model                  |          | Validation     |       |       |      |      |
|------------------------|----------|----------------|-------|-------|------|------|
|                        |          | R <sup>2</sup> | NSE   | PBias | KGE  | RSR  |
| basin-by-basin<br>(BB) | BB:B1    | 0.36           | 0.19  | 0.10  | 0.59 | 0.90 |
|                        | BB:B3    | 0.07           | -2.17 | 0.00  | 0.17 | 1.78 |
|                        | BB:B2    |                |       |       |      |      |
|                        | BB:B2-R1 | 0.29           | 0.22  | 0.00  | 0.51 | 0.88 |
|                        | BB:B2-R3 | 0.09           | 0.05  | 6.30  | 0.13 | 0.97 |

Table 6-2 and Table 6-3 demonstrate the calibrated and validated SWAT model results using the basin-by-basin technique. Overall, the calibrated result of the BB technique in basin-1 (BB:B1) indicates an acceptable range in all performance ratings, as illustrated in Figure 6-3. However, BB:B1 did not perform well during the validation process, with three performance ratings ( $R^2$ , NSE, and RSR) falling within an inadequate range. The performance ratings of KGE and Pbias indicated the 'satisfactory' and 'very good' range, respectively. From the simulated results, the BB:B1 underestimated the SRB low flow but significantly reflected the high flow during the calibration and validation processes. According to the result, from June 2014 to June 2015 and from June 2016 to December 2016 are the critical periods where the simulated streamflow fails to match the observed streamflow. These periods are also recorded in SSC method results as a critical simulated period.

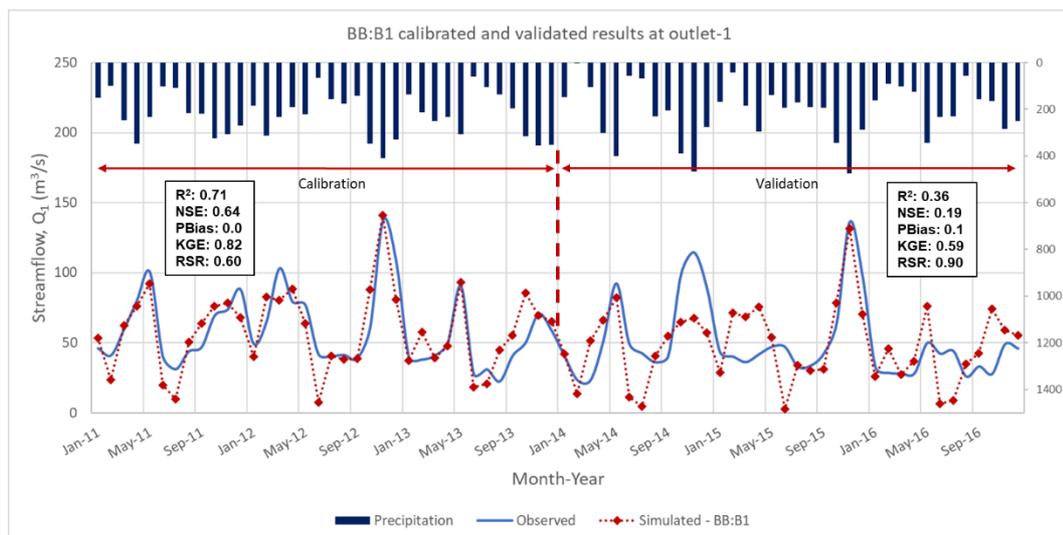


Figure 6-3: Calibrated and validated results using basin-by-basin technique in basin-1 (BB:B1)

The simulated outcomes in Figure 6-4 show that the model struggled throughout the calibration and validation processes in B3. The calibrated model had trouble matching the streamflow data, and the best fit was obtained with 0.17 of  $R^2$ , -0.80 of NSE, 0.0 of Pbias, 0.40 of KGE, and 1.04 of RSR. Except for Pbias, all outcomes for the objection functions were rated as 'unsatisfactory' performance. The validated outcomes were likewise in the

'unsatisfactory' range, with  $R^2$  dropping to 0.07, NSE dropping to -2.17, KGE dropping to 0.17, and RSR rising to 1.78. In both the calibrated and validated results, Pbias was the only objective function that remained in the 'very good' range with a value of 0.0.

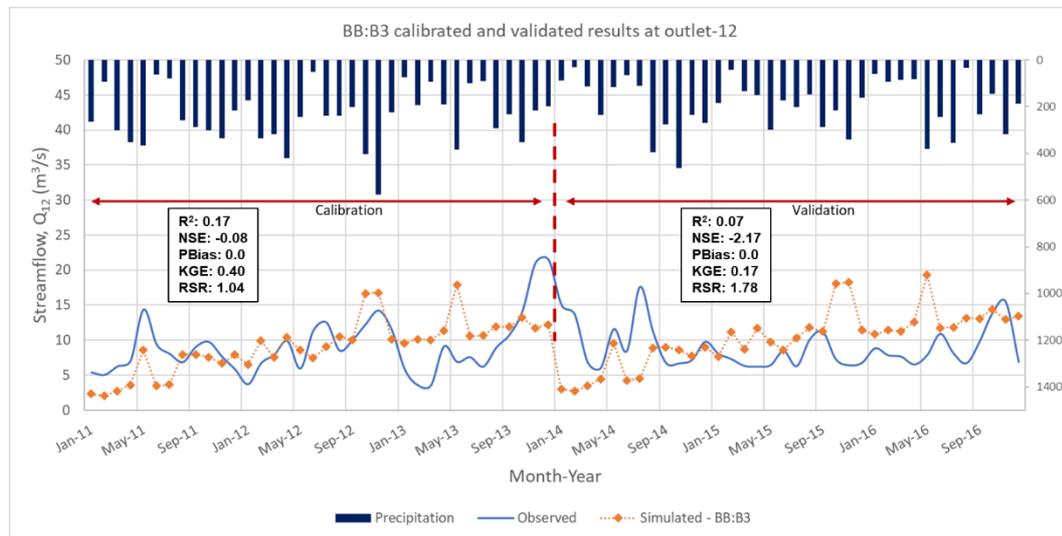


Figure 6-4: Calibrated and validated results using basin-by-basin technique in basin-3 (BB:B3)

The calibrated result for basin-2 (BB:B2) reveals inconsistent performance ratings, with  $R^2$ , NSE, and RSR indicating an 'unsatisfactory' range, and Pbias and KGE indicating 'good' and 'satisfactory' ranges, respectively. As shown in Figure 6-5, the calibrated model generally represented both high and low flow in B2 and struggled to match observed data from June 2014 to June 2015 and May to December 2016.

According to the regionalisation technique from B1 (BB:B2-R1), Table 6-3 demonstrates that the  $R^2$ , NSE, and RSR were in the 'unsatisfactory' range, with values of 0.29, 0.22, and 0.88, respectively. Pbias and KGE have scores that are, respectively, 'good' and 'satisfactory' at 0.0 and 0.51. The regionalisation technique from B3 (BB:B2-R3) shows an 'unsatisfactory' range for most of the performance ratings, except Pbias. The performance scores for this technique are  $R^2 = 0.09$ , NSE = 0.05, KGE = 0.13, and RSR = 0.97.

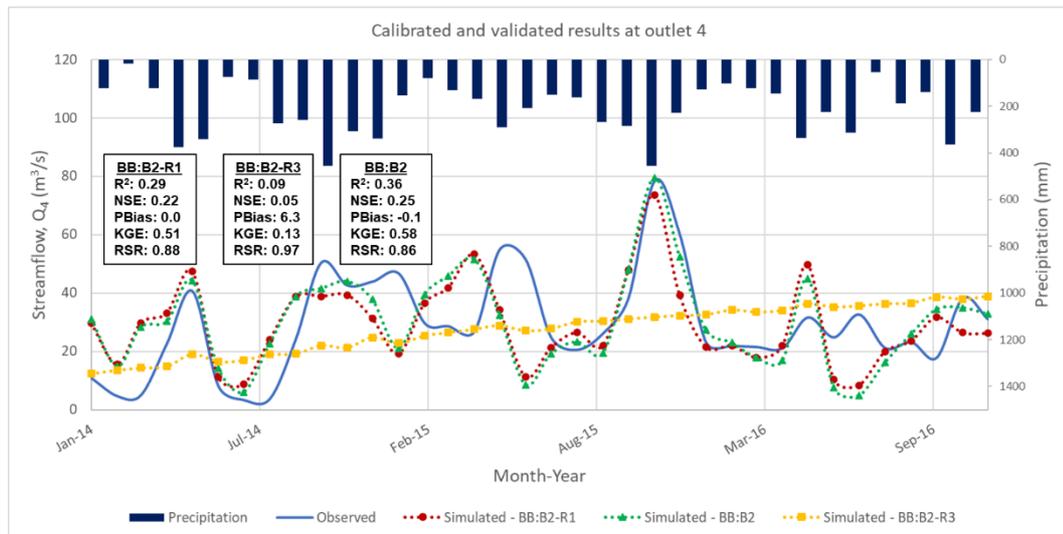


Figure 6-5: Calibrated and validated results using basin-by-basin technique in basin-3

In the regionalisation method, BB:B2-R1 outperformed BB:B2-R3 in all performance ratings, even though many of them were not within an acceptable range. According to Figure 6-5, the simulated graph of BB:B2-R1 is more akin to the observed graph compared to BB:B2-R3. The best parameters in B3 are not applicable for B2 as it does not respond to simulated streamflow at outlet-4. In fact, the best calibrated model for basin-2 (BB:B2) resembled the BB:B2-R1 simulated graph. Additionally, the parameters shown in Table 6-4 demonstrate that the regionalisation parameters from B1 are mostly within the same range as the calibrated parameters in B2. As a result, compared to BB:B2-R3, the regionalisation technique from basin-1 (BB:B2-R1) is more relevant in reflecting the conditions in basin-2 (B2).

Table 6-4: Fitted parameters in calibrated model basin-2 and regionalisation method from basin-1 and basin-3

| Parameter           | BB:B2  | BB:B2-R1 | BB:B2-R3 |
|---------------------|--------|----------|----------|
| 1:R_CN2.mgt         | -0.75  | -0.53    | 0.15     |
| 2:R_SOL_AWC(..).sol | 0.21   | 0.58     | 0.91     |
| 3:V_ESCO.hru        | 0.58   | 0.02     | 0.12     |
| 4:V_GW_DELAY.gw     | 371.16 | 343.17   | 696.38   |
| 5:V_GWQMN.gw        | 415.16 | 3355.00  | 2291.46  |
| 6:V_GW_REVAP.gw     | 0.06   | 0.15     | 0.25     |
| 7:R_HRU_SLP.hru     | -0.23  | 0.41     | 0.36     |
| 8:R_OV_N.hru        | 25.20  | 30.86    | 6.29     |
| 9:R_SLSUBBSN.hru    | 43.68  | 67.42    | 125.71   |
| 10:R_SOL_K(..).sol  | 510.49 | 1802.11  | 799.66   |
| 11:R_SOL_BD(..).sol | 1.90   | 1.45     | 0.85     |
| 12:V_CANMX.hru      | 92.47  | 53.28    | 48.02    |
| 13:V_CH_K2.rte      | 28.05  | 26.51    | 652.11   |
| 14:V_ALPHA_BF.gw    | -0.42  | -1.28    | 0.00     |
| 15:V_CH_N2.rte      | 0.00   | 0.00     | 0.00     |
| 16:V_LAT_TTIME.hru  | 88.42  | 240.25   | 101.68   |
| 17:V_RCHRG_DP.gw    | 0.00   | 0.00     | 0.01     |
| 18:V_SURLAG.bsn     | 11.76  | 7.20     | 48.79    |
| 19:V_EPCO.hru       | 1.03   | 1.18     | 1.11     |
| 20:V_REVAPMN.gw     | 168.93 | 32.84    | -204.05  |

#### *Simultaneous calibration*

The second multi-site calibration method used in this research is the simultaneous calibration (SM) approach. As illustrated in Figure 6-6, this technique was utilised to calibrate all basins simultaneously. There were three different simultaneous calibration techniques. The first technique combined all parameters from the three different basins into one basin (SM:B1B2B3). In contrast to the SSC method, this methodology included many calibration and validation points. The second technique combines basin-1 and 2 parameters and simultaneously performs the calibration process with an independent basin-3 (SM:B1B2+B3). The third technique combines basin-2 and 3 parameters and simultaneously runs the calibration process with basin-1 (SM:B1+B2B3). Two tests have been conducted as part of the validation process; the first test uses the exempted outlet-4 (T1) in the SWAT model as a validation point, while the second test is the opposite (T2). This test is performed to determine whether any significant impacts were discovered during the calibration and validation procedures and to increase confidence in the parameters utilised.

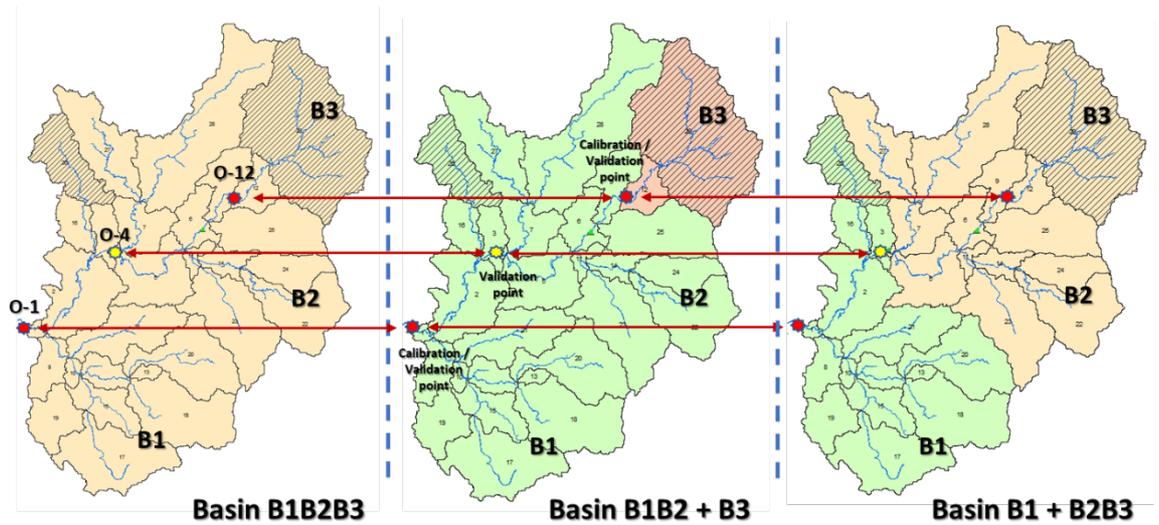


Figure 6-6: multi-site calibration for simultaneous technique

Table 6-5: multi-site calibrated results using the simultaneous technique.

| Model             |                        | Calibration    |       |        |      |      |
|-------------------|------------------------|----------------|-------|--------|------|------|
|                   |                        | R <sup>2</sup> | NSE   | PBias  | KGE  | RSR  |
| Simultaneous (SM) | SM:B1B2B3-T1 (O-12)    | 0.06           | -1.73 | -12.10 | 0.11 | 1.65 |
|                   | SM:B1B2B3-T1 (O-1)     | 0.85           | 0.37  | 12.10  | 0.60 | 0.79 |
|                   | SM:B1B2B3-T1 (O-4)     |                |       |        |      |      |
|                   | SM:B1B2B3-T2 (O-12)    | 0.06           | -1.73 | -12.10 | 0.11 | 1.65 |
|                   | SM:B1B2B3-T2 (O-1)     | 0.85           | 0.37  | 12.10  | 0.60 | 0.79 |
|                   | SM:B1B2B3-T2 (O-4)     |                |       |        |      |      |
|                   | SM:B1B2 + B3-T1 (O-12) | 0.43           | -0.32 | -11.80 | 0.37 | 1.15 |
|                   | SM:B1B2 + B3-T1 (O-1)  | 0.49           | 0.34  | 8.50   | 0.67 | 0.82 |
|                   | SM:B1B2 + B3-T1 (O-4)  |                |       |        |      |      |
|                   | SM:B1B2 + B3-T2 (O-12) | 0.43           | -0.32 | -11.80 | 0.37 | 1.15 |
|                   | SM:B1B2 + B3-T2 (O-1)  | 0.49           | 0.34  | 8.50   | 0.67 | 0.82 |
|                   | SM:B1B2 + B3-T2 (O-4)  |                |       |        |      |      |
|                   | SM:B1+B2B3-T1 (O-12)   | 0.06           | -2.37 | 14.40  | 0.05 | 1.84 |
|                   | SM:B1+B2B3-T1 (O-1)    | 0.86           | 0.27  | 24.40  | 0.58 | 0.85 |
|                   | SM:B1+B2B3-T1 (O-4)    |                |       |        |      |      |
|                   | SM:B1+B2B3-T2 (O-12)   | 0.06           | -2.37 | 14.40  | 0.05 | 1.84 |
|                   | SM:B1+B2B3-T2 (O-1)    | 0.86           | 0.27  | 24.40  | 0.58 | 0.85 |
|                   | SM:B1+B2B3-T2 (O-4)    |                |       |        |      |      |

The calibrated findings of the simultaneous approach are given in Table 6-5. Overall, test-1 and test-2 findings are identical at all outlets and for all calibrating techniques. At the outlet-1, simultaneous techniques from SM:B1B1B3 and SM:B1+B2B3 outperformed simultaneous techniques from SM:B1B2+B3, with three of the five performance ratings ( $R^2$ , Pbias and KGE) reaching the acceptable range. The  $R^2$ , Pbias, and KGE of SM:B1B1B3 were in 'very good', 'good', and 'satisfactory' ranges, while SM:B1+B2B3 were in 'very good' and 'satisfactory' ranges. Meanwhile, SM:B1B2+B3 achieved 'very good' ( $8.5 < \pm 10$ ) and 'satisfactory' range ( $0.5 < 0.67 < 0.75$ ) for Pbias and KGE performance, respectively. Figure 6-7 demonstrates that all the simulated graphs reflected the observed, but slightly lower than the observed, particularly from January 2011 until June 2012.

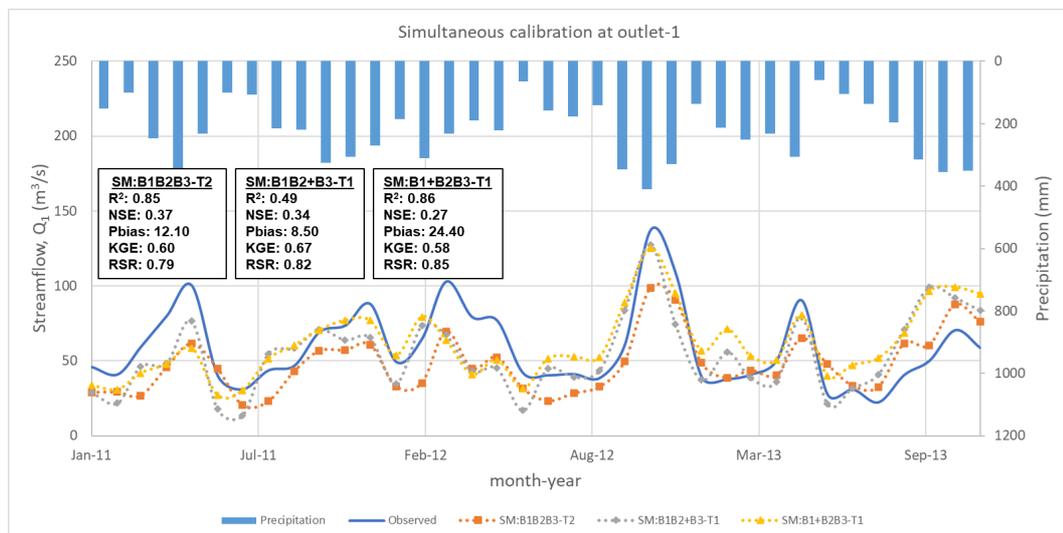


Figure 6-7: simulated graph on simultaneous techniques at outlet-1

SM:B1B2+B3 performs better than SM:B1B1B3 and SM:B1+B2B3 at outlet-12 despite poor performance ratings from most objective functions, with  $R^2 = 0.43$ , NSE = -0.32, Pbias = -11.8, KGE = 0.37, and RSR = 1.15. The SM:B1+B2B3 has the lowest  $R^2 = 0.06$ , NSE = -2.37, Pbias = 14.4, KGE = 0.05, and RSR = 1.84 values of performance ratings. Figure 6-8 demonstrates that, after May 2011, models no longer accurately represent observed data, with most simulated graphs being overestimated.

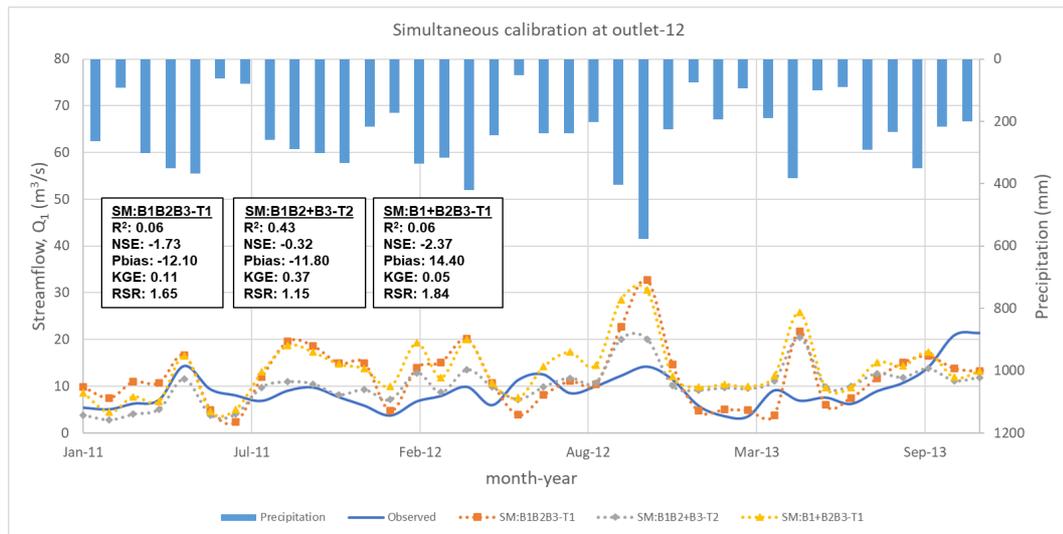


Figure 6-8: simulated graph on simultaneous techniques at outlet-12

Outlet-4 was solely utilised for the validation process of the application of the regionalisation method in basin-2 since there will not be any streamflow data available until 2014.

Table 6-6: multi-site validated results using the simultaneous technique.

| Model             |                        | Validation     |       |        |       |      |
|-------------------|------------------------|----------------|-------|--------|-------|------|
|                   |                        | R <sup>2</sup> | NSE   | PBias  | KGE   | RSR  |
| Simultaneous (SM) | SM:B1B2B3-T1 (O-12)    | 0.09           | -3.88 | 9.50   | -0.97 | 2.40 |
|                   | SM:B1B2B3-T1 (O-1)     | 0.52           | -0.68 | -8.90  | 0.70  | 0.92 |
|                   | SM:B1B2B3-T1 (O-4)     |                |       |        |       |      |
|                   | SM:B1B2B3-T2 (O-12)    | 0.09           | -4.74 | 12.70  | -0.97 | 2.40 |
|                   | SM:B1B2B3-T2 (O-1)     | 0.52           | 0.15  | 2.10   | 0.70  | 0.92 |
|                   | SM:B1B2B3-T2 (O-4)     | 0.47           | 0.19  | -14.50 | 0.59  | 0.90 |
|                   | SM:B1B2 + B3-T1 (O-12) | 0.13           | -2.91 | 1.60   | -1.05 | 1.98 |
|                   | SM:B1B2 + B3-T1 (O-1)  | 0.40           | 0.16  | 0.20   | 0.56  | 0.92 |
|                   | SM:B1B2 + B3-T1 (O-4)  |                |       |        |       |      |
|                   | SM:B1B2 + B3-T2 (O-12) | 0.13           | -2.91 | 13.90  | 0.08  | 1.98 |
|                   | SM:B1B2 + B3-T2 (O-1)  | 0.40           | 0.16  | -5.60  | 0.54  | 0.92 |
|                   | SM:B1B2 + B3-T2 (O-4)  | 0.35           | 0.10  | -10.20 | 0.50  | 0.95 |
|                   | SM:B1+B2B3-T1 (O-12)   | 0.07           | -6.98 | 12.80  | -0.50 | 2.84 |
|                   | SM:B1+B2B3-T1 (O-1)    | 0.58           | 0.22  | 10.70  | 0.58  | 0.86 |
|                   | SM:B1+B2B3-T1 (O-4)    |                |       |        |       |      |
|                   | SM:B1+B2B3-T2 (O-12)   | 0.07           | -7.77 | -51.70 | -1.27 | 2.96 |
|                   | SM:B1+B2B3-T2 (O-1)    | 0.58           | 0.01  | -22.30 | 0.54  | 0.90 |
|                   | SM:B1+B2B3-T2 (O-4)    | 0.53           | 0.18  | 7.10   | 0.11  | 0.99 |

Table 6-6 demonstrates the validated results of all simultaneous techniques with some differences between test-1 and 2 at outlet-1 and 12. However, this difference has no impact on the model's performance rating. At outlet-1, SM:B1B2B3-T2 improved NSE and Pbias performance rating values from SM:B1B2B3-T1. The NSE and Pbias of SM:B1B2B3-T2 increased to 0.15 from -0.68 and 2.10 from -8.90, respectively. However, various performance ratings values, such as NSE, Pbias, KGE, and RSR in SM:B1+B2B3-T2 and SM:B1B2+B3-T2, were lower in SM:B1+B2B3-T1 and SM:B1B2+B3-T1. Overall, SM:B1B2B3-T2 outperforms other simultaneous techniques, achieving  $R^2 = 0.52$ , NSE = 0.15, Pbias = 2.10, KGE = 0.70, and RSR = 0.92.  $R^2$ , Pbias, and KGE, three objective functions, were in 'satisfactory', 'very good', and 'satisfactory', respectively. The performance rating from SM:B1B2+B3-T2 at outlet-1, on the other hand, is the lowest, with  $R^2 = 0.40$ , NSE = 0.16, Pbias = -5.60, KGE = 0.54, and RSR = 0.92. Figure 6-9 displays the best verified streamflow for outlet-1 using all simultaneous techniques. The simulated graph effectively represented the observed data; however, it underestimated the low flow that occurs in June every year.

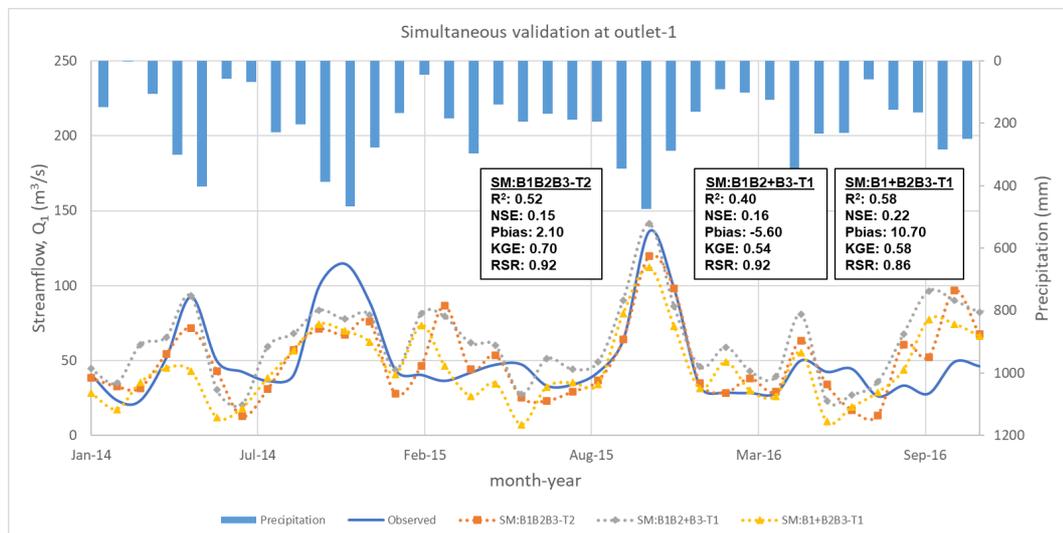


Figure 6-9: the best validated streamflow results from each simultaneous technique at outlet-1

Overall, none of the simultaneous techniques used at outlet-12 achieved performance ratings for  $R^2$ , NSE, KGE and RSR that were within an acceptable range, except Pbias. Despite having poor performance ratings, the SM:B1B2+B3 technique outperformed the SM:B1B2B3 and SM:B1+B2B3 techniques in terms of performance ratings, with  $R^2 = 0.13$ , NSE = -2.91, Pbias = 13.90, KGE = 0.08, and RSR = 1.98. The validated results in Table 6-6 demonstrate that test-2 was mostly used to lower performance rating values for NSE, Pbias, KGE, and RSR. However, the SM:B1B2+B3-T2 increased KGE's performance rating from -1.05 to 0.08. Figure 6-10 illustrates the simultaneous technique applied at outlet-12 with all techniques barely reflecting the observed data.

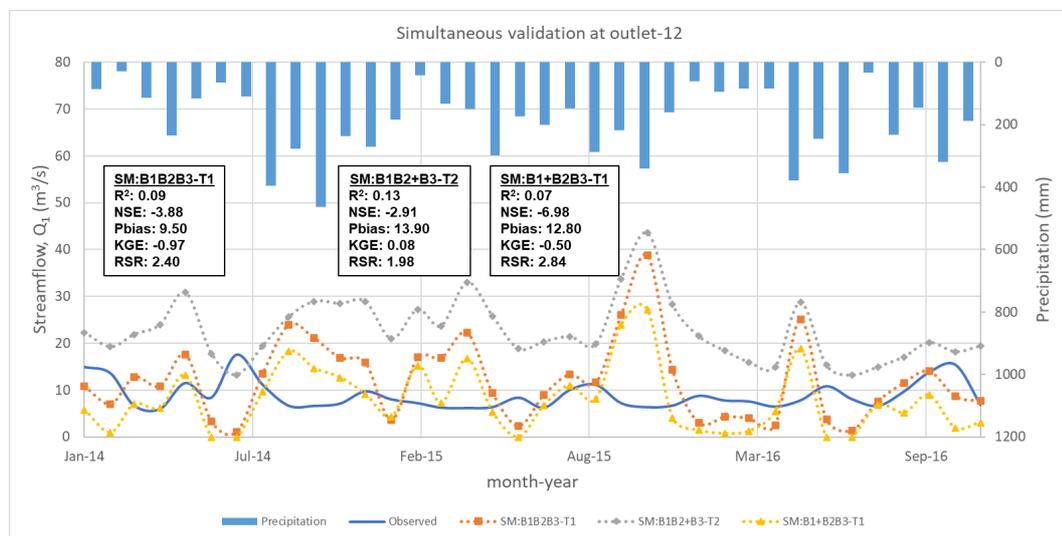


Figure 6-10: validated results for simultaneous techniques at outlet-12

The basin-by-basin technique used parameters that were calibrated from a neighbouring basin in basin-2 (B2) and verified at the outlet-4. In this technique, basin-2 was merged with either basin-1 (B1) or basin-3 (B3) to obtain the optimal parameters while doing the calibration procedure simultaneously. According to Table 6-6, all simultaneous techniques achieved two out of five acceptable ranges of performance ratings. The SM:B1B2B3-T2 and SM:B1B2+B3-T2 techniques achieved a 'satisfactory' range on KGE, while the SM:B1+B2B3-T2 achieved a 'satisfactory' range on  $R^2$ . All simultaneous techniques attained the 'good' range of Pbias.

Figure 6-11 illustrates how all simultaneous techniques substantially mirrored the observed data at outlet-4. However, SM:B1B2B3-T2 and SM:B1B2+B3-T2 are the techniques that best match the streamflow that was seen. This is in contrast to SM:B1+B2B3-T2, where both high and low flow were successfully captured. Meanwhile, the SM:B1+B2B3 simulated graph was slightly higher than the observed graph. According to the validated results in Table 6-6 and the simulated graph in Figure 6-11, the SM:B1B2B3-T2 is the best approach for outlet-4 since it outperformed the others in all three objective functions (NSE, KGE, and RSR).

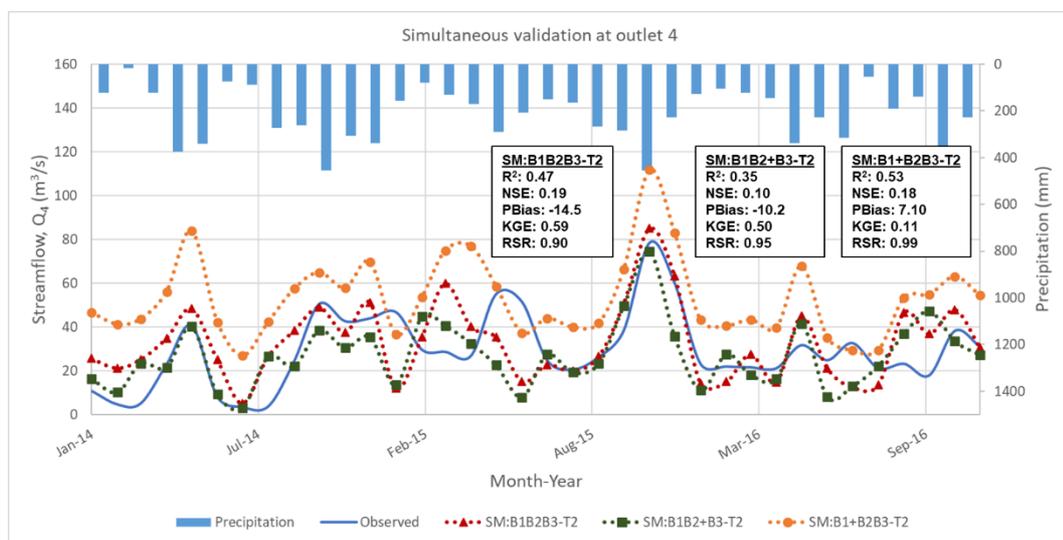


Figure 6-11: validated result of simultaneous technique at outlet-4

### *Sequential calibration*

The sequential technique is the third multi-site calibration approach used in this study, after the basin-by-basin and simultaneous methods. This technique adhered to the top-to-bottom principles, calibrating the upstream parameters in basin-3 (B3) first, followed by the downstream parameters in basin-2 (B2) and basin-1 (B1). Due to the restricted streamflow data and only being significant during the validation phase, basin-2 (B2) was similarly excused from the calibration process like other calibration procedures. The parameters for B2 were determined by combination with B1 or B3 in the calibration process.

Figure 6-12 illustrates the sequential calibration approach using two different techniques. The first sequential technique (SE:B1B2+B3) combines basin-1 (B1) and 2 (B2), while the second sequential technique (SE:B1+B2B3) combines basin-2 (B2) and basin-3 (B3). Two tests have been conducted, with the first validation test excluding outlet-4 (T1) in the calibration process and the second test including outlet-4 (T2).

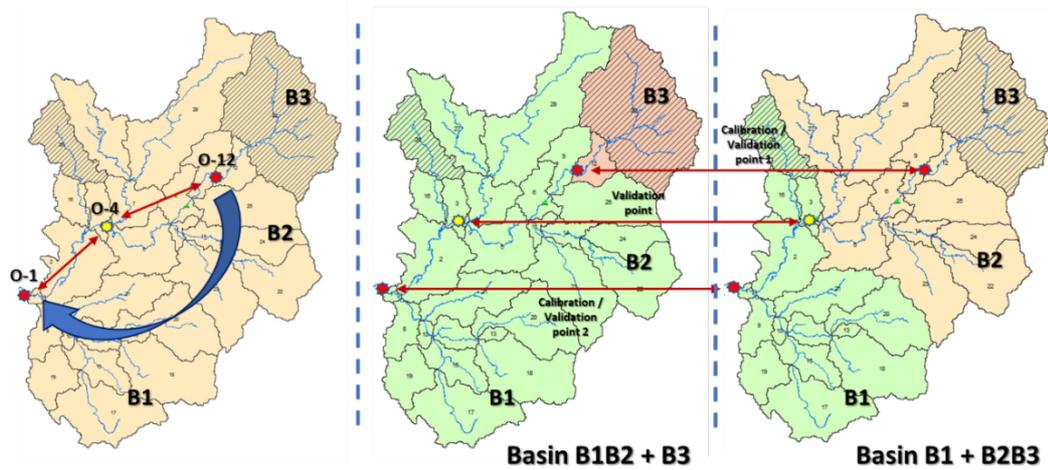


Figure 6-12: multi-site calibration approach in sequential method

Table 6-7: multi-site calibrated result using the sequential technique.

| Model           |                        | Calibration    |       |       |      |      |
|-----------------|------------------------|----------------|-------|-------|------|------|
|                 |                        | R <sup>2</sup> | NSE   | PBias | KGE  | RSR  |
| Sequential (SE) | SE:B1B2 + B3-T1 (O-12) | 0.17           | -0.07 | 0.50  | 0.39 | 1.03 |
|                 | SE:B1B2 + B3-T1 (O-1)  | 0.72           | 0.65  | -0.50 | 0.83 | 0.59 |
|                 | SE:B1B2 + B3-T1 (O-4)  |                |       |       |      |      |
|                 | SE:B1B2 + B3-T2 (O-12) | 0.17           | -0.07 | 0.50  | 0.39 | 1.03 |
|                 | SE:B1B2 + B3-T2 (O-1)  | 0.72           | 0.65  | -0.50 | 0.83 | 0.59 |
|                 | SE:B1B2 + B3-T2 (O-4)  |                |       |       |      |      |
|                 | SE:B1 + B2B3-T1 (O-12) | 0.18           | -0.07 | 2.10  | 0.40 | 1.03 |
|                 | SE:B1 + B2B3-T1 (O-1)  | 0.57           | 0.48  | -2.10 | 0.66 | 0.72 |
|                 | SE:B1 + B2B3-T1 (O-4)  |                |       |       |      |      |
|                 | SE:B1 + B2B3-T2 (O-12) | 0.18           | -0.07 | 2.10  | 0.40 | 1.03 |
|                 | SE:B1 + B2B3-T2 (O-1)  | 0.57           | 0.48  | -2.10 | 0.66 | 0.72 |
|                 | SE:B1 + B2B3-T2 (O-4)  |                |       |       |      |      |

Generally, test-1 (T1) and test-2 (T2) reveal a similar result to all sequential techniques, as shown in Table 6-7. At outlet-12, calibrated parameters in basin-3 (B3) have difficulty achieving at least a ‘satisfactory’ range, except for the Pbias. Both sequential techniques, SE:B1B2+B3 and SE:B1+B2B3 failed to reflect the actual streamflow condition, as illustrated in Figure 6-13. The calibration process for both techniques was continued at outlet-1, by assuming the calibrated results at outlet-12 are the optimum results representing basin-3 (B3).

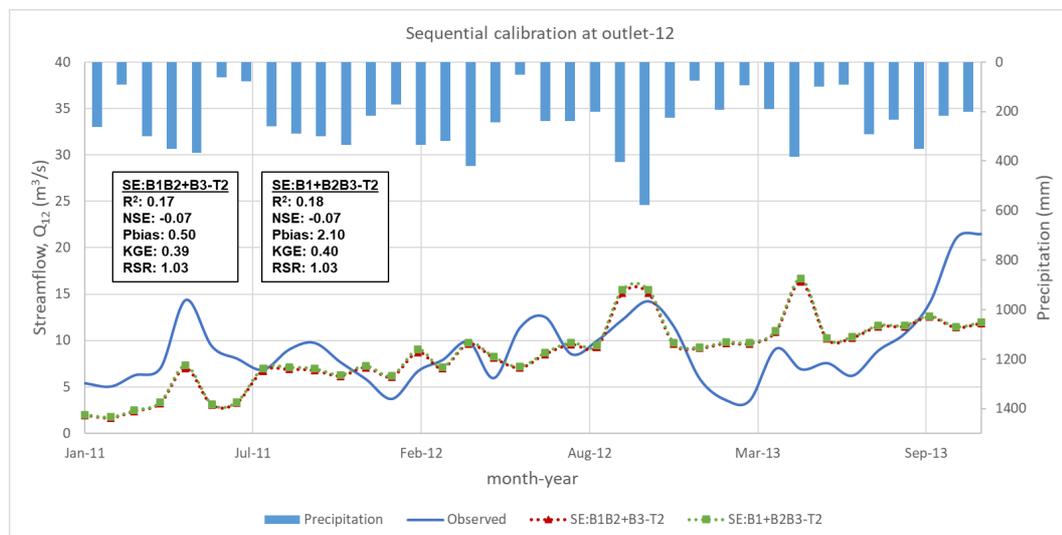


Figure 6-13: calibrated results for sequential techniques at outlet-12

Meanwhile at outlet-1, the SE:B1B2+B3 outperformed the SE:B1+B2B3 by obtaining 0.72 in  $R^2$  (good), 0.65 in NSE (good), -0.50 in Pbias (very good), 0.83 in KGE (good), and 0.59 in RSR (good). SE:B1+B2B3 received a  $R^2$  scored of 0.57 (satisfactory), a NSE score of 0.48 (unsatisfactory), a Pbias score of -2.10 (very good), a KGE score of 0.66 (satisfactory), and a RSR scored of 0.72 (unsatisfactory). In Figure 6-16, simulated graphs from SE:B1B2+B3 and SE:B1+B2B3 to observed data are shown. Both simulated graphs precisely captured the observed streamflow pattern. Overall, SE:B1B2+B3 properly identified the circumstances of high and low flow, whereas SE:B1+B2B3 accurately captured the low flow condition but slightly lower than the high flow.

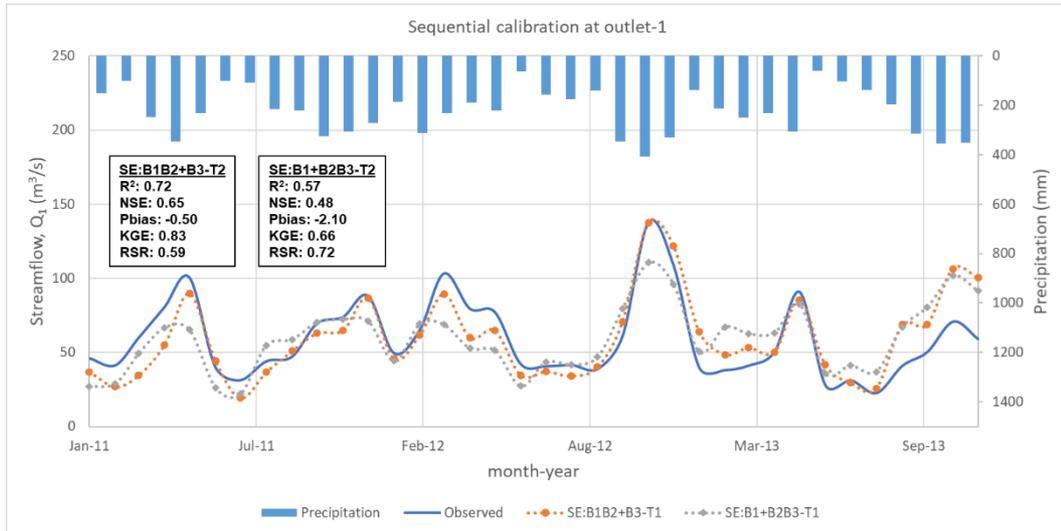


Figure 6-14: calibrated results for sequential techniques at outlet-1

Table 6-8: multi-site validated result for sequential technique.

| Model           |                        | Validation     |       |        |      |      |
|-----------------|------------------------|----------------|-------|--------|------|------|
|                 |                        | R <sup>2</sup> | NSE   | PBias  | KGE  | RSR  |
| Sequential (SE) | SE:B1B2 + B3-T1 (O-12) | 0.05           | -2.09 | 11.70  | 0.03 | 1.76 |
|                 | SE:B1B2 + B3-T1 (O-1)  | 0.46           | 0.27  | -11.70 | 0.65 | 0.86 |
|                 | SE:B1B2 + B3-T1 (O-4)  |                |       |        |      |      |
|                 | SE:B1B2 + B3-T2 (O-12) | 0.08           | -2.12 | 40.00  | 0.03 | 1.77 |
|                 | SE:B1B2 + B3-T2 (O-1)  | 0.45           | 0.21  | -16.40 | 0.65 | 0.89 |
|                 | SE:B1B2 + B3-T2 (O-4)  | 0.48           | 0.34  | -23.60 | 0.63 | 0.81 |
|                 | SE:B1 + B2B3-T1 (O-12) | 0.06           | -2.09 | 13.80  | 0.12 | 1.79 |
|                 | SE:B1 + B2B3-T1 (O-1)  | 0.37           | 0.19  | -13.80 | 0.53 | 0.86 |
|                 | SE:B1 + B2B3-T1 (O-4)  |                |       |        |      |      |
|                 | SE:B1 + B2B3-T2 (O-12) | 0.06           | -2.28 | -12.20 | 0.15 | 1.81 |
|                 | SE:B1 + B2B3-T2 (O-1)  | 0.37           | 0.15  | -27.00 | 0.39 | 0.92 |
|                 | SE:B1 + B2B3-T2 (O-4)  | 0.04           | -0.61 | 39.50  | 0.01 | 1.27 |

Test-1 (T1) and test-2 (T2) have a small variance, according to validated results in Table 6-8. Nevertheless, the rating of overall performance remains unchanged. In most objective functions, T1 performed better than T2 based on performance rating value.

Generally, the majority of the objective functions, with the exception of Pbias, have 'unsatisfactory' performance ratings according to the validated results at outlet-12. The performance rating values for SE:B1+B2B3-T1 were greater than those of the other models in certain objective functions, with  $R^2 = 0.06$ , NSE = -2.09, Pbias = 13.80, KGE = 0.12, and RSR = 1.79. Figure 6-15 illustrates how the calibrated parameters failed to accurately represent the streamflow pattern during validation. It was discovered that the simulated results produced an identical streamflow pattern at outlet-12. A minimum flow was initially modelled at outlet-12 from Jan to June 2014, where it may be impacted by the Sg. Selangor Dam operation. During this time, basin-3 received less amount of rainfall, but the streamflow is high. As mentioned in Chapter 3, about 90% of this area covered by dam area.

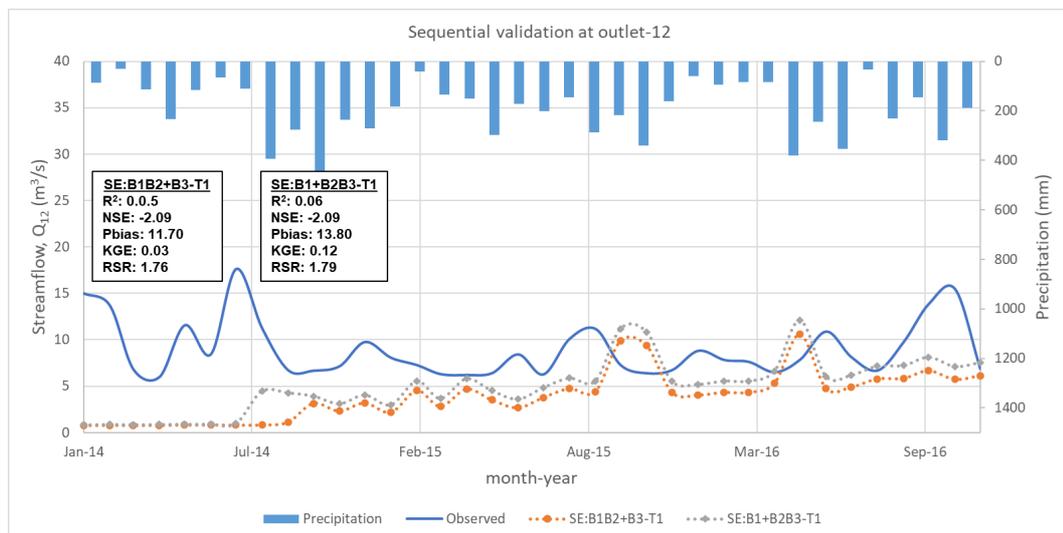


Figure 6-15: validated result of sequential technique at outlet-12

The best validated result at outlet-1 was SE:B1B2+B3-T1, with  $R^2 = 0.46$ , NSE = 0.27, Pbias = -11.7, KGE = 0.65, and RSR = 0.86. Meanwhile, the SE:B1+B2B3-T1 is slightly lower compared to SE:B1B2+B3-T1 with  $R^2 = 0.37$ , NSE = 0.19, Pbias = -13.80, KGE = 0.53, and RSR = 0.86. The simulated graph in Figure 6-16 demonstrates how the sequential techniques employed at outlet-1 significantly captured the observed flow pattern. Both the high and low flows were mirrored by the SE:B1B2+B3-T1, with SE:B1+B2B3-T1 being slightly lower on the high flow. The graph shows that SE:B1B2+B3-T1

performed better than SE:B1+B2B3-T1. The calibrated parameters in the sequential method have trouble reflecting the observed graph from July 2014 to July 2015, similar to the basin-by-basin and simultaneous methods.

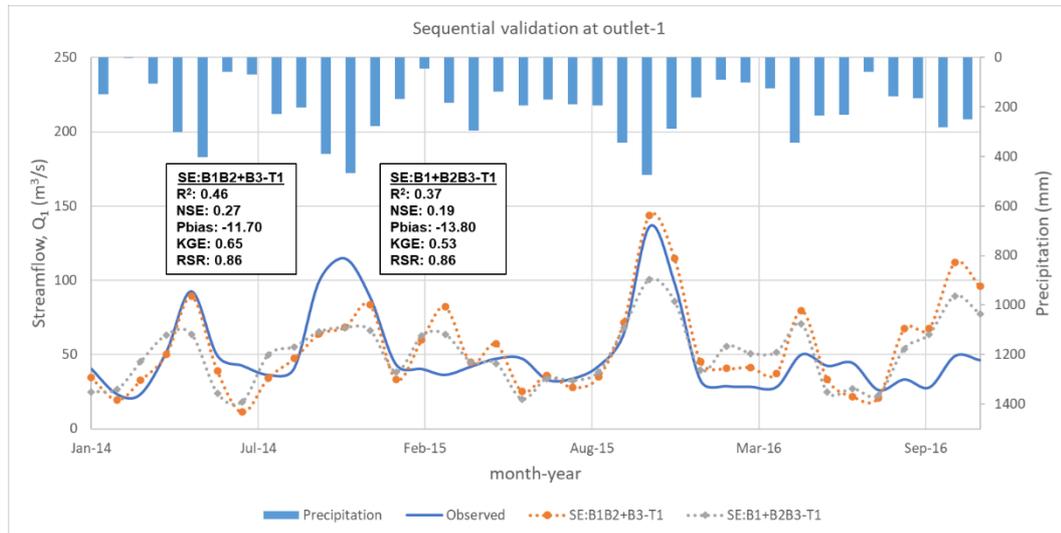


Figure 6-16: validated result of sequential technique at outlet-1

A calibration procedure has not been performed inside basin-2 (B2) since it has been integrated with other basins owing to data limitations. By using the combination technique with B1 and B3, parameters in B2 were calibrated, and they were then verified at the B2 outlet (outlet-4). The SE:B1B2+B3-T2 outperformed the SE:B1+B2B3-T2 according to the validated results in Table 6-8, which demonstrate that SE:B1B2+B3-T2 achieved two out of five ‘satisfactory’ performance ratings (Pbias and KGE). The SE:B1B2+B3-T2 performance ratings value for  $R^2$  is 0.48, NSE is 0.34, Pbias is -23.60, KGE is 0.63, and RSR is 0.81. The SE:B1+B2B3-T2 performance ratings, however, were deemed to be ‘unsatisfactory’ with  $R^2 = 0.04$ , NSE = -0.61, Pbias = 39.50, KGE = 0.01, and RSR = 1.27. Figure 6-17 illustrates how the simulated graph of SE:B1B2+B3-T2 significantly reflected the observed data, whereas the simulated graph of SE:B1+B2B3-T2 did not reflect at all. SE:B1B2+B3-T2 was identical to observed streamflow especially on Jan to July 2014 and August 2015 to March 2016. However, some of the duration is marginally different, either as a result of a drought occurrence or the operation of a dam in SRB. At outlet-4, regionalisation between B2 and B3 is not significant where the

model did not reflect to this combination. As a result, compared to SE:B1+B2B3-T2, calibrated parameters from SE:B1B2+B3-T2 were appropriate in representing basin-2 (B2).

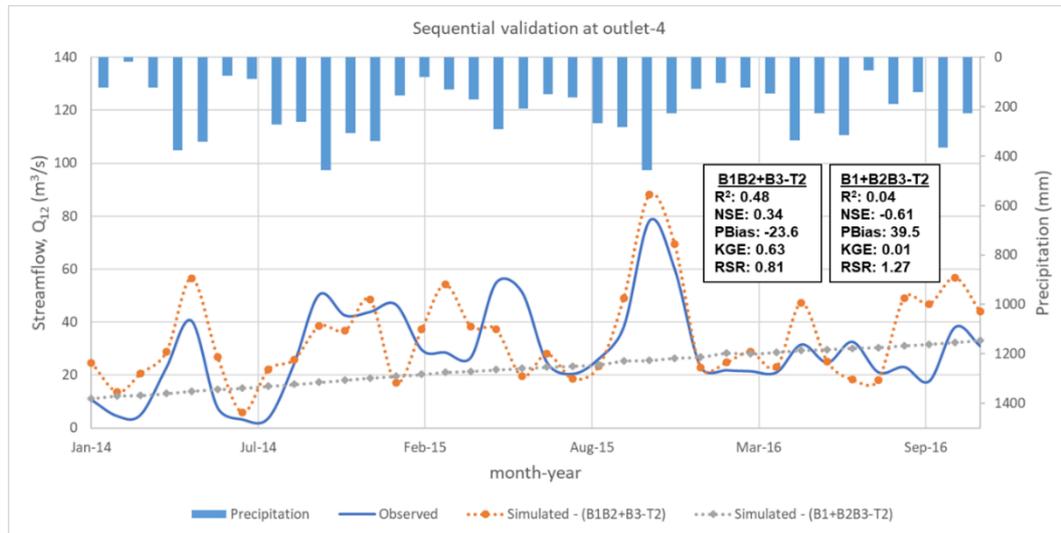


Figure 6-17: validated result of sequential technique at outlet-4

### 6.3 Model performances and uncertainty analysis

In order to create a successful rainfall-runoff model, the calibration and validation phases are essential because they establish the ideal model parameters to reflect the basin situation. The calibrated and validated parameters applied in the models were analysed and calculated using five objective functions in order to determine the ranges. The best MSC and SSC validation results, as well as the model's uncertainty analysis, were shown in Table 6-9 along with five performance rating values. Overall, the sequential technique outperformed the simultaneous method and basin-by-basin at all outputs, earning a greater range in performance ratings and uncertainty analysis of *p-factor* and *r-factor*. Table 6-10 displays the uncertainty analysis recommended by Abbaspour (2008) as well as the minimal range of performance ratings that are permitted.

At outlet-1, the simultaneous technique outperformed the sequential technique considerably, with  $R^2 = 0.52$  (satisfactory),  $NSE = 0.15$  (unsatisfactory),  $Pbias = 2.10$  (very good),  $KGE = 0.70$  (satisfactory), and  $RSR = 0.92$  (unsatisfactory). However, simultaneous uncertainty analysis is higher with a  $0.94$   $p$ -factor and  $3.28$   $r$ -factor, compared to the sequential techniques, which are  $0.64$  and  $1.55$ , respectively.

Table 6-9: Single-site and multi-site validation results

| SWAT Model |             | SSC  | MSC:BB | MSC:Sequential | MSC:Simultaneous |
|------------|-------------|------|--------|----------------|------------------|
| Outlet-1   | $R^2$       | 0.51 | 0.36   | 0.46           | 0.52             |
|            | NSE         | 0.46 | 0.19   | 0.27           | 0.15             |
|            | Pbias       | 0.10 | 0.10   | -11.7          | 2.10             |
|            | KGE         | 0.68 | 0.59   | 0.65           | 0.70             |
|            | RSR         | 0.74 | 0.90   | 0.86           | 0.92             |
|            | $p$ -factor | 0.67 | 0.31   | 0.64           | 0.94             |
|            | $r$ -factor | 1.29 | 0.89   | 1.55           | 3.28             |
| Outlet-4   | $R^2$       |      | 0.29   | 0.48           | 0.47             |
|            | NSE         |      | 0.22   | 0.34           | 0.19             |
|            | Pbias       |      | 0.00   | -23.6          | -14.50           |
|            | KGE         |      | 0.51   | 0.63           | 0.59             |
|            | RSR         |      | 0.88   | 0.81           | 0.90             |
|            | $p$ -factor |      | 0.67   | 0.53           | 1.00             |
|            | $r$ -factor |      | 1.30   | 1.47           | 3.42             |
| Outlet-12  | $R^2$       |      | 0.07   | 0.05           | 0.09             |
|            | NSE         |      | -2.17  | -2.09          | -4.74            |
|            | Pbias       |      | 0.00   | 11.7           | 12.70            |
|            | KGE         |      | 0.17   | 0.03           | -0.97            |
|            | RSR         |      | 1.78   | 1.76           | 2.40             |
|            | $p$ -factor |      | 0.58   | 0.72           | 0.97             |
|            | $r$ -factor |      | 1.65   | 2.76           | 6.67             |

Table 6-10: Minima range of performance rating and uncertainty analysis

|                      |                 | Minima acceptable range |
|----------------------|-----------------|-------------------------|
| Performance rating   | $R^2$           | $\geq 0.5$              |
|                      | NSE             | $\geq 0.5$              |
|                      | PBias           | $\leq \pm 25$           |
|                      | KGE             | $\geq 0.5$              |
|                      | RSR             | $\leq 0.7$              |
| Uncertainty analysis | <i>p-factor</i> | $> 0.7$                 |
|                      | <i>r-factor</i> | $< 1.5$                 |

Aside from the simultaneous techniques, SSC also performed well at outlet-1 with  $R^2 = 0.51$  (satisfactory), NSE = 0.46 (unsatisfactory), Pbias = 0.10 (good), KGE = 0.68 (satisfactory), and RSR = 0.74 (unsatisfactory). With *p-factor* = 0.67 and *r-factor* = 1.29, the uncertainty analysis is likewise within acceptable bounds. However, due to the possibility of missing regional variability and failing to fulfil the demands of the whole watershed (Bai et al., 2017), these results may not correctly reflect the SRB condition. This is demonstrated by the basin-by-basin technique (MSC:BB) in Table 6-9, where the performance rating results of  $R^2$  dramatically drop to 0.36, NSE to 0.19, Pbias to 0.10, KGE to 0.59, and RSR to 0.90 at outlet-1. Figure 6-18 showed how every single simulated graph closely matched the actual graph at outlet-1. The SSC and MSC methods also effectively captured the high and low flow. According to the plotted graph, models had a difficult time reproducing the observed data from June 2014 to June 2015 and June to December 2016.

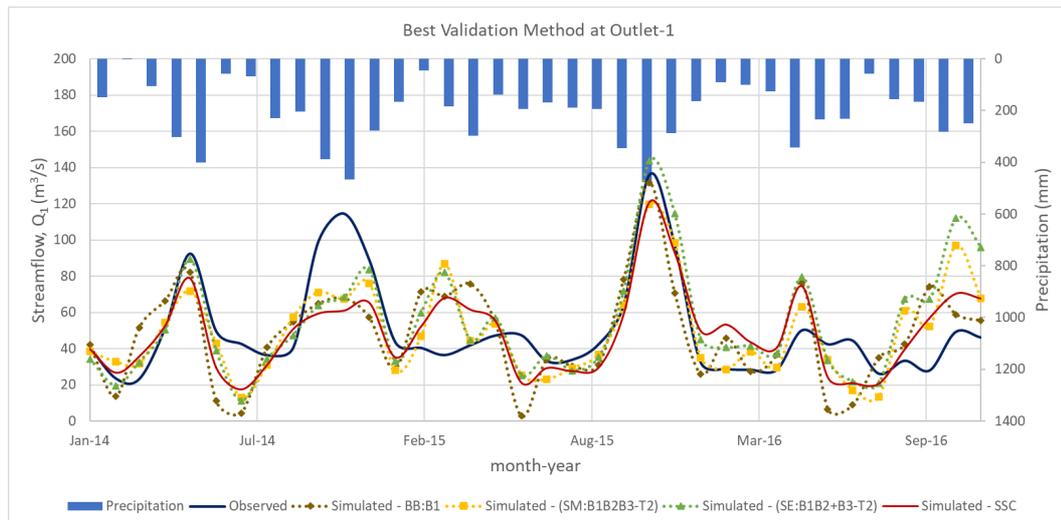


Figure 6-18: single-site and multi-site validation results at outlet-1

For basin-2, streamflow data at outlet-4 starts to be available from 2014 onwards. Due to this, outlet-4 is applicable only for the validation process in the MSC method. Basin-by-basin applied the neighbouring technique to define the best parameters in basin-4 while sequential and simultaneous applied the basin combination approach.

Sequential techniques outperformed simultaneous and basin-by-basin techniques in MSC methods in terms of performance ratings. It obtained 0.48, 0.34, -23.6, 0.63, and 0.81 for  $R^2$ , NSE, Pbias, KGE, and RSR, respectively. although the  $p$ -factor and  $r$ -factor of the uncertainty analysis are 0.53 and 1.47, respectively. The simultaneous technique attained a comparable range to the sequential technique, but it performed poorly in the uncertainty analysis, with  $p$ -factor and  $r$ -factor scores of 1.00 and 3.42, respectively. While the uncertainty analysis for the basin-by-basin technique nearly reached the minimum range of  $p$ -factor and  $r$ -factor with scores 0.67 and 1.30, respectively. However, the performance ratings were in the 'unsatisfactory' range, which is shown in Table 6-9. Figure 6-19 demonstrates that the simulated streamflow patterns produced by MSC methods are identical to one another and match the observed high and low flows.

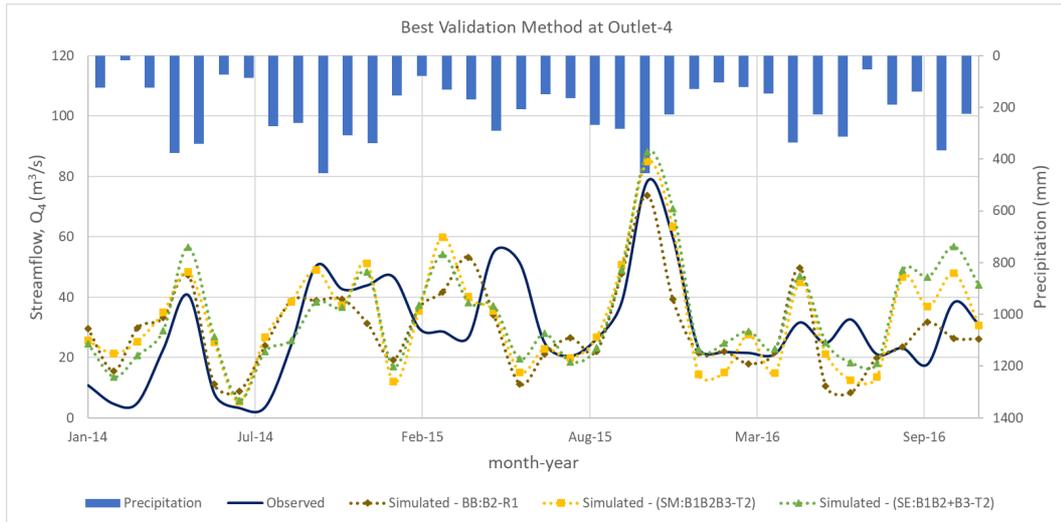


Figure 6-19: multi-site validation result at outlet-4

The MSC method used in basin-3 did not adequately duplicate the streamflow data at outlet-12. Only Pbias showed an acceptable range, with 0.00 in basin-by-basin, 11.7 in sequential, and 12.7 in simultaneous techniques. According to Table 6-9, the *p*-factor and *r*-factor have ranges of 0.58, 0.72 and 0.97, and 1.65, 2.76, and 6.67, respectively. As illustrated in Figure 6-20, the sequential and simultaneous output are fundamentally the same but have different magnitudes.

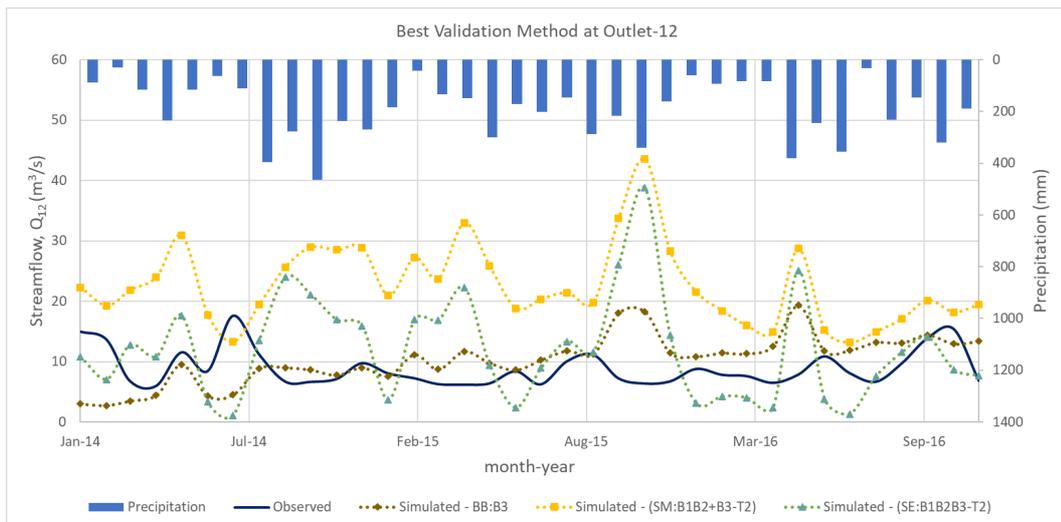


Figure 6-20: multi-site validation results at outlet-12

### 6.3.1 Summary of the results

According to the results in Table 6-9, it can be concluded that in the MSC method, the sequential techniques produced the best rainfall-runoff model for SRB compared to basin-by-basin and simultaneous techniques. This result is similar to the findings from Leta et al., (2017) and Nkiaka et al., (2017). Table 6-11 lists the optimum sequential technique parameters in the SRB SWAT model. The simultaneous technique performed as well as sequential according to five objective functions, but poorly performed in uncertainty analysis with an *r-factor* higher than 1.5 at all outlets. SSC effectively performed at outlet-1 with four objective functions achieved beyond the minimum level as listed in Table 6-9. However, the performance ratings abruptly decreased when the SWAT model used the basin-by-basin technique. It shows how crucial the calibration method is determining the optimum parameters to represent the watershed conditions. As stated by Anderson et al., (2002) and Bai et al., (2017), SSC has a minimal impact, especially when it comes to the big watershed's geographic variability and its multiple characteristics. In this study, all simulated analyses indicated that the objective functions of Pbias and RSR were in the "very good" and "unsatisfactory" ranges, respectively. For instance, the simulated graphs at outlet-12 did not precisely represent the observed data, although the Pbias results showed that all calibration techniques had "good" performance ratings. The RSR showed an "unsatisfactory" range even though the simulated graph significantly resembled the observed data in the sequential procedure at outlet-1. As a result, the objective functions of  $R^2$ , NSE, and KGE as well as the uncertainty analysis were referred to in order to determine the best SRB model. These objective functions provide more realistic indicators of model performance. The SWAT model had trouble replicating streamflow conditions in 2014, notably at outlet-1 and 4 in both SSC and MSC techniques, since there was a problem with water depletion as a consequence of prolonged dry seasons, leading to a water crisis in the state of Selangor (Anang et al., 2017).

Table 6-11: SRB SWAT model parameters

| Parameter_Name      | Outlet-1      | Outlet-4      | Outlet-12     |
|---------------------|---------------|---------------|---------------|
|                     | SE:B1B2+B3-T2 | SE:B1B2+B3-T2 | SE:B1B2+B3-T1 |
| 1:R_CN2.mgt         | -0.311        |               | 0.133         |
| 2:R_SOL_AWC(..).sol | 1.688         |               | 0.767         |
| 3:V_ESCO.hru        | 0.173         |               | -0.282        |
| 4:V_GW_DELAY.gw     | 17.24         |               | 466           |
| 5:V_GWQMN.gw        | 3715          |               | 1995          |
| 6:V_GW_REVAP.gw     | 0.104         |               | 0.357         |
| 7:R_HRU_SLP.hru     | 0.898         |               | 0.480         |
| 8:R_OV_N.hru        | 16.64         |               | 14.51         |
| 9:R_SLSUBBSN.hru    | 116.5         |               | 127.1         |
| 10:R_SOL_K(..).sol  | -717.3        |               | 58.43         |
| 11:R_SOL_BD(..).sol | 3.289         |               | 1.424         |
| 12:V_CANMX.hru      | 92.73         |               | 38.80         |
| 13:V_CH_K2.rte      | 13.69         |               | 639.6         |
| 14:V_ALPHA_BF.gw    | 1.151         |               | 0.001         |
| 15:V_CH_N2.rte      | -0.009        |               | -0.001        |
| 16:V_LAT_TTIME.hru  | 38.71         |               | 99.47         |
| 17:V_RCHRG_DP.gw    | 0.010         |               | 0.003         |
| 18:V_SURLAG.bsn     | 12.36         |               | 30.11         |
| 19:V_EPCO.hru       | 0.378         |               | 1.114         |
| 20:V_REVAPMN.gw     | 263.9         |               | -132.0        |

There are just two stations available in SRB to assess the model performance. It is a challenging task to determine the accuracy of each gauge since there are no additional references points for comparison. With limited data from only two gauges, the uncertainty estimates may be less precise or less representative of the true uncertainty. Therefore, a regionalisation method was introduced in order to increase reliability and improve accuracy by getting an optimum parameter to represent the ungauged basin. It is anticipated that there will be an unusually high number of water abstraction operations in SRB during these challenging times. At outlet-12, MSC performed poorly, while in basin-3, none of the techniques worked well with the observed streamflow. The sequential and simultaneous procedures both created a similar simulated flow pattern with a different amplitude, even if it does not exactly match the observed flow. Additionally, it was found that there are discrepancies between streamflow data and rainfall data, with outlet-12 streamflow data not mirroring rainfall data. Basin-3 covered 90% of the dam catchment area, therefore it was a challenging basin for SWAT to simulate when dam operation was involved. A specific model or technique applicable for dam operation should be applied in future research.

## **CHAPTER 7**

### **CLIMATE CHANGE ASSESSMENT**

In chapter 6, the SWAT (Soil and Water Assessment Tool) rainfall-runoff model for the SRB (Selangor River Basin) was developed utilising the best combination of datasets, while enhancing model accuracy using the MSC (multi-site calibration) technique. The optimum hydrological model for SRB is defined by the combination of the datasets of the gauge station and the global reanalysis datasets from CMADS (China Meteorological Assimilation Driving Dataset), with the sequential technique producing a satisfactory simulation to depict the SRB environment. Evaluation of the effects of climate change has emerged as a major issue of discussion in Malaysia in recent years. However, the literature only has a few research on the influence that climate change would have on the hydrologic processes in Malaysia.

Modelling interactions between the atmosphere and land surface hydrologic systems in a range of conditions is one of the most fundamentally difficult areas of the science of assessing climate change. The hydrologic and atmospheric sciences are both up against this obstacle (Kavvas et al., 2013). In addition to that, the purpose of this research is to evaluate the impacts of climate change on future SRB water resources by using the optimal SWAT model in conjunction with an ensemble of future climate prediction datasets obtained from the atmospheric boundary layer. To achieve these goals, future precipitation, and temperature projections were obtained for 10-year periods starting from 2030 to 2050. These projections were based on the maximum, minimum, and average values of an ensemble of 15 different future climate realisations that were derived from coarse-resolution General Climate Model (GCM) projections for the 21st century.

## **7.1 First Generation Coupled General Circulation Model (CGCM1)**

General Circulation Models (GCM) provide numerous climate change projection datasets around the world, and they are available at very coarse resolutions for Peninsular Malaysia. In order to simulate the First Generation Coupled General Circulation Model (CGCM1) at a spatial resolution of approximately 410 km (Flato et al., 2000), the Canadian Center for Climate Modelling and Analysis (CCCma) used climate change simulation datasets corresponding to rising greenhouse gas (GHG) concentrations and changing sulphate aerosol (A) loadings (GHG+A) of the IPCC IS92a Scenario Run, as illustrated in Figure 7-1. The CCCma second generation atmospheric GCM and a version of the GFDL (Geophysical Fluid Dynamics Laboratory) modular ocean model MOM1 are combined in the CGCM1 three-dimensional climate model (Chen et al., 2006). Because there is only one CGCM1 model grid with one nodal point that falls within the Peninsular Malaysia area, the grid resolution of the CGCM1 model is too coarse for hydrological research to be conducted there. As a result, the GCM is unable to accurately depict the specifics of the local climate in Peninsular Malaysia.

To address this issue, the CGCM1 for Peninsular Malaysia was downscaled to a finer gridded resolution in the Regional Hydro Climate Model. There are numerous regional climate models (RCMs) available from worldwide laboratories with over 400 publications per year in the late 2010s since they first originated in the late 1980s (Giorgi, 2019). The Regional Hydro Climate Model for Peninsular Malaysia (RegHCM-PM) was used in order to apply the climate change predictions data to Peninsular Malaysia, and specifically to SRB. The Regional Hydro Climate Model for Peninsular Malaysia (RegHCM-PM) was created based on the UCD-PWRI Integrated Regional-Scale Hydrologic/Atmospheric Model (IRSHAM), which was used in Japan (Chen et al., 2006)

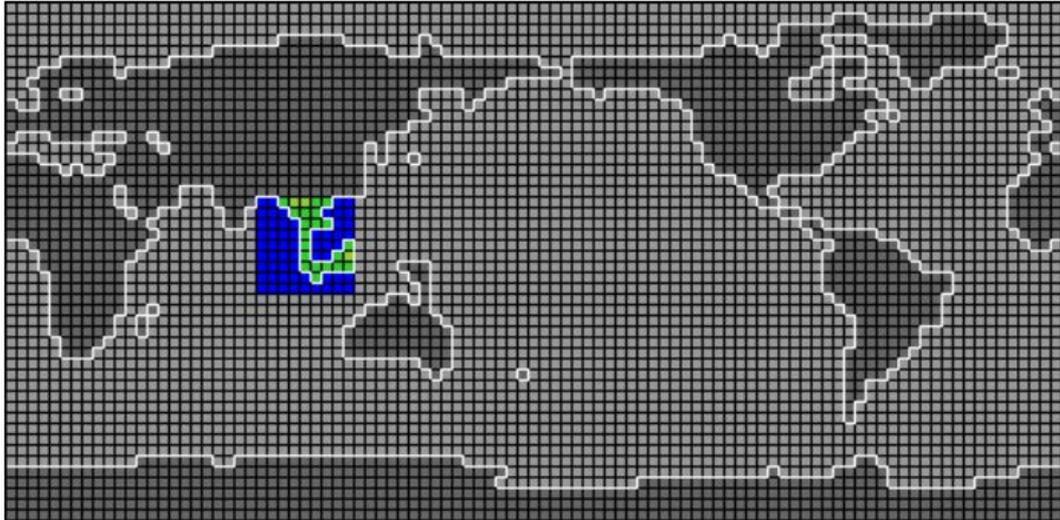


Figure 7-1: CGCM1 grid data used in the RegHCM-PM. Blue represents the ocean grids. While green represents the land grids. (source; Chen et al., (2006))

## 7.2 Integrated Regional-Scale Hydrologic/Atmospheric Model (IRSHAM)

IRSHAM is the first regional-scale climate model used for downscaling research that has investigated the influence of climate change on the hydrologic regime in an area the same size as Peninsular Malaysia. IRSHAM used a 20 km grid resolution in Japan, South Korea, and California to account for the influence of local topography and land surface conditions on the local climates of its many subregions (Chen et al., 2006). However, IRSHAM cannot be used with a spatial grid resolution of less than 20 kilometres since the atmospheric component of IRSHAM is a modified version of the hydrostatic regional atmospheric model (Mathur, 1983). A nonhydrostatic regional-scale (mesoscale) atmospheric model, such as the US National Center for Atmospheric Research's (NCAR) fifth Generation Mesoscale Model (MM5), is required to translate the findings of the CGCM1 simulations to the local scale of watersheds and sub-regions in Peninsular Malaysia.

### **7.3 The Fifth-Generation Mesoscale Model (MM5)**

The Fifth-Generation Mesoscale Model (MM5) is the newest version in a series after Richards et al., (1978) and it is based on the Fourth-Generation Mesoscale Model (MM4). The nonhydrostatic model MM5, which has a spatial resolution of up to 0.5 km, can simulate the effects of Malaysia's steep terrain on the regional climate at the watershed scale (spatial resolution <10 km) (Amin et al., 2016). Standard Cartesian grids are used by MM5 in the horizontal directions, while a terrain-following  $\sigma$ -coordinate is used in the vertical directions. When combined with IRSHAM's land surface hydrology component, MM5 was successfully used to recreate historical atmospheric data over the 67 watersheds of Lake Tahoe, California, from 1965 to 2001 with a spatial resolution of 3 km (Anderson et al., 2007), and showed excellent agreement between the simulated and actual hydroclimatic conditions in terms of temperature, precipitation, and river runoff over a 960,000 km<sup>2</sup> Tigris-Euphrates basin (Chen et al., 2006).

### **7.4 Ensemble approach**

In order to quantify the uncertainty from three sources based on Hawkins and Sutton (2009), which are the internal variability of the climate system (natural fluctuations that occur in the absence of any radiative forcing of the planet), model uncertainty (uncertainty due to different responses of climate models in response to the same radiative forcing), and scenario uncertainty (uncertainty in future emissions of greenhouse gases), the ensemble approach of multiple future projections from multiple GCMs (General Circulation Models) based on multiple scenarios was applied in RegHCM-PM. Three distinct coupled land-atmosphere-ocean general circulation models (GCMs) based on the Fourth Assessment Report (AR4) of the United Nations Intergovernmental Panel on Climate Change (IPCC) were used to account for the model uncertainty in climate change simulation. The applied models include the coupled ocean-atmosphere general circulation model (CGCM2.3.2) from the Meteorological Research Institute (MRI) of Japan (Yukimoto et al., 2001), the fifth generation GCM ECHAM5 from the Max Planck Institute (MPI) for Meteorology in Germany (Roeckner et al., 2006), and the third generation Community Climate

System Model version 3 (CCSM3) from the University Corporation for Atmospheric Research (UCAR) in the (Collins et al., 2006). To account for the uncertainty in future greenhouse gas emissions, the full range of future greenhouse emission scenarios from the Special Report on Emission Scenarios (SRES) (Nakicenovic et al., 2000) were covered, including the best possible scenario (SRES B1), the worst-case scenario (SRES A1FI), the most likely scenario (SRES A1B), and the second worst scenario (SRES A2). These hypothetical situations are based on numerous projections for future estimates such as growth in population, economic and social development, technology advancements, and energy resources Table 7-1. Table 7-1 below describes the future greenhouse emission scenarios of SRES with an ensemble of 15 different future climate realisations applied in SRB.

Table 7-1: Special Report on Emission Scenarios (SRES) of 15 different future climate realisations. source: (IPCC, 2014b)

| SRES | Scenario  | Realization | Model  |
|------|---|-------------|--|
| A1B  | The most plausible scenario, describes a future world of rapid economic growth, a global population that peaks in mid-century and declines thereafter, and increase cultural and social interactions.<br><br>The technological emphasis of this scenario is on a balance across all energy sources, not relying too heavily on any particular energy source.  | ccsm3a1b    | CCSM3: The Community Climate System Model (CCSM) is a coupled model for simulating past, present and future climates. The Community Climate System Model version 3 (CCSM3) is a coupled climate model with components representing the atmosphere, ocean, sea ice, and land surface connected by a flux coupler.<br><br>ECHM5: The atmospheric GCM ECHM5 has a spectral dynamical core where vorticity, divergence, temperature and surface pressure are represented in the horizontal by a truncated series of harmonics. A semi-Lanrangian scheme is used for the transport of water components (water vapour, cloud liquid water and cloud ice).<br><br>MRI: The AGCM component of MRI-CGCM2.3.2 is based on a version of the operational weather forecasting model of the Japan Meteorological Agency (JMA). It has a spectral dynamical core where vorticity, divergence, temperature, specific humidity and surface pressure represented in the horizontal by a truncated series of spherical harmonics. |
|      |   | echam5a1b1  |  |
|      |   | echam5a1b2  |  |
|      |   | echam5a1b3  |  |
|      |   | mria1b      |  |
| A2   | Very heterogeneous world. The underlying theme is self-reliance and preservation of local identities.<br><br>Fertility patterns across regions of the world converge very slowly, resulting in continuously increasing world population. Economic development is primary regionally oriented and slower than in other scenarios.  | ccsm3a2     |  |
|      |   | echam5a2_1  |  |
|      |   | echam5a2_2  |  |
| B1   | A convergent world with the global population peaking around mid-century and declining thereafter.<br><br>Rapid changes in economic structures toward a service and information economy, with reductions in materials intensity, and the introduction of clean and resource-efficient technologies.   | ccsm3b1     |  |
|      |   | echam5b1_1  |  |
|      |   | echam5b1_2  |  |
|      |   | echam5b1_3  |  |
| A1F1 | A future world of very rapid economy growth, low population growth and rapid introduction of new and more efficient technology.<br><br>Major underlying themes are economic and cultural convergence and capacity building, with a substantial reduction in regional differences in per capita income.<br><br>People pursue personal wealth rather than environmental quality. The worst case scenario among all scenarios. | mrib1       |  |
|      |   | ccsm3a1fi   |  |

The RegHCM-PM model and the SRB SWAT rainfall-runoff model were combined with 15 distinct future climate scenario models in order to conduct the evaluation of watershed-scale hydrologic processes that occur in SRB. These climate predictions were utilised in the future climate projection simulations in order to account for internal variability. These climate forecasts were derived from three GCMs and were run under four distinct emission scenarios. In fact, these climate projection realisations have been calibrated and validated at streamflow stations throughout the defined watershed by Amin et al., (2019), with SRB being one of the locations, and it produces a credible result when compared to the observed streamflow hydrograph. The reliable climate projection datasets were applied to the SRB SWAT model in order to project the future hydrologic condition in SRB as well as to determine the water availability within the SRB watershed.

## **7.5 Regional Hydro Climate Model of Peninsular Malaysia**

The RegHCM-PM was created to downscale the climate change simulations run by the Canadian Center for Climate Modelling and Analysis (CCCma) using the First-Generation Couple General Circulation Model (CGCM1). RegHCM-PM incorporates both a regional land hydrology model and a mesoscale atmospheric model, same as the original IRSHAM for Japan. However, the RegHCM-PM atmospheric model is the MM5 (nonhydrostatic), as opposed to IRSHAM, which is hydrostatic. The IRSHAM model was chosen as the reference model because it was the only model used to study how climate change affected the hydrologic regime in a region similar to Peninsular Malaysia (Chen et al., 2006). IRSHAM has been preserved in earlier RCM research because it is the only regional land hydrology model that is physically based and has upscaled hydrologic conversion equations (Chen et al., 2006). However, Amin et al. (2017) and Amin et al. (2019) integrated a physically based hydrology model, the Watershed Environmental Hydrology Model (WEHY), with a regional climate model, MM5, to generate more realistic model interactions.

This study proposed using the optimum SWAT model described in Chapter 6 as a regional land hydrology model and coupled with the MM5 regional model because the RegHCM-PM produced reasonable simulations of historical hydroclimatic trends over Peninsular Malaysia at the scales of watersheds and subregions, and because the simulated hydroclimatic variables (air temperature, precipitation, river flow), are of similar order of magnitude and seasonal trends as their observed counterparts. SWAT is one of the most used hydrological models for evaluating model quality, land use, and climate change (Mou Leong Tan et al., 2019). The RegHCM-PM downscaled a coarse spatial resolution of ~410 km CCCma to a smaller gridded resolution of 6 km in order to analyse the influence of climate change on Peninsular Malaysia's hydrology and water resources (Ir. Mohd Zaki bin Mat Amin et al., 2019).

Three layered domains were used in the RegHCM-PM downscaling method. In the first run of the RegHCM-PM, the whole Peninsular Malaysia area and the surrounding territories were covered by an outer domain of the world scale atmospheric data with initial and boundary conditions. The Southeast Asia area is covered by the outer domain's 54 km-squared grid. The second domain (intermediate domain) of RegHCM-PM has a grid size of 18 km nested inside the outer domain and runs over the inner domain with its initial conditions derived from global scale atmospheric data interpolated at 6 km. Boundary conditions of the second domain derived from the simulation results of the second domain RegHCM-PM simulations. Figure 7-2 illustrates the results of RegHCM-PM simulations using these initials, boundary conditions, and inner domain to produce comprehensive atmospheric data (precipitation, air temperature, radiation, wind, relative humidity, and evapotranspiration) at a spatial grid resolution of 6 km and time steps of 1 hr.

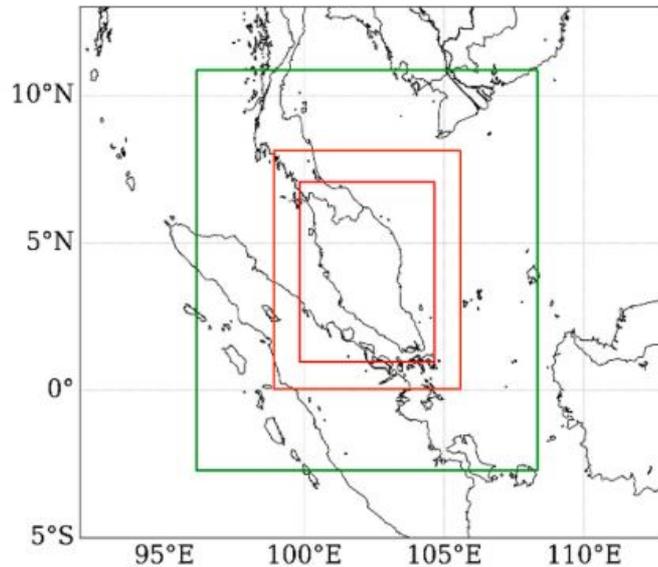


Figure 7-2: The modelling domains of the Peninsular Malaysia hydroclimate model: The green box covered by 54 km x 54 km grids is the larger outer domain; the orange box covered by 18 km x 18 km grids is the intermediate domain; and the red box is covered by 6 km x 6 km grids, which is the inner modelling domain. (source: Amin et al., (2019))

An ensemble method consisting of multiple future projections based on historical data from multiple GCMs (CCSM3, ECHAM5, and MRI-CGCM2.3.2) at coarse grid resolutions of approximately 1.8°, 2.8°, and 1.4° were dynamically downscaled to a finer spatial resolution of approximately 6 km using the MM5 model over the course of the entire 21st century. This resolution was chosen based on research by Jang et al., (2017), which found that banded structures and orographic effects on precipitation and wind fields can be well described by a mesoscale model at 3 km and 9 km grid resolutions, instead of 27 km and 81 km, which may not be enough for watershed-scale or sub-watershed-scale studies. Since these GCMs have been downscaled to a finer spatial resolution, the model-simulated historical atmospheric conditions from 1970-2000 were validated against their corresponding historical observations to evaluate the model performance. Due to the availability of large-scale historical atmospheric data from CGCM1, the years 1970-2000 were chosen as the historical simulation periods for the RegHCM-PM hydroclimate simulations over Peninsular Malaysia.

Between 1970 and 2000, monthly mean gauge observations and downscaled control simulations were compared to assess precipitation simulation biases. For future hydroclimate simulations, the CGCM1 data from the GHG+A1 IPCC IS92a Scenario Run were used, which corresponds to a progressive yearly rise of 1 percent in CO<sub>2</sub> after 1993. Validated results demonstrated that after bias correction was applied, in Peninsular Malaysia, the RegHCM-PM delivered accurate simulations of historical hydroclimatic trends at the scale of watersheds and subregions, and the simulated hydroclimatic variables (air temperature, precipitation, river flow) showed similar seasonal patterns to their real-world equivalents (Ir. Mohd Zaki bin Mat Amin et al., 2019). The RegHCM-PM requires the atmospheric conditions that its atmospheric module reconstructs as an input to its hydrologic module in order to forecast surface hydrologic conditions in SRB.

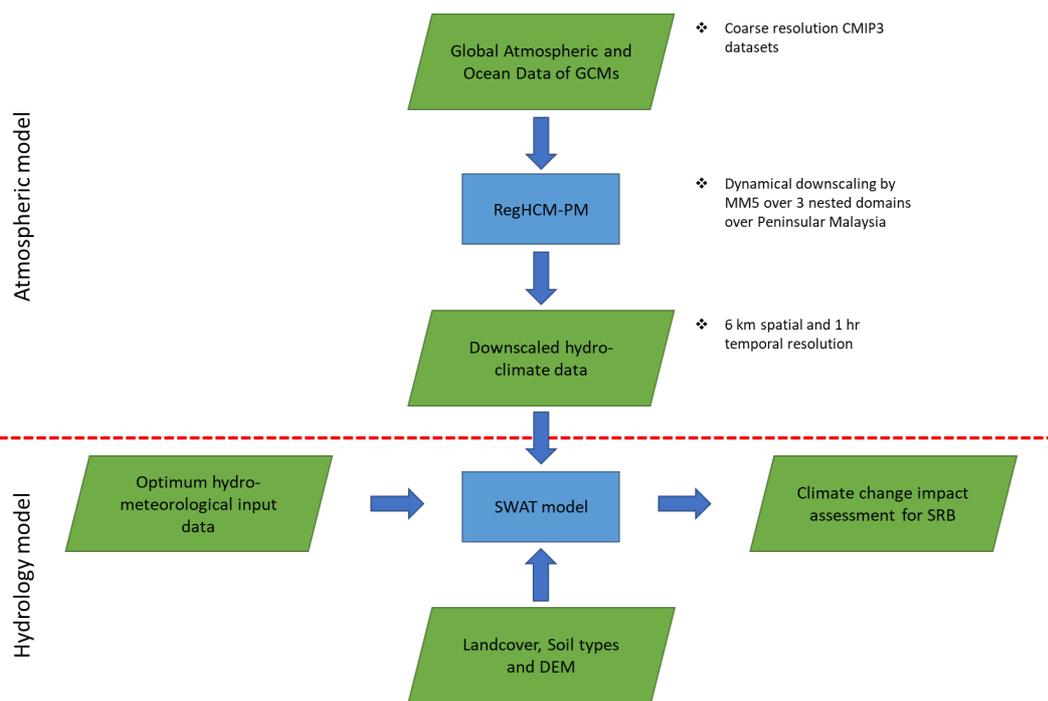


Figure 7-3: Schematic description of the modelling approach

## 7.6 Simulation on future changes in hydrological components

In order to predict the surface hydrologic conditions, this study employed the optimum SWAT hydrological model outlined in Chapter 6 coupled with the outcome of the ensemble data from 15 realisations dynamically downscaled from the RegHCM-PM model. As shown in Figure 7-4, the centroids of the grid cells from the final grid resolution of 6 km of the ensemble data were then converted into grid points using the bilinear interpolation method, and the data has been set up into a SWAT database for the model simulation process. This future climate data was secondary data provided by the National Water Research Institute of Malaysia (Nahrim), the agency which produces and manages the climate change datasets for Malaysia. This study focuses on the rainfall, maximum and minimum temperature data which are the main hydrological components.

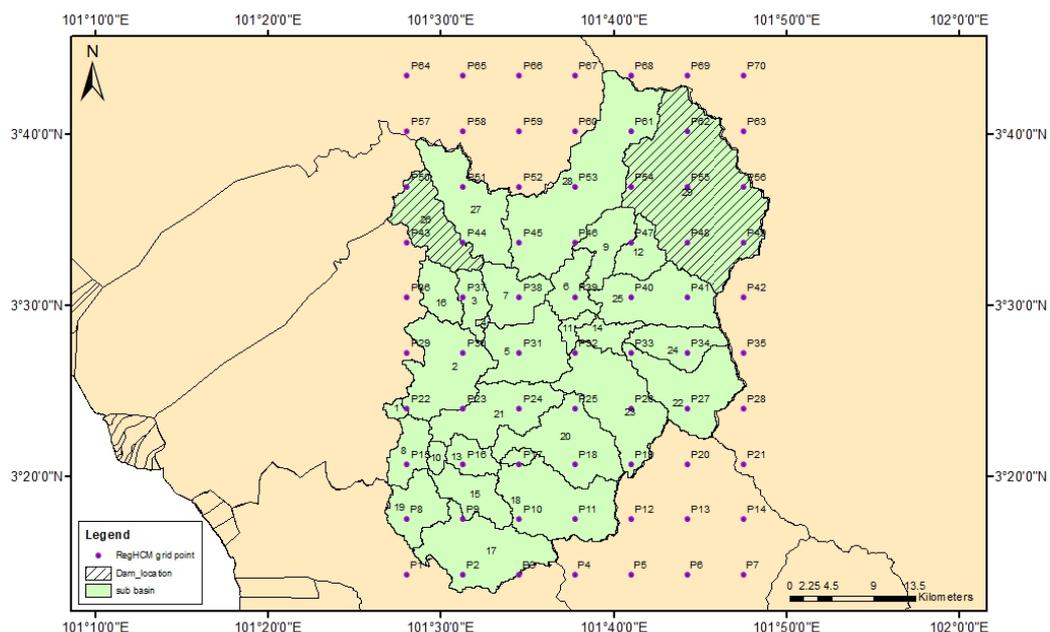


Figure 7-4: RegHCM-PM grid point downscale at 6 km spatial resolution in the SRB. Purple dots represent the centroid of the grid cell.

According to the simulation using the ensemble climate data, the annual precipitation in SRB can be classified into two range groups, with the first group showing the most realisations anticipated to be in the range of 1500 mm to 3500 mm. While the SRES ccsm3a1fi is the second range, which predicts yearly precipitation to range between 400 mm and 1500 mm, is the lowest from other realisations, as depicted in Figure 7-5. According to Table 7-1, the SRES ccsm3a1fi is the worst-case emission scenario. It portrays a future society with extremely rapid economic development, a sluggish population rise, and rapid introduction of new technology that is both more efficient and more advanced. The projected precipitation range of SRES ccsm3a1fi was less than the average annual rainfall range in SRB, which was between 1600 mm and 2500 mm. Overall, most realisations projected an increasing trend pattern of SRB precipitation from 2030 to 2050.

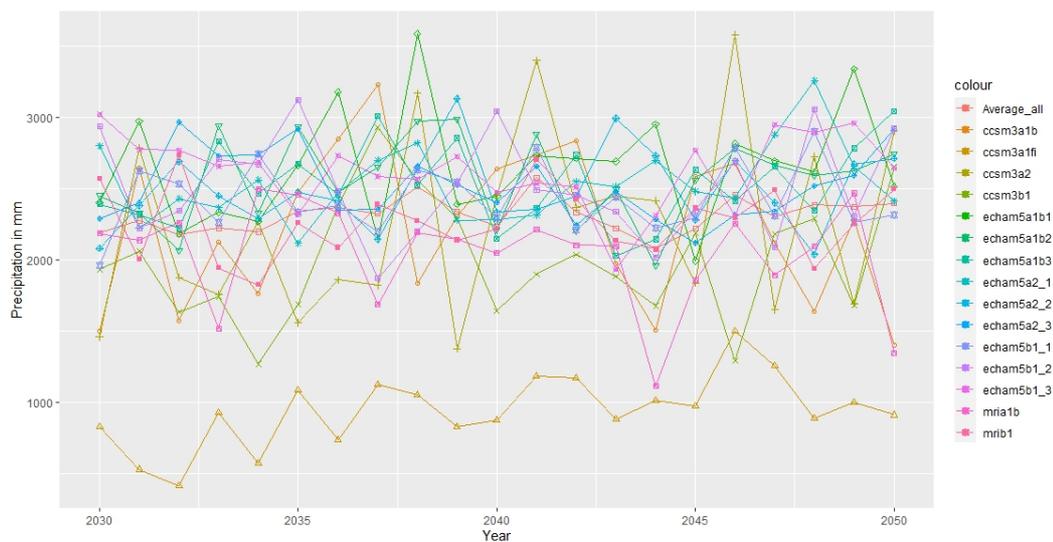


Figure 7-5: Annual precipitation on individual projections of the ensemble data

As illustrated in Figure 7-6, it is predicted that the annual average temperature in SRB will vary between 24.5 and 27 degrees Celsius. The SRES ccsm3b1 realisation has the lowest projected temperature, while the SRES mria1b realisation has the highest. The results indicate a significant increase in linear trendline patterns from the majority of realisations, except for SRES ccsm3b1 and SRES mrib1, which patterns are practically flattening.

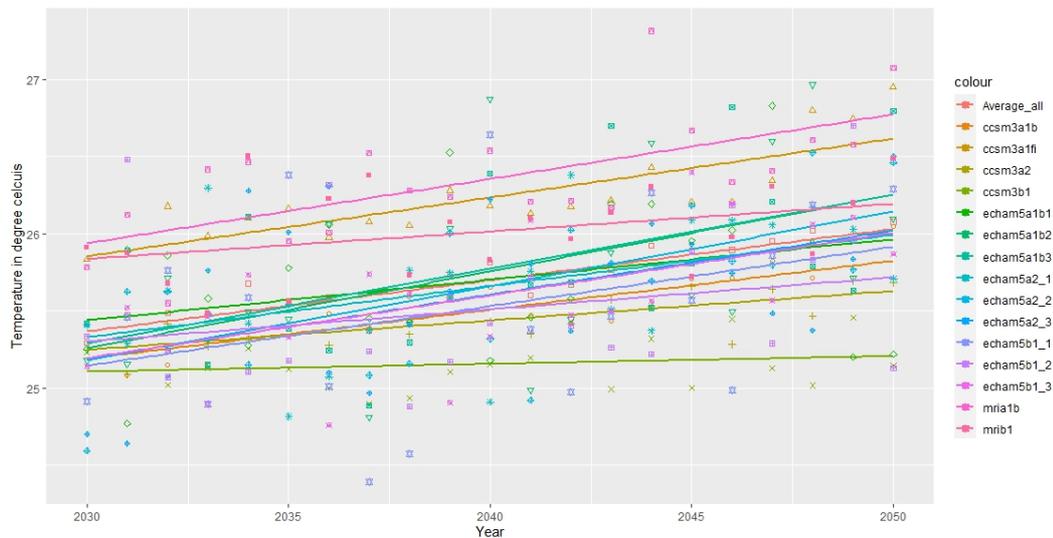


Figure 7-6: Annual temperature on an individual projection of the ensemble data

### 7.6.1 Hydrological model simulation

This study focused on future climate data from the ‘ensemble average’ to represent the first group in precipitation analysis, and the SRES ccsm3alfi to represent the second group, which also indicated the worst-case scenario. The hydrological model simulation was concentrated on the precipitation and temperature data as it were the main datasets needed to run the SWAT hydrological model and the only available dynamically downscaled climate data provided by Nahrim. A monthly streamflow simulation was performed from 2030 to 2050. However, the analysis focused on the years 2030, 2040, and 2050. Three streamflow stations were chosen to analyse how climate change has affected streamflow since the multi-site calibration model is the best model for SRB.

#### ‘Ensemble average’

Different realisations had a different model outcome, and it could jeopardise the flow analysis if used a limited number of projections. In fact, there are no trends in terms of where the regions are located, and each one has a different increment magnitude. Therefore, in addition to the 15 various future climate realisations in SRES, Nahrim has created an ‘ensemble average’ realisation

from the 15 climate forecasts for the 21st century that have been calibrated and validated using RegHCM-PM. The study Vulnerability and Adaptation to Climate Change in Malaysia under Third National Communication by the National Hydraulic Research Institute of Malaysia describes this ‘ensemble average’ (NAHRIM, 2016). The yearly mean flow corresponding to the 15 dynamically downscaled GCM forecasts was calculated as an ‘ensemble average’ utilising simulated monthly mean flows and statistical confidence ranges based on the t-statistic. The forecasts for the annual mean flow conditions in the Muda watershed, one of the study regions in (NAHRIM, 2016), are shown in Figure 7-7 for (a) the individual and (b) the ensemble average with a 10-year moving average and a 95% confidence range.

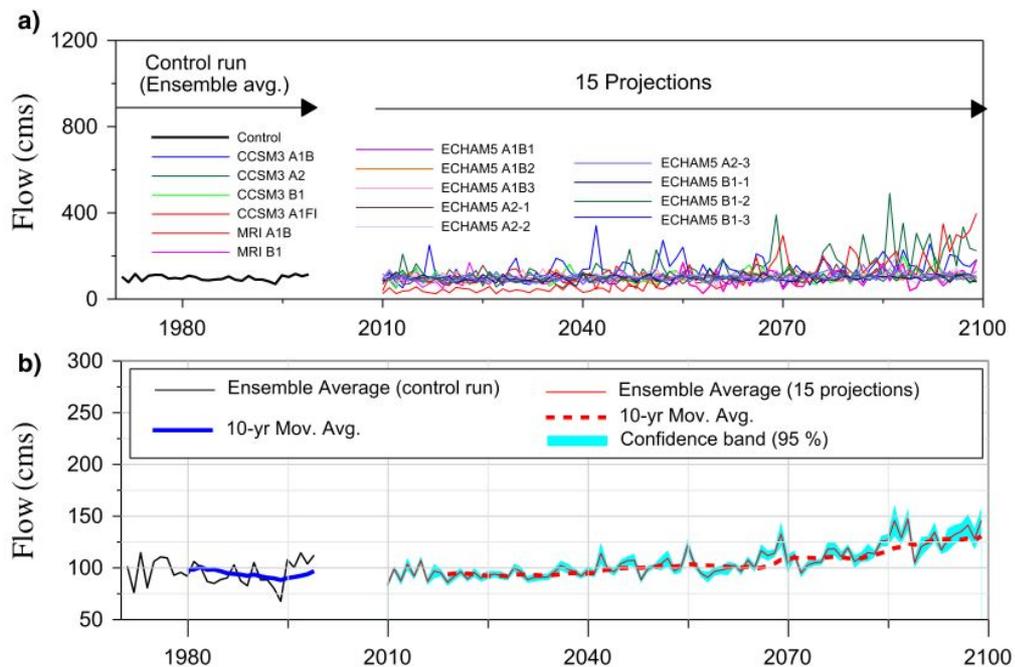


Figure 7-7: Annual mean flow projections in the Muda watershed for (a) individual and (b) ensemble averages using a 10-year moving average and a 95% confidence band. (source: Amin et al., (2017))

### Simulation on the 'ensemble average' datasets

The projection of precipitation based on the 'ensemble average' data shows an increasing linear trend on precipitation throughout the year, as illustrated in Figure 7-8. The highest precipitation is predicted in 2041 with 2,575 mm and the lowest is 2,074 mm in 2044. Similar to the Mann Kendall test with Sen's Slope as shown in Figure 7-9, the Sen's slope is 6.275 with *p-value* and *z* value are 0.075 and 1.782 respectively.

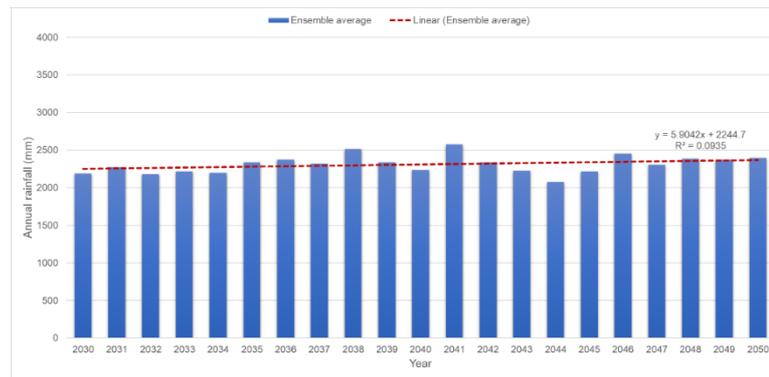


Figure 7-8: Linear trend analysis of precipitation for SRB using the 'ensemble average'.

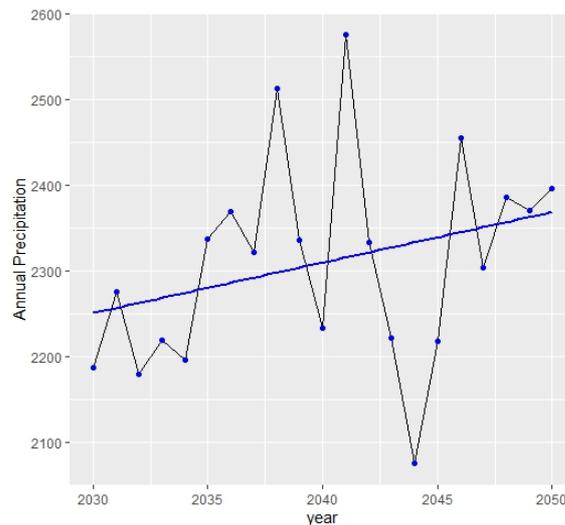


Figure 7-9: Annual rainfall variation of 'ensemble average' from 2030-2050 using the Mann-Kendall test with Sen's slope.

The SRB SWAT model simulation demonstrates that the maximum amount of precipitation forecast for 2030 is 2,382 mm, increasing to 2,505 mm in 2040 and 2,618 mm in 2050. Whereas the minimum precipitation decreased from 1,855 mm in 2030 to 1,825 mm in 2040, with a peak of 2,029 mm in 2050, as shown in Figure 7-10. The SRB's upstream area (sub-basin 28) is predicted to receive the most precipitation during the year, followed by sub-basins 13, 18, and 20 while sub-basin 29 is predicted to get less precipitation during the year and some in sub-basin 8 and sub-basin 17.

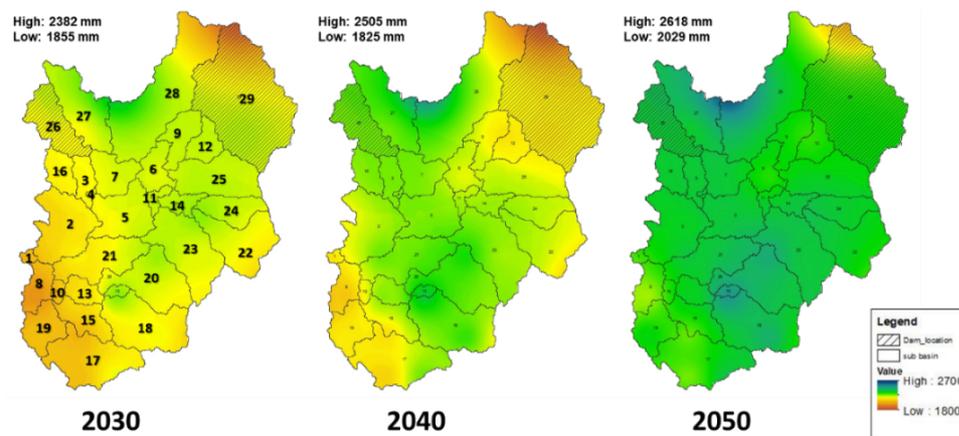


Figure 7-10: Precipitation projection from the 'ensemble average' realisation

The rainfall anomaly for the simulation periods from 2030-2050 was depicted in Figure 7-11. The results shows that the extremely wet rainfall anomaly was predicted to occur in the year 2038, 2041, and 2046 with corresponding values of 6.25, 8.66, and 4.76 respectively. Very wet and moderately wet were predicted to occur from the year 2048 to 2050 with corresponding values of 2.56, 1.99, and 2.58 respectively. On the aspect of dryness anomalies of rainfall within the period of 2030-2050, the result predicted that in the years 2030, 2032, 2034, and 2044 having values of -3.72, -4.22, -3.26, and -7.23 respectively. for the very dry and moderately dry were predicted to occur in the year 2031, 2033, 2040, 2043, 2045, with anomaly values of -1.06, -2.87, -2.07, -2.46, -2.65 respectively.

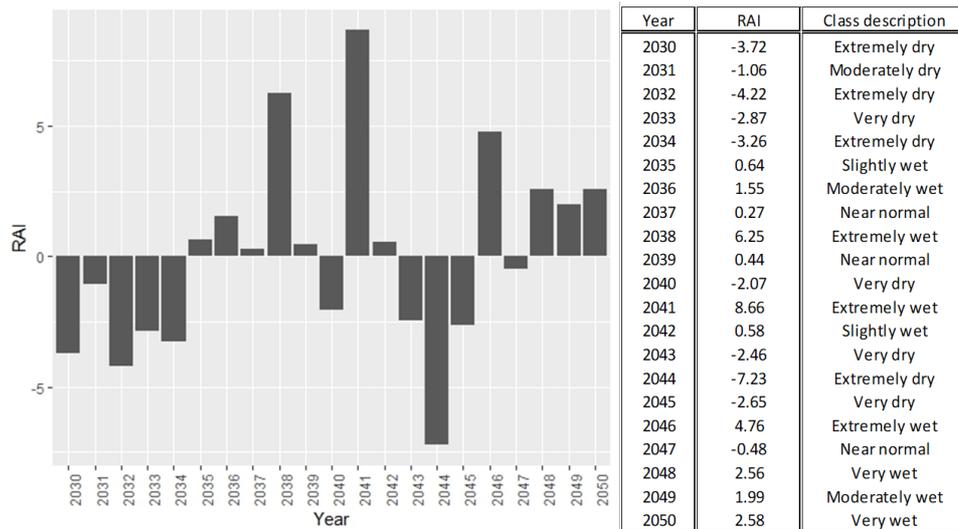


Figure 7-11: Rainfall Anomaly Index (RAI) prediction on the ‘ensemble average’ realisation.

As shown in Figure 7-12, the average temperature projection using the ‘ensemble average’ shows an increasing linear pattern throughout the simulation year. In 2030, the temperature in SRB is predicted to be 25.3 °C and increase by about 0.5 degrees to 25.8 °C in 2040. From 2030 to 2040, the temperature drops about 0.2 degrees, from 25.6 °C in 2036 to 25.38 °C in 2037. In 2041, the temperature drops about 0.2 degree to 25.6 °C from the previous year and increases until it reaches the highest temperature of 26.1 °C in 2050. In Mann-Kendall test with Sen’s slope, the *p-value* and z value are  $1.576 \times 10^{-6}$  and 4.801 respectively. The Sen’s slope is 0.033 as shown in Figure 7-13.

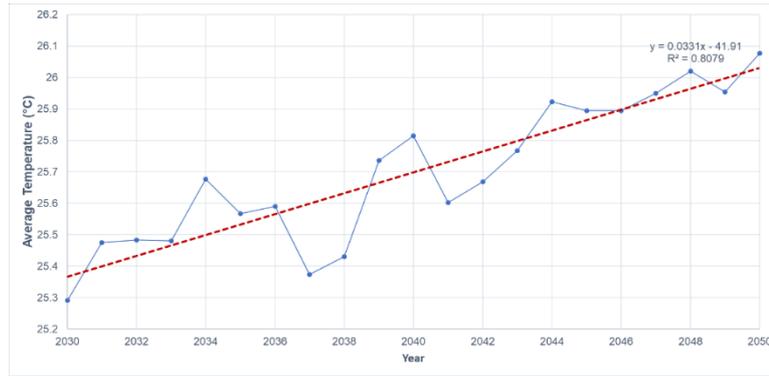


Figure 7-12: Linear trend analysis of average temperature for the SRB using SRES 'ensemble average'.

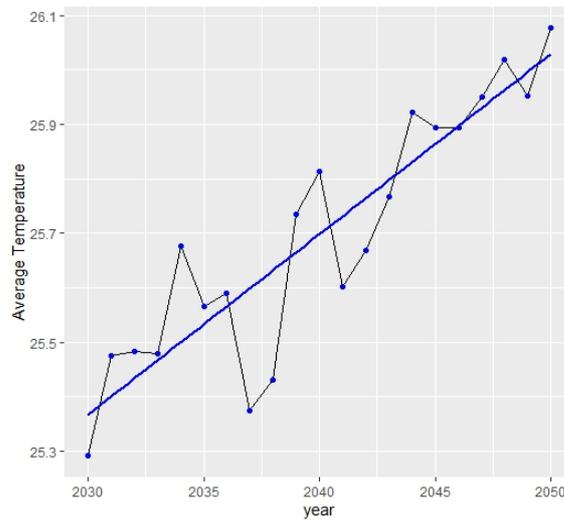


Figure 7-13: Average temperature variation on 'ensemble average' from 2030-2050 using the Mann-Kendall test with Sen's slope.

Figure 7-14 illustrates the temperature pattern in SRB for both maximum and minimum temperatures in 2030, 2040 and 2050. In 2030, the high temperature is 26.9 °C and the low temperature is 20.8 °C, then increases to a high temperature of 27.4 °C and a low temperature of 21.2 °C in 2040. The high and low temperatures in 2050 are 27.7 °C and 21.5 °C, respectively. Most downstream areas are expected to endure high temperatures all year round, while sub-basin 29 will experience milder temperatures.

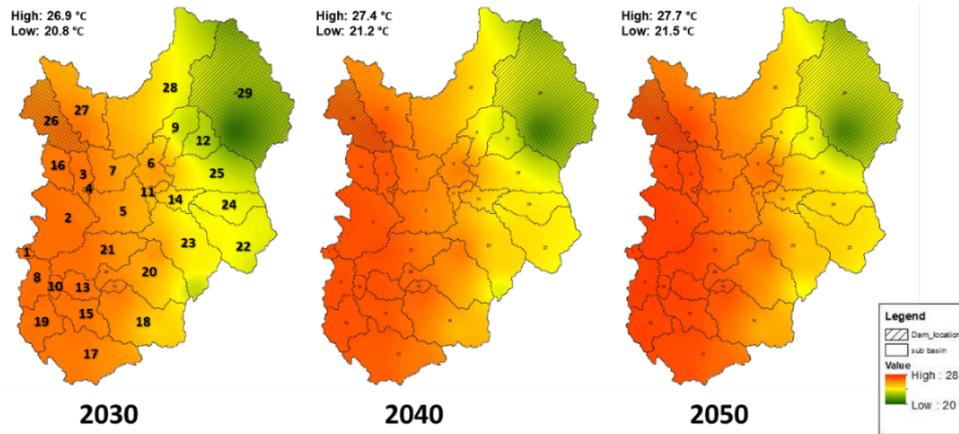


Figure 7-14: Temperature projection from the ‘ensemble average’ realisation

The ‘ensemble average’ in the monthly assessment is shown in Figure 7-15, which indicates that more precipitation is expected to occur in April and November, with November being the wettest month. Contrarily, less precipitation is anticipated in the months of January, February, and June, with less than 150 mm.

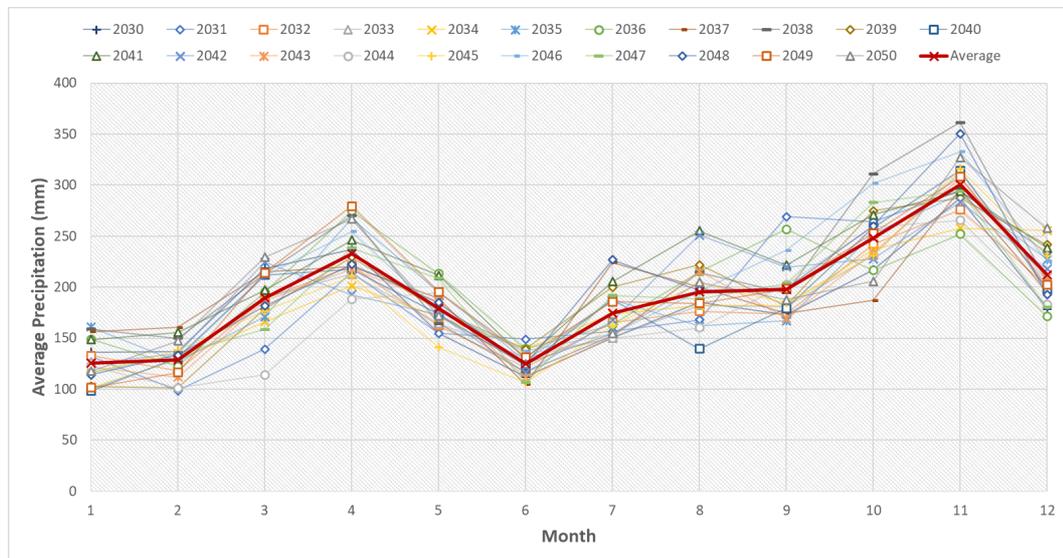


Figure 7-15: Average monthly precipitation from 2030 to 2050 on the ‘ensemble average’ realisation

The SRB temperature pattern is identical from 2030 to 2050, as illustrated in Figure 7-16. Temperatures are predicted to rise from January to June, and then fall from July until December. The maximum temperature anticipated by the ‘ensemble average’ realisation is around 26.25 °C, while the minimum temperature is approximately 25.25 °C.

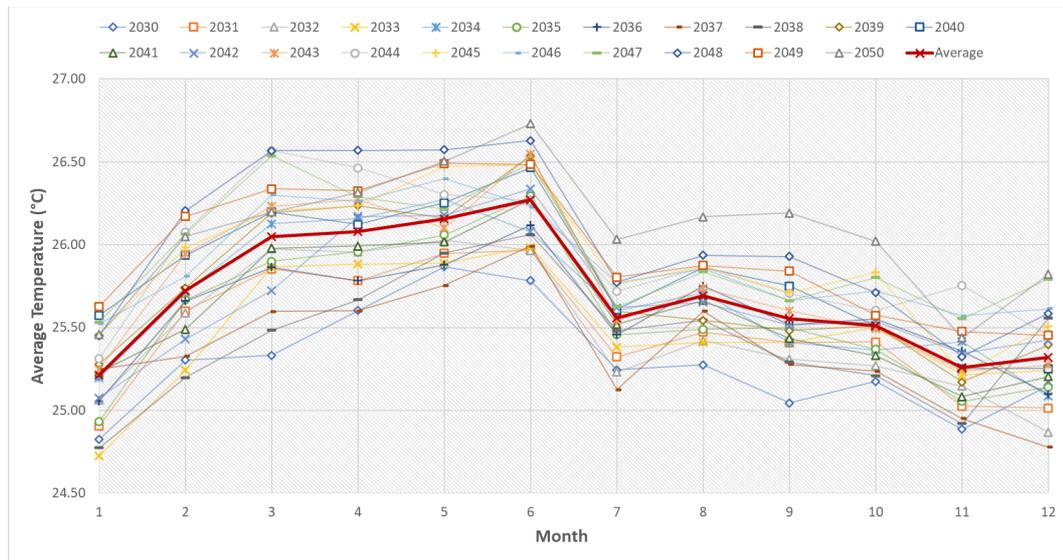


Figure 7-16: Average monthly temperature from 2030 to 2050 on the ‘ensemble average’ realisation

From the realisation of the ‘ensemble average’ simulation, Figure 7-17 illustrates the water availability in the SRB sub basin in 2030, 2040, and 2050. By the year 2050, more water is anticipated, with basins 1 to 5 projected to have more than 15 m<sup>3</sup>/s. Eleven sub basins are expected to have water availability of more than 5 m<sup>3</sup>/s in 2030, 14 subbasins in 2040, and 16 sub basins in 2050. Except for sub basins 16, 26, and 27, water availability in the SRB north area is consistent with more than 5 m<sup>3</sup>/s throughout the year, compared to less than 5 m<sup>3</sup>/s in the south area.

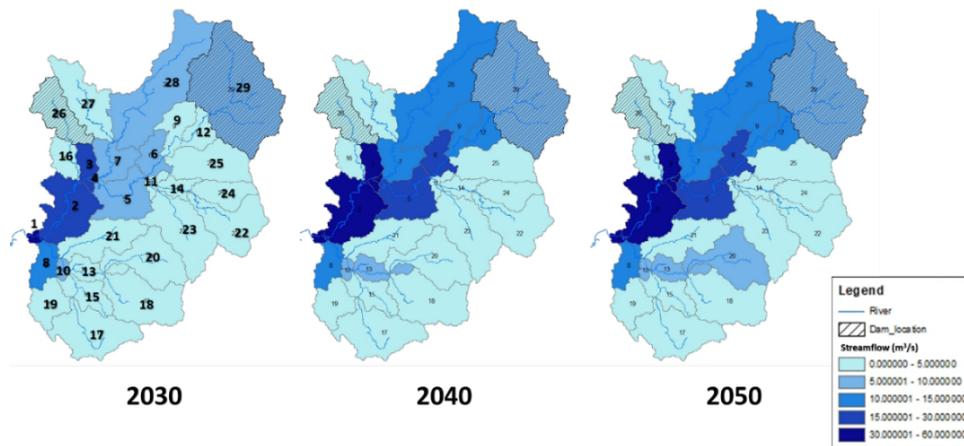


Figure 7-17: Water availability in the SRB sub basin using the ‘ensemble average’ realisation.

The ‘ensemble average’ projected datasets based on three sub basins known as outlet-1, outlet-4, and outlet-12 described in chapter 6 were used to simulate monthly flow at three streamflow stations. Figure 7-18 and Figure 7-19 depicts an identical trend with a different magnitude for streamflow projections at outlet-1 and outlet-4 throughout the year. For both outlet-1 and 4 in 2030, the flow peaks in April and November. In April, the streamflow at outlet-1 is expected to be 45.91 m<sup>3</sup>/s and 54.12 m<sup>3</sup>/s in November. While at outlet-4, streamflow is expected to be 25.84 m<sup>3</sup>/s and 30.77 m<sup>3</sup>/s, respectively. Meanwhile, Figure 7-20 shows the flow at outlet-12 is predicted to remain stable throughout the year, with streamflow of less than 3 m<sup>3</sup>/s. The expected streamflow in 2040 and 2050 is identical at all flow outlets, with the highest flow predicted in November and the lowest in February and June. The highest flow is around 80-81 m<sup>3</sup>/s at outlet-1, 53-54 m<sup>3</sup>/s at outlet-4 and 21-23 m<sup>3</sup>/s at outlet-12. In February and June, the lowest flow is projected to be around 36-39 m<sup>3</sup>/s at outlet-1, 20-25 m<sup>3</sup>/s at outlet-4, and 5-7 m<sup>3</sup>/s at outlet-12.

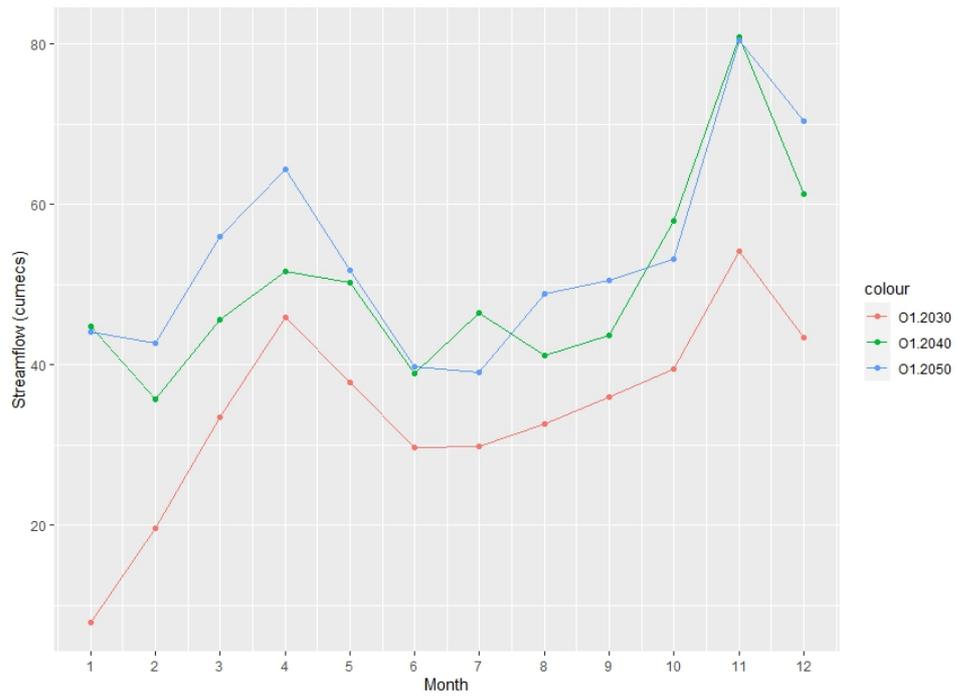


Figure 7-18: Mean monthly flow simulation on ‘ensemble average’ datasets for year 2030, 2040 and 2050 at outlet-1.

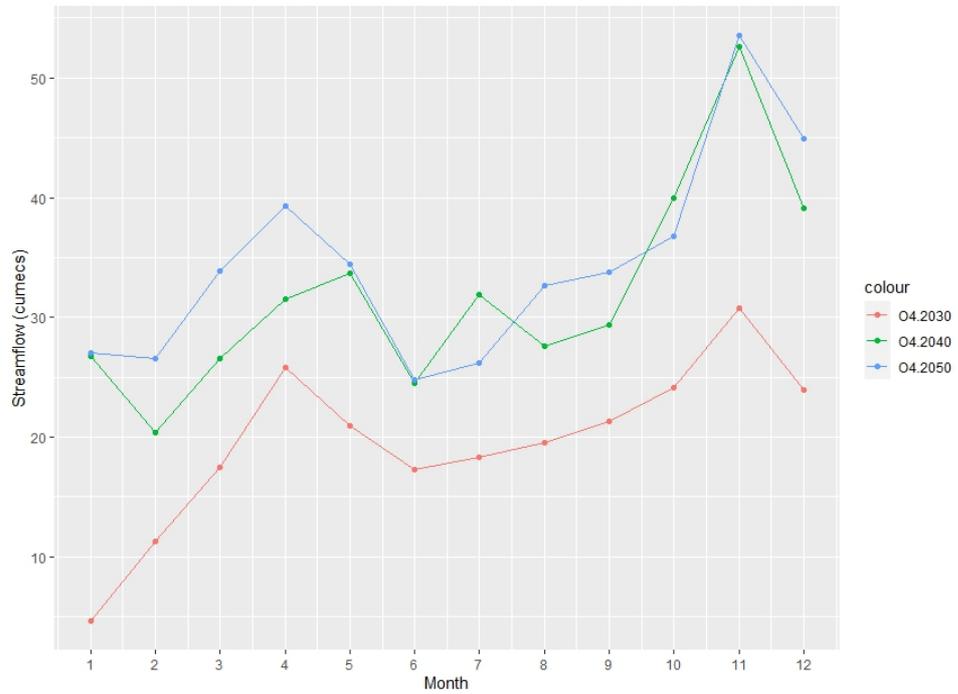


Figure 7-19: Mean monthly flow simulation on ‘ensemble average’ datasets for year 2030, 2040 and 2050 at outlet-4.

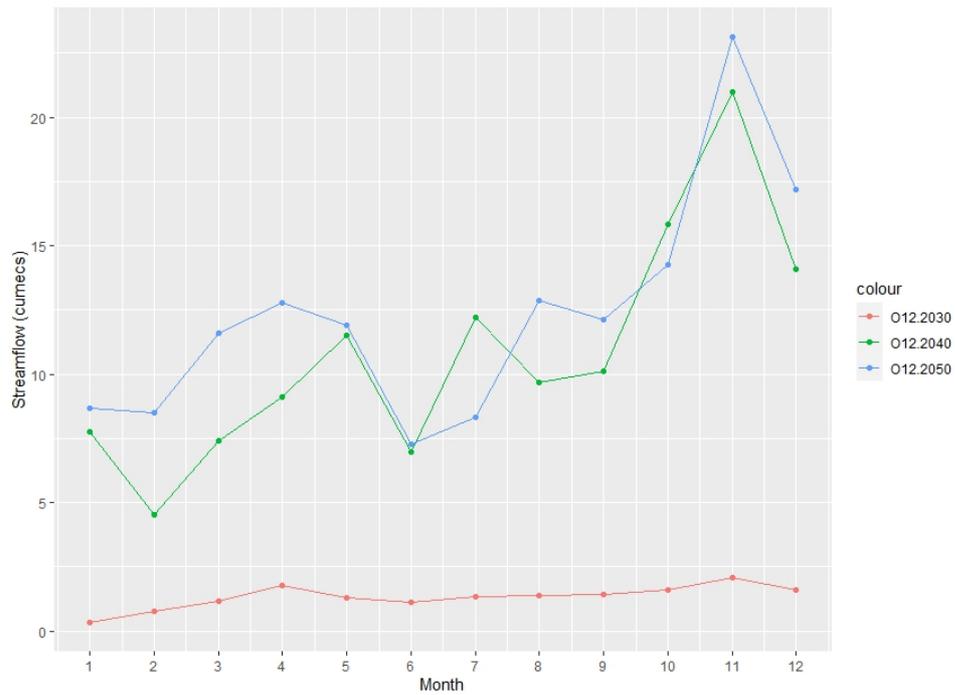


Figure 7-20: Mean monthly flow simulation on ‘ensemble average’ datasets for year 2030, 2040 and 2050 at outlet-12.

*SRES ccs3a1fi dataset simulation*

The precipitation ranges from 2030 to 2050 of SRES ccs3a1fi is about 400 to 1500 mm, as depicted in Figure 7-21, with the highest 1496 mm in 2046 and the lowest 415 mm in 2032. The trend analysis shows an increasing precipitation pattern in this scenario. In Figure 7-22, the Sen’s slope value is 20.192 with *p-value* and z value are 0.02 and 2.325 respectively.

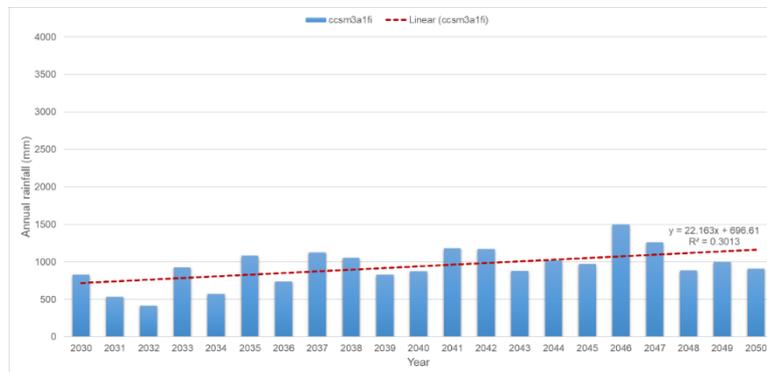


Figure 7-21: Linear trend analysis of precipitation for SRB using the SRES ccs3a1fi.

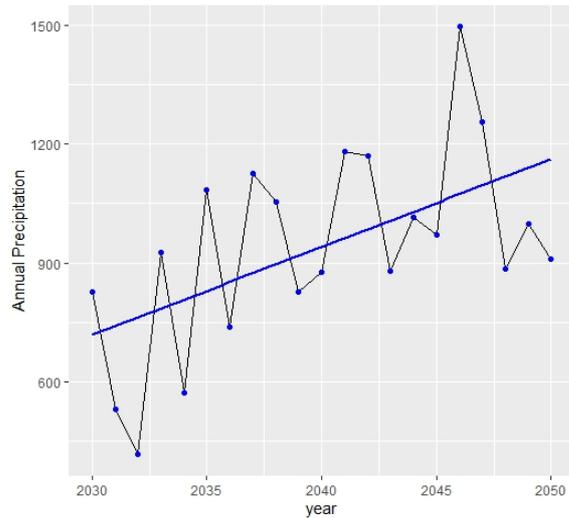


Figure 7-22: Annual rainfall variation of ‘ccsm3a1fi’ from 2030-2050 using the Mann-Kendall test with Sen’s slope.

Figure 7-23 depicts the precipitation projections for the years 2030, 2040 and 2050 based on SRES ccs3alfi realisation. Sub-basin 29 is anticipated to get the most precipitation, with 1,246 mm in 2030, 1163 mm in 2040, and 1242 mm in 2050. The lowest precipitation is expected at sub basin 17 with 402 mm in 2030, 701 mm in 2040, and at sub basin 19 with 497 mm in 2050. In this case, the area downstream of SRB is more likely to get less rain over the course of the year than the area upstream.

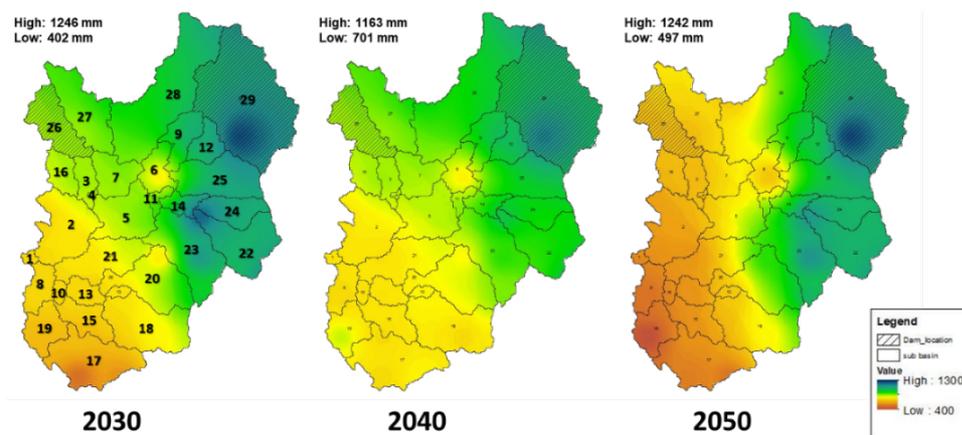


Figure 7-23: Precipitation projection from SRES ccs3alfi realisation

The rainfall anomaly for the simulation periods of ccsm3a1fi from 2030-2050 was depicted in Figure 7-24. The results predicted a negative rainfall anomaly in this realisation. The near normal was the most range on this realisation with rainfall anomaly predicted to occur in 2033, 2035, 2037, 2038, 2041, 2042, 2044, 2046, 2047, 2049, and 2050 with corresponding values of -0.18, 0.05, 0.05, 0.04, 0.08, 0.07, 0.01, 0.17, 0.09, 0.01, and -0.48. The extremely dry rainfall anomaly was predicted to occur in the years 2031, 2032, 2034, and 2036 having a value of -6.41, -8.37, -5.28, and -3.39 respectively.

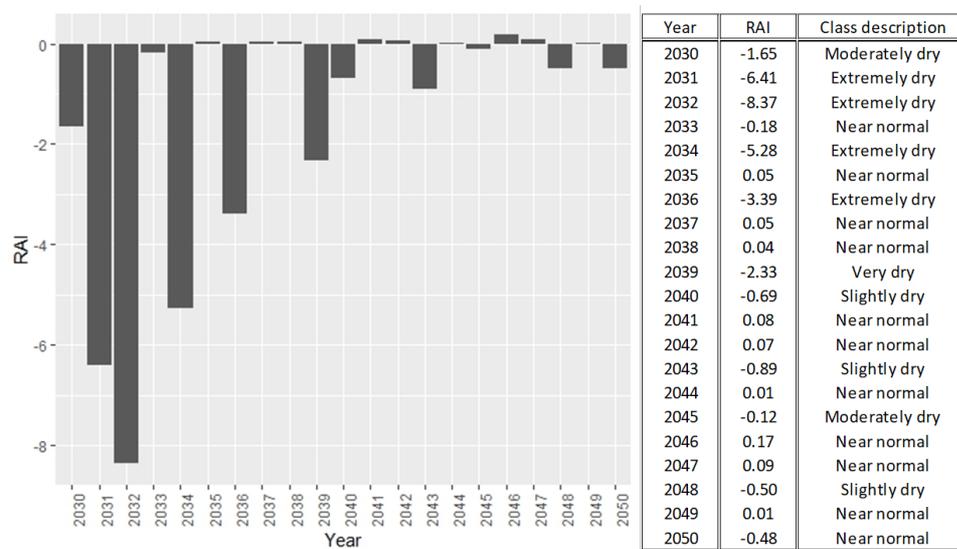


Figure 7-24: Rainfall Anomaly Index (RAI) prediction on the 'ccsm3a1fi' realisation.

The SRES ccsm3a1fi temperature projection trend study indicates an increasing pattern, with average temperatures increasing from 25.8 °C in 2030 to 26.1 °C in 2040 and 26.9 °C in 2050. The average temperature rises by around 1.1 degrees from 2030 to 2050 as depicted in Figure 7-25. Meanwhile in Mann-Kendall test with Sen's slope, the slope value is 0.0321 as shown in Figure 7-26 with *p-value* and z value are  $9.04 \times 10^{-6}$  and 4.439 respectively.

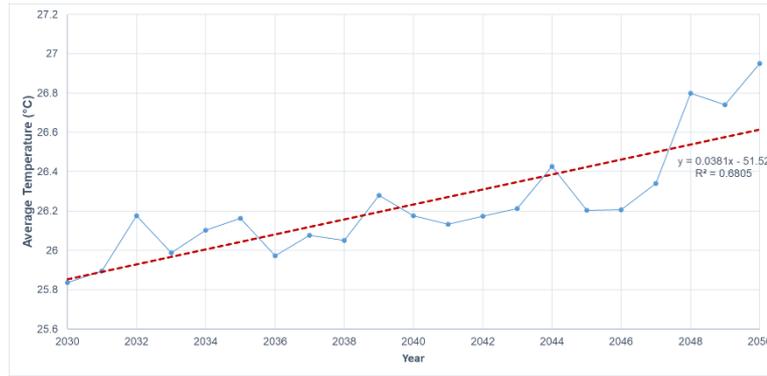


Figure 7-25: Linear trend analysis of temperature for the SRB using SRES ccsm3alfi.

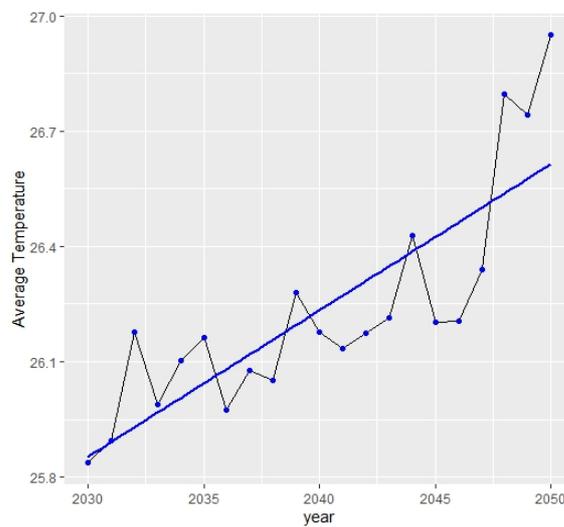


Figure 7-26: Average temperature variation on 'ccsm3a1fi' from 2030-2050 using the Mann-Kendall test with Sen's slope.

In temperature analysis, the highest temperature is expected to be at a downstream area of SRB throughout the year, with 27.7°C in 2030, 28°C in 2040, and 28.8°C in 2050. While the upstream area (sub basin 29) is expected to have the lowest temperature with 21°C in 2030, 21.3°C in 2040 and 21.9°C in 2050, as illustrated in Figure 7-27. From 2030 to 2050, the temperature in SRB is predicted to rise about  $\pm 1.1^\circ\text{C}$  in the maximum and  $\pm 0.9^\circ\text{C}$  in the minimum.

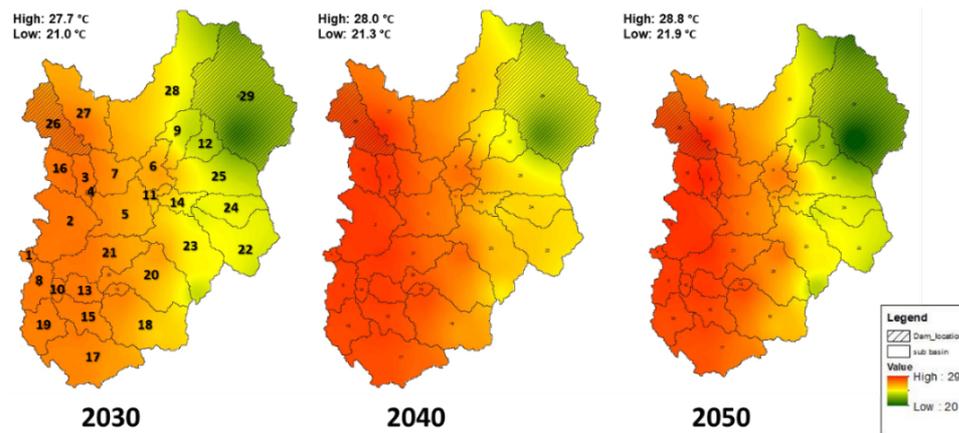


Figure 7-27: Temperature projection based on SRES ccsm3a1fi implementation.

In the SRES ccsm3a1fi realisation, October and November are predicted to have more precipitation than the other months, with more than 200 mm, as depicted in Figure 7-28. The average rainfall from January to September, including December, was expected to be less than 100 mm. The rainfall pattern from this realisation is different to 'ensemble average' where a minimum rainfall predicted in April with predict about 50 mm compared to 230 mm in 'ensemble average'. Even the highest rainfall in November is less about 100 mm compared to realisation from 'ensemble average'. As mentioned in Table 7-1, the ccsm3a1fi is the worst-case scenario compared to the other scenarios. This realisation is therefore anticipated to have the lowest projection when compared to the "ensemble average". According to this forecast, SRB will experience high temperatures and the least amount of rainfall from March through June.

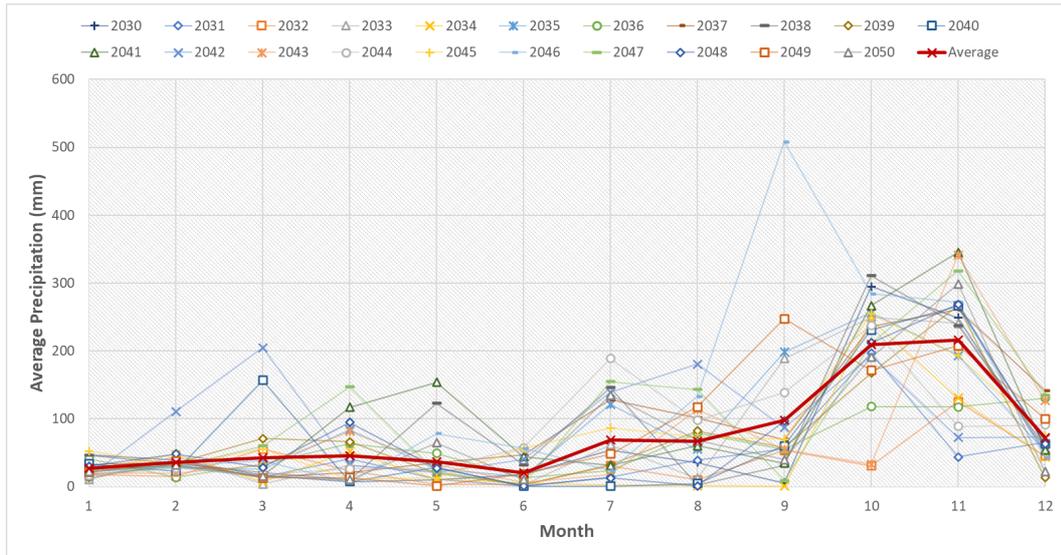


Figure 7-28: Monthly precipitation averages from 2030 to 2050 based on the SRES ccs3a1fi realisation.

For temperature prediction, most of the temperature patterns are identical throughout the year, with a slight variation in the temperature trends. From 2030 to 2050, the average temperature from SRES ccs3a1fi realisation was consistent at 25.5 °C to 26.5 °C, as shown in Figure 7-29. However, temperatures reached a high of 28 °C in June and December 2032, and a low of 24.5 °C in February 2031.

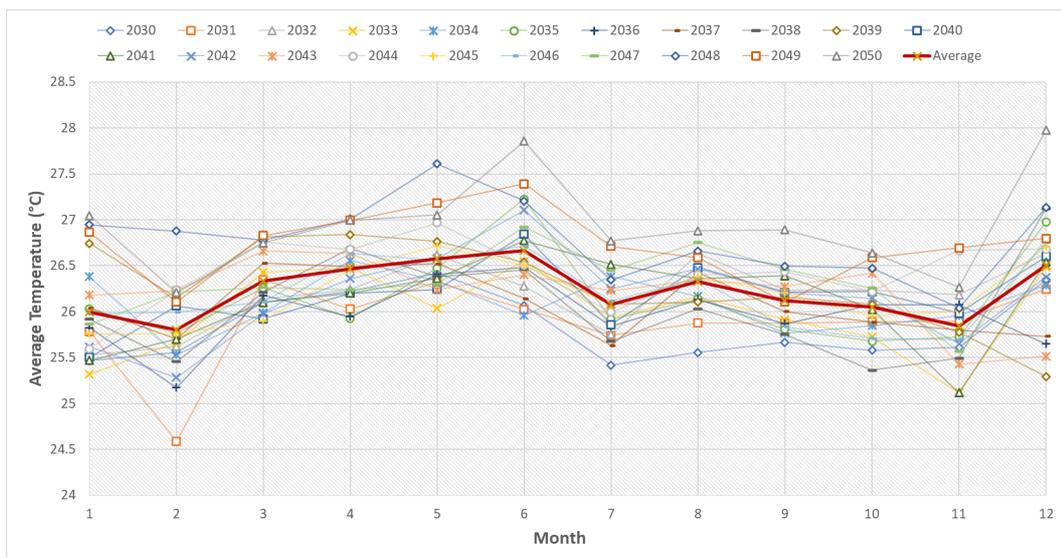


Figure 7-29: Monthly average temperature on the SRES ccs3a1fi realisation from 2030 to 2050.

Water availability simulations using the SRES ccsm3a1fi realisation demonstrate that in 2030, three of 29 sub basins will have more than 5 m<sup>3</sup>/s of water available. In 2040, seven sub basins with more than 5 m<sup>3</sup>/s will become available, and in 2050, nine sub basins will become available, as illustrated in Figure 7-30. In this simulation, water is more abundant in the downstream area of SRB than in the upstream area.

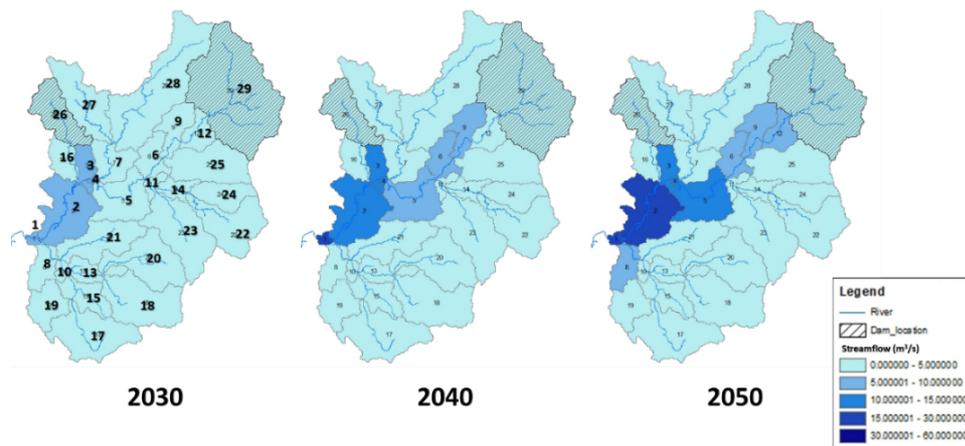


Figure 7-30: Water availability in the SRB sub basin using the SRES ccsm3a1fi realisation.

Figure 7-31 to Figure 7-33 illustrates a streamflow simulation using SRES ccsm3a1fi that displays an identical trend with a variable magnitude at three outlets, except for streamflow in 2030 at outlet-12, which is constant throughout the year. The highest projected flow at outlet-1 in 2030 is 50.7 m<sup>3</sup>/s in November, followed by 35.11 m<sup>3</sup>/s in October, 12.23 m<sup>3</sup>/s in December, and 8.22 m<sup>3</sup>/s in July. The projected streamflow at outlet-4 in the same year is 24.59 m<sup>3</sup>/s in October, followed by 23.44 m<sup>3</sup>/s in November, 9.18 m<sup>3</sup>/s in December, and the same flow magnitude as outlet-1 in July. Flow patterns in 2040 are projected to be similar, with peaks in Mac and November at all outlets. The peak flow at outlet-1 in March and November is 44.67 m<sup>3</sup>/s and 74.51 m<sup>3</sup>/s, respectively whereas the highest flows at outlet-4 and 12 are 34.33 and 38.92 m<sup>3</sup>/s and 11.24 and 15.31 m<sup>3</sup>/s, respectively. The flow patterns in 2050 also projected to be the highest in November with 82.37 m<sup>3</sup>/s at outlet-1, 46.76 m<sup>3</sup>/s at outlet-4 and 19 m<sup>3</sup>/s at outlet-12. Under SRES ccsm3a1fi datasets, streamflow in SRB is predicted to be less than 15 m<sup>3</sup>/s.

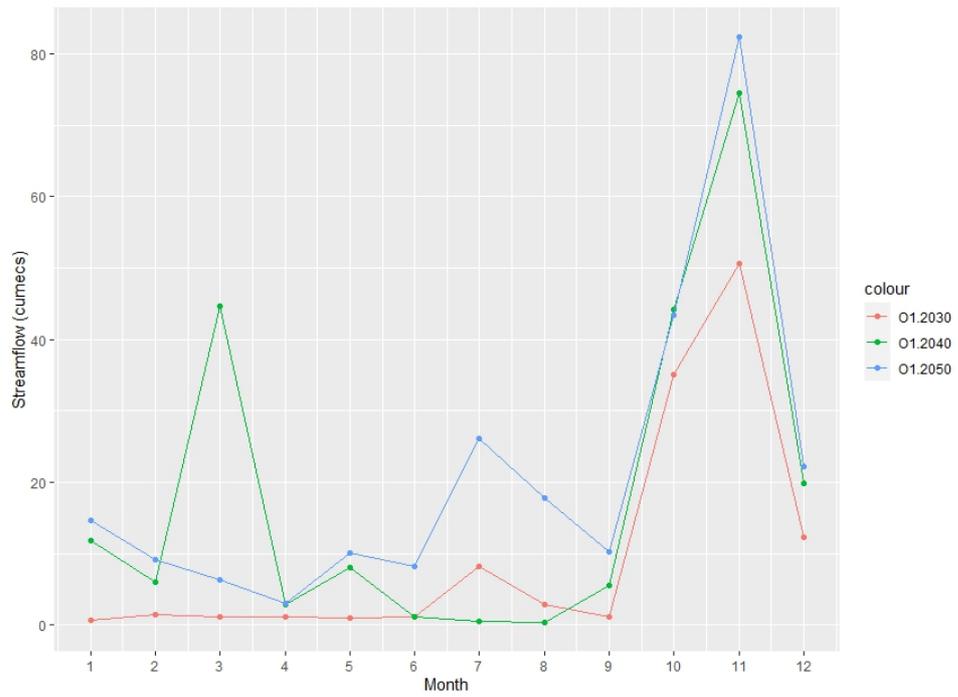


Figure 7-31: Mean monthly flow simulation on SRES ccs3alfi datasets for the years 2030, 2040 and 2050 at outlet-1.

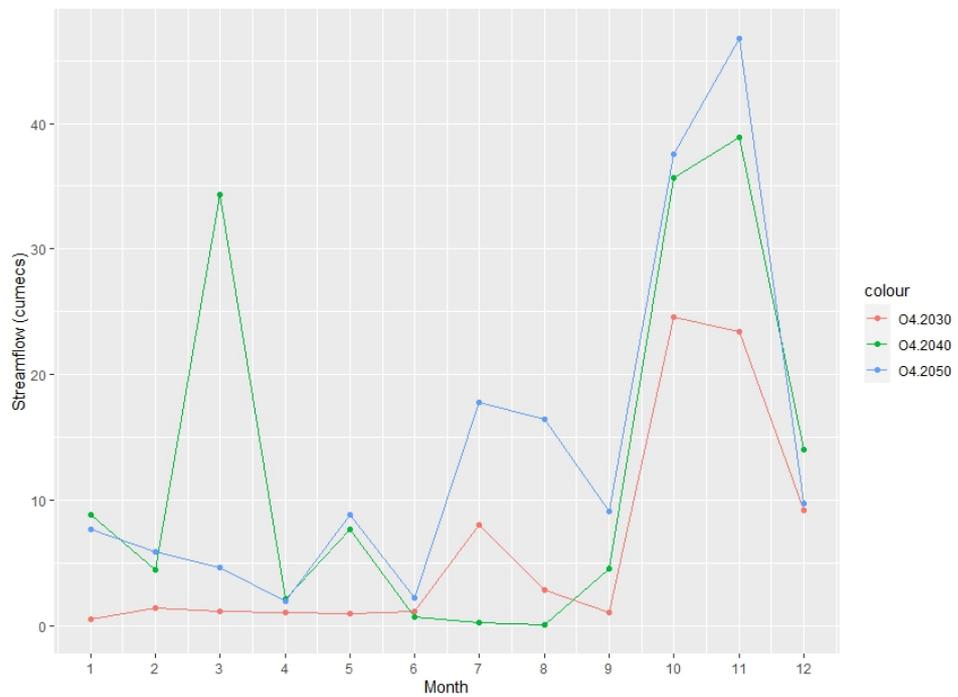


Figure 7-32: Mean monthly flow simulation on SRES ccs3alfi datasets for the years 2030, 2040 and 2050 at outlet-4.

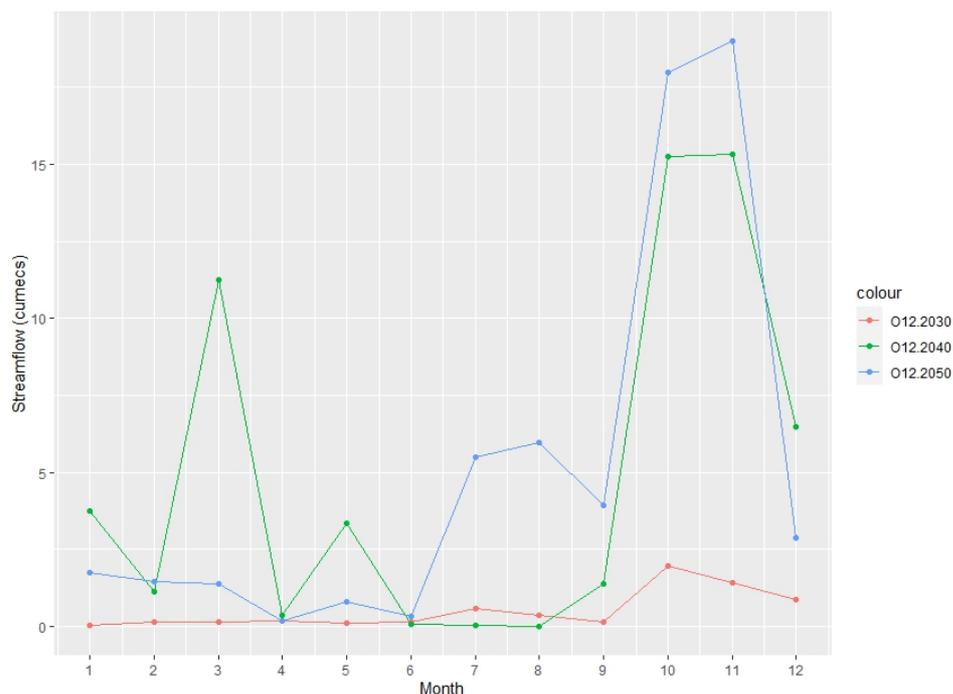


Figure 7-33: Mean monthly flow simulation on SRES ccsm3alfi datasets for the years 2030, 2040 and 2050 at outlet-12.

## 7.7 Summary of the results

As a result, analysis of precipitation and temperature data from the ‘ensemble average’ and SRES ccsm3a1fi realisations from 2030 to 2050 reveals a rising tendency. In the ‘ensemble average’ realisation, precipitation and temperature remain in the same range and pattern throughout the year. By that, SRB is expected to receive a huge amount of precipitation in November and some in April each year, as well as an elevated temperature range from February to June. In terms of area, the SRB upstream area will have less precipitation and warmer temperatures throughout the year compared to the downstream area. Therefore, more water is expected to be accessible downstream of SRB than upstream, as seen by the flow at outlet-12, which is predicted to be less than 3 m<sup>3</sup>/s in 2030.

In the worst-case scenario of the climate projection dataset (SRES ccsm3a1fi), the analysis of precipitation and temperature was separated into two zones. The upstream section of SRB is predicted to have a lot of precipitation and a low temperature, while the downstream area is predicted to have less precipitation and an elevated temperature throughout the year. Precipitation is expected to be higher in November and lower from January until June every year, similarly to the '*ensemble average*' realisation whereas the average temperature is about the same for each month. As a result, only a few sub basins are expected to have water availability for more than 5 m<sup>3</sup>/s in this realisation. Sub basin 1 to 4 have water availability of more than 5 m<sup>3</sup>/s in three consecutive years, while the rest, particularly in the southern part of the SRB, has less than 5 m<sup>3</sup>/s.

In precipitation analysis, the '*ensemble average*' and SRES ccsm3a1fi indicates that November has the most precipitation, while January, February, and June have the least. In terms of temperatures, both indicate that the elevated temperatures occur from February to June, while the lowest occur from August to November. Water is more abundant in the northern part of SRB compared to the southern part, according to simulations based on both realisations.

## CHAPTER 8 CHAPTER 8

### SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

#### 8.1 Summary

This chapter reports some research results that relate to the study's objectives. Some recommendations are made based on the results and also for future research.

##### 8.1.1 *Input data*

A realistic hydrological model outcome depends on how consistent and reliable the input data is. The quality of the input data used to describe the region is crucial for a successful rainfall-runoff model. The optimum input data is that which can represent the local weather with reasonable local spatial and temporal accuracy. Establishing a hydro-meteorological station that covers the whole region might be difficult in a developing country like Malaysia. Therefore, using satellite datasets as a complement to the limited number of gauged stations is another alternative. Input data for the SRB rainfall-runoff model was used from three different data sources in this study: gauged data, CMADS (China Meteorological Assimilation Driving Dataset), and CFSR (Climate Forecast System Reanalysis). In order to evaluate the daily and monthly scale data from 2008 to 2016, point-to-point analysis was utilised. To ensure that the input data was at the optimum level, several analyses were used, including the Double Mass Curve (DMC) method, error and missing data analysis, and bias correction. The findings of the Double Mass Curve analysis showed that the data on precipitation and streamflow were reliable, with an  $r^2$  nearly to 1.0. The error and missing data analysis reveals that 47% and 38% of the data are missing, for the maximum and minimum temperatures, respectively. The bias correction analysis reveals that the difference in precipitation data is  $\pm 30\%$ . The bias of CMADS is  $0.6^\circ\text{C}$  and  $0.4^\circ\text{C}$ , respectively, in the maximum and minimum temperature values, compared to  $0.7^\circ\text{C}$  and  $0.5^\circ\text{C}$ , respectively, for CFSR.

Evaluation of precipitation data from CMADS and CFSR showed both were highly significant to gauged data with a *p-value* less than 0.01. However, CMADS performed better on temperature data with an *r-value* of 0.6 (maximum) and 0.7 (minimum), compared to CFSR with 0.31 (maximum) and 0.18 (minimum). Meanwhile, the Mean Absolute Error (MAE) and the Root-Mean-Square Error (RMSE) results showed the CMADS outperformed the CFSR in terms of maximum and minimum temperature by obtaining lower values. In addition to single-source datasets, cross-combined datasets were assessed. Seven scenarios were developed and evaluated in the SWAT (Soil and Water Assessment Tool) model for the optimum results based on the three data sources. The observed datasets outperformed the global reanalysis datasets (CMADS and CFSR) for single-source datasets analysis, with 0.53 of  $R^2$ , 0.45 of NSE, 0.00 of Pbias, 0.72 of KGE, and 0.74 of RSR. Evaluation of the cross-combined datasets showed a good performance rating in  $R^2$ , NSE, and KGE. In rainfall-runoff modelling, cross-combined datasets outperformed single-source datasets, which is an excellent method to boost model performance.

### **8.1.2 Calibration and validation techniques**

The single-site calibration (SSC) method is commonly used to simulate streamflow at the catchment outlet with the assumption that the spatial heterogeneity of the model is uniform throughout the catchment. Meanwhile, the multi-site calibration (MSC) method is a new approach to enhancing the calibration process by dividing a large watershed into smaller watersheds and obtaining the parameters from each of the smaller watersheds. In the MSC method, three techniques were - basin-by-basin, simultaneous, and sequential. Since scenario 5 (SC-5) had the optimum result when compared to the other scenarios, it is utilised as the input data in MSC modelling. Five statistical analyses, including the Kling Gupta Efficiency (KGE), Nash-Sutcliffe Efficiency (NSE), coefficient of determination ( $R^2$ ), Percent Bias (Pbias), and the RMSE-observation standard deviation ratio (RSR), were used to assess the model's performance. Additionally, a comparison between SSC and MSC was evaluated.

In the MSC method, sequential technique outperformed simultaneous and basin-by-basin, achieving a 'satisfactory' range for most of the objective functions at outlets-1 and 4. Despite the fact that outlet-12 shows an 'unsatisfactory' range in this technique, the performance rating results are still better than the other techniques. The simultaneous technique performed as well as sequential, but poorly performed in uncertainty analysis with an *r-factor* higher than 1.5 at all outlets. Furthermore, the regionalisation method applied at outlet-4 revealed a highly significant difference between the calibrated parameters from basin-1 compared to basin-3. In fact, it is identical to the 2014-2016 calibrated model parameters in basin-2.

The SSC effectively performed at outlet-1, achieving four objective functions beyond the minimum level (0.51 of  $R^2$ , 0.46 of NSE, 0.10 of Pbias, and 0.68 of KGE). However, when using the basin-by-basin technique, the performance ratings at outlet-1 dramatically dropped, with 0.36 of  $R^2$ , 0.19 of NSE, and 0.59 of KGE. In fact, the results at outlet-4 and 12 did not reach the minimum range in basin-by-basin technique. Compared to sequential and simultaneous techniques, the objective function results were better than basin-by-basin. In fact, sequential achieved a 'satisfactory' result at outlet-1 with 0.46 of  $R^2$ , 0.27 of NSE, -11.7 of Pbias, 0.65 of KGE, and 0.86 of RSR. As a result, the MSC method is more reliable than the SSC method for determining the model accuracy.

### **8.1.3 Climate change assessment**

The hydrologic cycle is significantly impacted by climate change, which also increases uncertainty in water availability at all scales. According to the IPCC's AR6 study (Sixth Assessment Report of the Intergovernmental Panel on Climate Change), drought conditions in Asian countries may worsen by 5-20%, by the end of this century. In fact, since the 20<sup>th</sup> century, Asia's surface air temperature has been rising, and intense precipitation events have increased the frequency of floods. The purpose of this research is to evaluate the impacts of climate change on streamflow in the Selangor River basin (SRB), since there is limited study on this topic in Malaysia,

In order to predict the surface hydrologic conditions, this study employs the optimum SWAT hydrological model coupled with the ensemble data of 15 realisations dynamically downscaled from the RegHCM-PM model. Analysis of 15 realisations shows a rising trendline pattern of precipitation and temperature data from 2030 to 2050. The annual average temperature will vary between 24.5 and 27 degrees Celsius, with the SRES (Special Report on Emission Scenarios) mria1b projected to be the higher temperature and SRES ccsm3b1 the lowest. The projection of precipitation from the '*ensemble average*' shows the SRB is expected to receive a huge amount of precipitation in November and some in April each year. Meanwhile, the SRB average temperature trend shows a yearly rise from February to a high of 26.3°C in June and a decline to 25.5°C from July to December. Based on Rainfall Anomaly index (RAI), SRB expected to have an extremely dry and very dry condition in 2030 and 2040. This trend will impact on water resources and worsen the water crisis in SRB as projected water demand will be increased to 306 mm/year in 2030 and 328.7 in 2040.

In terms of area, the SRB upstream area will have less annual precipitation and higher temperatures compared to the downstream area. The analysis of precipitation and temperature is separated into two zones in the worst-case scenario of the climate projection dataset. The SRB's upstream region is expected to have heavy precipitation and low temperatures throughout the year, while the downstream region is expected to have less precipitation and high temperatures. As a result, only a few sub basins are expected to have water availability for more than 5 m<sup>3</sup>/s in this realisation. The sub-basin 1 to 4 have water availability of more than 5 m<sup>3</sup>/s in three consecutive years, while the rest, particularly in the southern part of the SRB, has less than 5 m<sup>3</sup>/s. The '*ensemble average*' and SRES ccsm3a1fi results of precipitation analysis shows that November has the highest precipitation, while January, February, and June have the least. In terms of temperature, both indicate that the highest temperatures occur between February and June, and the lowest between August and November. The simulation based on both realisations show that water is more plentiful in the northern part of SRB than the southern part particularly in sub-basin 1 to 6 and sub-basin 9.

## 8.2 Conclusions

The main conclusions from the study's results are as follows:

- Global reanalysis datasets (CMADS and CFSR) are reliable to be used in Malaysia environment with some adjustment particularly on temperature data. The CMADS datasets are highly significant to gauged datasets compared to CFSR.
- As input data to the SWAT model, evaluation of single-source datasets shows CMADS outperforms CFSR in four objective functions, which are  $R^2$ , NSE, Pbias and KGE.
- Evaluation on the cross-combined datasets showed the results were excellent compared to the single-source datasets with  $R^2$ , NSE, Pbias and KGE above the minimum model performance range. The combination of gauged and CMADS datasets (SC-5) was the optimum input data with an achieved 0.51 of  $R^2$ , 0.46 of NSE, 0.10 of Pbias, 0.68 of KGE, and 0.74 of RSR.
- The single-site calibration (SSC) method achieves a 'satisfactory' result at outlet-1, with 0.51 of  $R^2$ , 0.46 of NSE, 0.10 of Pbias, 0.68 of KGE, and 0.74 of RSR. However, in the basin-by-basin technique, the performance ratings drastically drop to 0.36 of  $R^2$ , 0.19 of NSE, 0.59 of KGE, and 0.90 of RSR.
- The multi-site calibration (MSC) method shows highly significant results compared to the SSC. The sequential and simultaneous techniques have an identical result in performance ratings. However, the simultaneous technique in uncertainty analysis of *p-factor* and *r-factor*.
- The regionalisation method was applied in basin-2 due to data limitation. The results reveal that the calibrated parameter from basin-1 is highly significant compared to basin-3. In fact, it is identical to the 2014-2016 calibrated model parameters in basin-2.

- The simulation on 15 realisations showed a rising trendline pattern of precipitation and temperature data from 2030 to 2050, with the SRB upstream area having less annual precipitation and higher temperatures compared to the downstream area based on the 'ensemble average' data projection.
- The SRES ccs3a1fi projection determines that SRB's upstream region is expected to have heavy precipitation and low temperatures throughout the year, while the downstream region is expected to have less precipitation and high temperatures.
- Both realisations show that water is more plentiful in the northern part of SRB than in the southern part, particularly in sub-basins 1 to 6 and sub-basin 9.

### **8.3 Contributions of this study**

- This study coupled gauged and gridded datasets (CMADS and CFSR) using five weather elements, including precipitation, temperature, relative humidity, solar radiation, and wind speed, as input data in the SRB SWAT model. The results show an improvement in modelling performance compared to single-source datasets. Other regions with a lack of data may use the approach framework that has been created.
- This is the first study to examine the suitability of global reanalysis datasets for the Selangor River basin, which supplies approximately 67% of the water to Kuala Lumpur City Centre and the state of Selangor in Malaysia. The findings help to clarify how gridded datasets can improve model performance in data-scarce regions.
- The evaluation of multi-site calibration (MSC) method was done for the first time on SRB.
- For the first time, the effects of climate change on SRB water resources were assessed using multiple General Circulation Models (GCMs), which take a wide range of uncertainty into account when making predictions about the future of temperature and precipitation and, consequently, the availability of water.

## **8.4 Recommendations**

The following suggestions are made based on the study's findings:

### **8.4.1 Based on the study results**

- To get the best results from rainfall-runoff modelling, it is suggested to add more sources to the cross-combined datasets.
- Water managers are encouraged to use the multi-site calibration method rather than the single-site calibration method since it shows great efficacy in rainfall-runoff models.
- For future development of water resources, water managers must consider the climate change scenarios in planning and managing water resources based on the projected water availability.

### **8.4.2 For future research**

- This study uses two global reanalysis datasets as input data in a rainfall-runoff model. It is recommended to further diversify other data that can enhance model performance, increase model accuracy, and reduce uncertainty.
- There are two large dams operating in the SRB, and this study assumes that the inflow and outflow of the dam area are normal like in other sub-basins. It is suggested to include information on dam operation in order to obtain a good rainfall-runoff model and expand the analysis to the agricultural region in SRB.
- This study used 2008 to 2016 datasets to develop the SWAT model. It is suggested to use the latest available data in order to obtain a relevant rainfall-runoff model that represents the catchment condition.
- To obtain the optimum model, the most recent landscape features data, including soil type, land use, and Digital Elevation Model (DEM) should be included.
- This study assesses the climate change impact on SB water resources using the Fourth Assessment Report (AR4). The results can be updated since the Sixth Assessment Report (AR6) has been published.

## REFERENCES

- Abbaspour, K. C., Johnson, C. A., & van Genuchten, M. T. (2004). Estimating Uncertain Flow and Transport Parameters Using a Sequential Uncertainty Fitting Procedure. *Vadose Zone Journal*, 3(4), 1340–1352. <https://doi.org/10.2136/vzj2004.1340>
- Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H., & Kløve, B. (2015). A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of Hydrology*, 524, 733–752. <https://doi.org/10.1016/j.jhydrol.2015.03.027>
- Abbaspour, Karim C., Vaghefi, S. A., & Srinivasan, R. (2017). A guideline for successful calibration and uncertainty analysis for soil and water assessment: A review of papers from the 2016 international SWAT conference. *Water (Switzerland)*, 10(1). <https://doi.org/10.3390/w10010006>
- Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986). An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology*, 87(1–2), 45–59. [https://doi.org/10.1016/0022-1694\(86\)90114-9](https://doi.org/10.1016/0022-1694(86)90114-9)
- Albert, J. M. (2004). *Hydraulic Analysis and Double Mass Curves of the Middle Rio Grande from Cochiti to San Marcial, New Mexico*. 1–224.
- Amin, M. Z. M., Shaaban, A. J., Ohara, N., Kavvas, M. L., Chen, Z. Q., Kure, S., & Jang, S. (2016). Climate Change Assessment of Water Resources in Sabah and Sarawak, Malaysia, Based on Dynamically-Downscaled GCM Projections Using a Regional Hydroclimate Model. *Journal of Hydrologic Engineering*, 21(1), 05015015. [https://doi.org/10.1061/\(asce\)he.1943-5584.0001242](https://doi.org/10.1061/(asce)he.1943-5584.0001242)
- Anang, Z., Padli, J., Kamaludin, M., & Sathasivam, S. (2017). The Effect of Climate Change on Water Resources Using Panel Approach: The Case of Malaysia. *International Journal of Academic Research in Business and Social Sciences*, 7(11), 141–152. <https://doi.org/10.6007/ijarbss/v7-i11/3446>
- Anderson, M. L., Chen, Z. Q., Kavvas, M. L., & Yoon, J. Y. (2007). Reconstructed Historical Atmospheric Data by Dynamical Downscaling. *Journal of Hydrologic Engineering*, 12(2), 156–162. [https://doi.org/10.1061/\(asce\)1084-0699\(2007\)12:2\(156\)](https://doi.org/10.1061/(asce)1084-0699(2007)12:2(156))
- Anderton, S., Latron, J., & Gallart, F. (2002). Sensitivity analysis and multi-response, multi-criteria evaluation of a physically based distributed model. *Hydrological Processes*, 16(2), 333–353. <https://doi.org/10.1002/hyp.336>
- Ang, R., & Oeurng, C. (2018). Simulating streamflow in an ungauged catchment of Tonlesap Lake Basin in Cambodia using Soil and Water

- Assessment Tool (SWAT) model. *Water Science*, 32(1), 89–101. <https://doi.org/10.1016/j.wsj.2017.12.002>
- Arnold, J. G., & Fohrer, N. (2005). SWAT2000: Current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes*, 19(3), 563–572. <https://doi.org/10.1002/hyp.5611>
- Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., Santhi, C., Harmel, R. D., Van Griensven, A., Van Liew, M. W., Kannan, N., & Jha, M. K. (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), 1491–1508. <https://doi.org/10.1029/2002WR001340>; Bradford, S.A., Šimůnek, J., Bettehar, M., Van, G.M.Th., Yates, S.R., Modeling colloid attachment, straining, and exclusion in saturated porous media (2003) *Environ. Sci. Tech.*, 37 (10), pp. 2242-2250; Bradford, S.A., Bettehar, M., Šimůnek, J., Van Genuchten, M.Th., Straining and attachment of colloids in physically heterogeneous porous media (2004) *Vadose Zone J.*, 3 (2), pp. 384-394; Brooks, R.H., Corey, A.T., (1964) *Hydraulic Properties of Porous Media*, Hydrology Paper
- Arnold, Kiniry, Srinivasan, Williams, Haney, & Neitsch. (2012). *Soil & Water Assessment Tool*.
- Awange, J. L., Hu, K. X., & Khaki, M. (2019). The newly merged satellite remotely sensed, gauge and reanalysis-based Multi-Source Weighted-Ensemble Precipitation: Evaluation over Australia and Africa (1981–2016). *Science of the Total Environment*, 670, 448–465. <https://doi.org/10.1016/j.scitotenv.2019.03.148>
- Azoulay, A., & Hougbo, G. F. (2018). The united nations world water development report 2018: Nature-based solutions for water. In *UN Water Report*. <https://doi.org/https://unesdoc.unesco.org/ark:/48223/pf0000261424>
- Bai, J., Shen, Z., & Yan, T. (2017). A comparison of single- and multi-site calibration and validation: a case study of SWAT in the Miyun Reservoir watershed, China. *Frontiers of Earth Science*, 11(3), 592–600. <https://doi.org/10.1007/s11707-017-0656-x>
- Beven, K. (2012). Rainfall-Runoff Modelling. In *Rainfall-Runoff Modelling*. <https://doi.org/10.1002/9781119951001>
- Blöschl, G. (2016). Predictions in ungauged basins – Where do we stand? *IAHS-AISH Proceedings and Reports*, 373, 57–60. <https://doi.org/10.5194/piahs-373-57-2016>
- Boelee, L., Bahrom, R., Amer, H., Sondor, N. Z., Brown, E., Ahmad, F., & Wilson, G. (2017). Operational decision support system for sustainable water resource management for Sungai Selangor. *37th IAHR World Congress*, 1–16.
- Bressiani, D. de A., Gassman, P. W., Fernandes, J. G., Garbossa, L. H. P., Srinivasan, R., Bonumá, N. B., & Mendiondo, E. M. (2015). A review of soil and water assessment tool (SWAT) applications in Brazil: Challenges

- and prospects. *International Journal of Agricultural and Biological Engineering*, 8(3), 1–27. <https://doi.org/10.3965/j.ijabe.20150803.1765>
- Brighenti, T. M., Bonumá, N. B., Grison, F., Mota, A. de A., Kobiyama, M., & Chaffe, P. L. B. (2019). Two calibration methods for modeling streamflow and suspended sediment with the swat model. *Ecological Engineering*, 127(May 2018), 103–113. <https://doi.org/10.1016/j.ecoleng.2018.11.007>
- Cao, Y., Zhang, J., Yang, M., Lei, X., Guo, B., Yang, L., Zeng, Z., & Qu, J. (2018). Application of SWAT model with CMADS data to estimate hydrological elements and parameter uncertainty based on SUFI-2 algorithm in the Lijiang River basin, China. *Water (Switzerland)*, 10(6). <https://doi.org/10.3390/w10060742>
- Carroll, S., Liu, A., Dawes, L., Hargreaves, M., & Goonetilleke, A. (2013). Role of Land Use and Seasonal Factors in Water Quality Degradations. *Water Resources Management*, 27(9), 3433–3440. <https://doi.org/10.1007/s11269-013-0356-6>
- Castaneda-Gonzalez, M., Poulin, A., Romero-Lopez, R., Arsenault, R., Brissette, F., Chaumont, D., & Paquin, D. (2018). Impacts of Regional Climate Model Spatial Resolution on Summer Flood Simulation. *HIC 2018. 13th International Conference on Hydroinformatics*, 3(November), 372–362. <https://doi.org/10.29007/hd8l>
- Chen, Z. Q., Kavvas, M. L., Ohara, N., Amin, M. Z. M., & Ahmad, J. S. (2006). Regional hydroclimate model of Peninsular Malaysia. *Seminar-Impact of Climate ...*, January 2006. <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:REGIONAL+HYDROCLIMATE+MODEL+OF+PENINSULAR+MALAYSIA#0>
- Chinnasamy, P., & Ganapathy, R. (2018). Long-term variations in water storage in Peninsular Malaysia. *Journal of Hydroinformatics*, 20(5), 1180–1190. <https://doi.org/10.2166/hydro.2017.043>
- Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., Chang, P., Doney, S. C., Hack, J. J., Henderson, T. B., Kiehl, J. T., Large, W. G., Mckenna, D. S., Santer, B. D., & Smith, R. D. (2006). The Community Climate System Model, version 3 (CCSM3). *Journal of Climate*, 19(19), 2122–2143. [https://doi.org/10.1175/1520-0442\(2004\)017<3666:TCCSMV>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<3666:TCCSMV>2.0.CO;2)
- Combalicer, E. A., Cruz, R. V. O., Lee, S. H., & Im, S. (2010). Modelling hydrologic processes distribution in a tropical forest watershed in the philippines. *Journal of Tropical Forest Science*, 22(2), 155–169.
- Cruz, F. T., Narisma, G. T., Dado, J. B., Singhruck, P., Tangang, F., Linarka, U. A., Wati, T., Juneng, L., Phan-Van, T., Ngo-Duc, T., Santisirisomboon, J., Gunawan, D., & Aldrian, E. (2017). Sensitivity of temperature to physical parameterization schemes of RegCM4 over the CORDEX-Southeast Asia region. *International Journal of Climatology*, 37(15), 5139–5153. <https://doi.org/10.1002/joc.5151>
- D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, & T.

- L. Veith. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, 50(3), 885–900. <https://doi.org/10.13031/2013.23153>
- Dao, D. M., Lu, J., Chen, X., Kantoush, S. A., Van Binh, D., Phan, P., & Tung, N. X. (2021). Predicting tropical monsoon hydrology using CFSR and CMADS data over the Cau river basin in Vietnam. *Water (Switzerland)*, 13(9). <https://doi.org/10.3390/w13091314>
- Datagy. (2022). *How to Calculate Mean Absolute Error (MAE) in Python*. Python in 30 Days Course. <https://datagy.io/mae-python/>
- Deng, P., Zhang, M., Bing, J., Jia, J., & Zhang, D. (2019). Evaluation of the GSMaP\_Gauge products using rain gauge observations and SWAT model in the Upper Hanjiang River Basin. *Atmospheric Research*, 219(January), 153–165. <https://doi.org/10.1016/j.atmosres.2018.12.032>
- Desa M, M. N., Noriah, A. B., & Rakhecha, P. R. (2001). Probable maximum precipitation for 24h duration over southeast Asian monsoon region - Selangor, Malaysia. *Atmospheric Research*, 58(1), 41–54. [https://doi.org/10.1016/S0169-8095\(01\)00070-9](https://doi.org/10.1016/S0169-8095(01)00070-9)
- Desai, S., Singh, D. K., Islam, A., & Sarangi, A. (2021). Multi-site calibration of hydrological model and assessment of water balance in a semi-arid river basin of India. *Quaternary International*, 571(September 2019), 136–149. <https://doi.org/10.1016/j.quaint.2020.11.032>
- Di Baldassarre, G., Elshamy, M., Van Griensven, A., Soliman, E., Kigobe, M., Ndomba, P., Mutemi, J., Mutua, F., Moges, S., Xuan, Y., Solomatine, D., & Uhlenbrook, S. (2011). Future hydrology and climate in the River Nile basin: a review. *Hydrological Sciences Journal*, 56(2), 199–211. <https://doi.org/10.1080/02626667.2011.557378>
- Dickinson, R. E., Errico, R. M., Giorgi, F., & Bates, G. T. (1989). A Regional Climate Model for the Western United States. *Climatic Change*, 15, 383–422.
- DID. (2009a). *Hydrology and Water Resources* (Vol. 4).
- DID. (2009b). River Management. In *DID Manual* (Vol. 2). Department of Irrigation and Drainage Malaysia.
- DID. (2011a). *Executive Summary: Review of the National Water Resources (2000-2050) and Formulation of National Water Resources Policy* (Vol. 5).
- DID. (2011b). Selangor, Federal Territory of Kuala Lumpur & Putrajaya. In *Review of The National Water Resources Study (2000-2050) and Formulation of National Water Resources Policy* (Vol. 14, Issue v). Ministry of Natural Resources and Environment Malaysia.
- DID. (2011c). *Selangor, Federal Territory of Kuala Lumpur and Putrajaya: Review of the National Water Resources Study (2000-2050) and Formulation of National Water Resources Policy* (Vol. 14, Issue August).
- DID. (2021). *Kompendium*.

[http://www.websm.org/uploadi/editor/Reja\\_2003\\_open\\_vs\\_close-ended\\_questions.pdf](http://www.websm.org/uploadi/editor/Reja_2003_open_vs_close-ended_questions.pdf)

- Dile, Y. T., & Srinivasan, R. (2014). Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: An application in the blue Nile river basin. *Journal of the American Water Resources Association*, 50(5), 1226–1241. <https://doi.org/10.1111/jawr.12182>
- Dlamini, N. S., Rowshon, M. K., Fikhri, A., Lai, S. H., & Mohd, M. S. F. (2017). Modelling the streamflow of a river basin using enhanced hydro-meteorological data in Malaysia. *Acta Horticulturae*, 1152(April), 291–297. <https://doi.org/10.17660/ActaHortic.2017.1152.39>
- Dlamini, Nkululeko Simeon, Kamal, M. R., Soom, M. A. B. M., Mohd, M. S. F., bin, Abdullah, A. F. Bin, & Hin, L. S. (2017). Modeling potential impacts of climate change on streamflow using projections of the 5th assessment report for the Bernam river basin, Malaysia. *Water (Switzerland)*, 9(3), 1–22. <https://doi.org/10.3390/w9030226>
- Doblas-Reyes, F.J., Sörensson, A. A., Almazroui, M., Dosio, A., Gutowski, W. J., Haarsma, R., Hamdi, R., Hewitson, B., Kwon, W.-T., Lamptey, B. L., Maraun, D., Stephenson, T. S., Takayabu, I., Terray, L., Turner, A., & Z. Zuo. (2021). *Linking Global to Regional Climate Change. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* <https://doi.org/10.1017/9781009157896.012>
- Du Plessis, A. (2017). Global Context of Freshwater Resources. *Freshwater Challenges of South Africa and Its Upper Vaal River*, 3–11. [https://doi.org/10.1007/978-3-319-49502-6\\_1](https://doi.org/10.1007/978-3-319-49502-6_1)
- Duan, Z., Tuo, Y., Liu, J., Gao, H., Song, X., Zhang, Z., Yang, L., & Mekonnen, D. F. (2019). Hydrological evaluation of open-access precipitation and air temperature datasets using SWAT in a poorly gauged basin in Ethiopia. *Journal of Hydrology*, 569(November 2018), 612–626. <https://doi.org/10.1016/j.jhydrol.2018.12.026>
- Eini, M. R., Javadi, S., Delavar, M., Monteiro, J. A. F., & Darand, M. (2019). High accuracy of precipitation reanalyses resulted in good river discharge simulations in a semi-arid basin. *Ecological Engineering*, 131(March), 107–119. <https://doi.org/10.1016/j.ecoleng.2019.03.005>
- Farmer, W. H. (2016). Ordinary kriging as a tool to estimate historical daily streamflow records. *Hydrology and Earth System Sciences*, 20(7), 2721–2735. <https://doi.org/10.5194/hess-20-2721-2016>
- Farsani, I. F., Farzaneh, M. R., Besalatpour, A. A., Salehi, M. H., & Faramarzi, M. (2018). Assessment of the impact of climate change on spatiotemporal variability of blue and green water resources under CMIP3 and CMIP5 models in a highly mountainous watershed. *Theoretical and Applied Climatology*, 1–16. <https://doi.org/10.1007/s00704-018-2474-9>
- Flato, G. M., Boer, G. J., Lee, W. G., McFarlane, N. A., Ramsden, D., Reader, M. C., & Weaver, A. J. (2000). The Canadian centre for climate modelling

- and analysis global coupled model and its climate. *Climate Dynamics*, 16(6), 451–467. <https://doi.org/10.1007/s003820050339>
- Fuka, D. R., Walter, M. T., Macalister, C., Degaetano, A. T., Steenhuis, T. S., & Easton, Z. M. (2014). Using the Climate Forecast System Reanalysis as weather input data for watershed models. *Hydrological Processes*, 28(22), 5613–5623. <https://doi.org/10.1002/hyp.10073>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, 2(December). <https://doi.org/10.1038/sdata.2015.66>
- Gao, X., Zhu, Q., Yang, Z., & Wang, H. (2018). Evaluation and hydrological application of CMADS against TRMM 3B42V7, PERSIANN-CDR, NCEP-CFSR, and gauge-based datasets in Xiang River basin of China. *Water (Switzerland)*, 10(9), 1–24. <https://doi.org/10.3390/w10091225>
- Gassman, P. W., Reyes, M. R., Green, C. H., & Arnold, J. G. (2007). THE SOIL AND WATER ASSESSMENT TOOL: HISTORICAL DEVELOPMENT, APPLICATIONS, AND FUTURE RESEARCH DIRECTIONS. *American Society of Agricultural and Biological Engineers*, 50(4), 1211–1250. <https://doi.org/10.1109/RTSS.2012.58>
- Gassman, P. W., Reyes, M. R., Green, C. H., Arnold, J. G., P. W. Gassman, M. R. Reyes, C. H. Green, & J. G. Arnold. (2007). The Soil and Water Assessment Tool: Historical Development, Applications, and Future Research Directions. *Transactions of the ASABE*, 50(4), 1211–1250. <https://doi.org/10.13031/2013.23637>
- Gebremeskel, S., Liu, Y. B., De Smedt, F., Hoffmann, L., & Pfister, L. (2005). Analysing the effect of climate changes on streamflow using statistically downscaled GCM scenarios. *International Journal of River Basin Management*, 2(4), 271–280. <https://doi.org/10.1080/15715124.2004.9635237>
- Ghulami, M. (2018). *Assessment of climate change impacts on water resources and agriculture in data-scarce Kabul basin, Afghanistan*.
- Giorgi, F. (2019). Thirty Years of Regional Climate Modeling: Where Are We and Where Are We Going next? *Journal of Geophysical Research: Atmospheres*, 124(11), 5696–5723. <https://doi.org/10.1029/2018JD030094>
- Giorgi, F., & Bates, G. T. (1989). The Climatological Skill of a Regional Model over Complex Terrain. *American Meteorological Society*, 117(November), 2325–2347.
- Gitau, M W, Mehan, S., & Guo, T. (2017). Weather generator utilization in climate impact studies: Implications for water resources modelling. *European Water*, 59, 69–75. [http://www.ewra.net/ew/pdf/EW\\_2017\\_59\\_10.pdf](http://www.ewra.net/ew/pdf/EW_2017_59_10.pdf)
- Gitau, Margaret W., & Chaubey, I. (2010). Regionalization of SWAT model

- parameters for use in ungauged watersheds. *Water (Switzerland)*, 2(4), 849–871. <https://doi.org/10.3390/w2040849>
- Golmohammadi, G., Prasher, S., Madani, A., & Rudra, R. (2014). Evaluating Three Hydrological Distributed Watershed Models: MIKE-SHE, APEX, SWAT. *Hydrology*, 1(1), 20–39. <https://doi.org/10.3390/hydrology1010020>
- González-Zeas, D., Erazo, B., Lloret, P., De Bièvre, B., Steinschneider, S., & Dangles, O. (2019). Linking global climate change to local water availability: Limitations and prospects for a tropical mountain watershed. *Science of the Total Environment*, 650, 2577–2586. <https://doi.org/10.1016/j.scitotenv.2018.09.309>
- Gu, P., Wu, Y., Liu, G., Chengcheng, X., Wang, G., Xia, J., Chen, K., Huang, X., & Daiyuan, L. (2020). Application of Meteorological Element Combination-Driven SWAT Model based on Meteorological Datasets in Alpine catchment: A case study in the Yellow River Source Region. *Authorea*, September 21, 1–22.
- Guo, B., Zhang, J., Xu, T., Croke, B., Jakeman, A., Song, Y., Yang, Q., Lei, X., & Liao, W. (2018). Applicability assessment and uncertainty analysis of multi-precipitation datasets for the simulation of hydrologic models. *Water (Switzerland)*, 10(11). <https://doi.org/10.3390/w10111611>
- Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999). Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *Journal of Hydrologic Engineering*, 4(April), 135–143.
- Hasan, H. H., Fatin, S., Razali, M., Muhammad, N. S., & Ahmad, A. (2019). *Research Trends of Hydrological Drought*: 1–19.
- Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, 90(8), 1095–1107. <https://doi.org/10.1175/2009BAMS2607.1>
- Holden, J. (2011). Physical Geography. In *Physical Geography* (Vol. 34, Issue 1). <https://doi.org/10.1080/02723646.2013.798531>
- Hong, D., & Hong, K. A. (2016). Drought Identification, Monitoring and Forecasting for Selangor River Basin. *International Journal of U- and e-Service, Science and Technology*, 9(3), 53–66. <https://doi.org/10.14257/ijunesst.2016.9.3.07>
- Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., ... Cudennec, C. (2013). A decade of Predictions in Ungauged Basins (PUB)-a review. *Hydrological Sciences Journal*, 58(6), 1198–1255. <https://doi.org/10.1080/02626667.2013.803183>
- Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., & Wolff, D. B. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear,

- combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1), 38–55. <https://doi.org/10.1175/JHM560.1>
- Huntington, T. G. (2006). Evidence for intensification of the global water cycle: Review and synthesis. *Journal of Hydrology*, 319(1–4), 83–95. <https://doi.org/10.1016/j.jhydrol.2005.07.003>
- Inoue, T., & Matsumoto, J. (2004). A comparison of summer sea level pressure over East Eurasia between NCEP-NCAR reanalysis and ERA-40 for the period 1960-99. *Journal of the Meteorological Society of Japan*, 82(3), 951–958. <https://doi.org/10.2151/jmsj.2004.951>
- IPCC. (2007a). Climate Change 2007: Impacts, Adaptation and Vulnerability. In *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. <https://doi.org/10.1016/B978-008044910-4.00250-9>
- IPCC. (2007b). Climate change 2007. In *The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. <https://doi.org/10.1256/wea.58.04>
- IPCC. (2014a). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In R. K. Pachauri & L. A. Meyer (Eds.), *IPCC, Geneva, Switzerland*. <https://doi.org/10.1046/j.1365-2559.2002.1340a.x>
- IPCC. (2014b). Summary for Policymakers: Emissions Scenarios. In *International Panel on Climate Change* (Issue April 2007).
- IPCC. (2019a). Climate Change and Land. In *Research Handbook on Climate Change and Agricultural Law*. <https://www.ipcc.ch/srccl/download/>
- IPCC. (2019b). Global warming of 1.5°C. In *Special Report on Global Warming of 1.5°C*.
- IPCC. (2019c). The Ocean and Cryosphere in a Changing Climate. In *Intergovernmental Panel on Climate Change*. <https://www.ipcc.ch/srocc/chapter/summary-for-policymakers/>
- IPCC. (2022a). *Data Distribution Centre*. Intergovernmental Panel on Climate Change.
- IPCC. (2022b). Summary for Policymakers. In *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. (Issue 1). <https://doi.org/10.1017/9781009157926.001>.
- Ir. Mohd Zaki bin Mat Amin, Ali Ercan, Kei Ishida, M., L. K., Chen, Z. Q., & Su-Hyung Jang. (2019). Impacts of climate change on the hydro-climate of peninsular Malaysia. *Water (Switzerland)*, 11(9). <https://doi.org/10.3390/w11091798>
- Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive

- microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5(3), 487–503. [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2)
- K. C., A., C. A., J., M. Th., V. G., Abbaspour, K. C., Johnson, C. A., & van Genuchten, M. T. (2004). Estimating Uncertain Flow and Transport Parameters Using a Sequential Uncertainty Fitting Procedure. *Vadose Zone Journal*, 3(4), 1340–1352. <https://doi.org/10.2113/3.4.1340>
- Kavvas, M. L., Kure, S., Chen, Z. Q., Ohara, N., & Jang, S. (2013). WEHY-HCM for Modeling Interactive Atmospheric-Hydrologic Processes at Watershed Scale. I: Model Description. *Journal of Hydrologic Engineering*, 18(10), 1262–1271. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000724](https://doi.org/10.1061/(asce)he.1943-5584.0000724)
- Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrology and Earth System Sciences*, 23(10), 4323–4331. <https://doi.org/10.5194/hess-23-4323-2019>
- Kondo, T., Sakai, N., Yazawa, T., & Shimizu, Y. (2021). Verifying the applicability of SWAT to simulate fecal contamination for watershed management of Selangor River, Malaysia. *Science of the Total Environment*, 774, 145075. <https://doi.org/10.1016/j.scitotenv.2021.145075>
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3), 259–263. <https://doi.org/10.1127/0941-2948/2006/0130>
- Kouchi, D. H., Esmaili, K., Faridhosseini, A., Sanaeinejad, S. H., Khalili, D., & Abbaspour, K. C. (2017). Sensitivity of calibrated parameters and water resource estimates on different objective functions and optimization algorithms. *Water (Switzerland)*, 9(6), 1–16. <https://doi.org/10.3390/w9060384>
- Kundzewicz, Z. W., Krysanova, V., Benestad, R. E., Hov, Piniewski, M., & Otto, I. M. (2018). Uncertainty in climate change impacts on water resources. *Environmental Science and Policy*, 79(October 2017), 1–8. <https://doi.org/10.1016/j.envsci.2017.10.008>
- Li, C., Tang, G., & Hong, Y. (2018). Cross-evaluation of ground-based, multi-satellite and reanalysis precipitation products: Applicability of the Triple Collocation method across Mainland China. *Journal of Hydrology*, 562(September 2017), 71–83. <https://doi.org/10.1016/j.jhydrol.2018.04.039>
- Ligaray, M., Kim, H., Sthiannopkao, S., Lee, S., Cho, K. H., & Kim, J. H. (2015). Assessment on hydrologic response by climate change in the Chao Phraya River basin, Thailand. *Water (Switzerland)*, 7(12), 6892–6909. <https://doi.org/10.3390/w7126665>
- Liu, C., Zhang, Z., & Balay, J. W. (2018). Posterior assessment of reference

- gages for water resources management using instantaneous flow measurements. *Science of the Total Environment*, 634, 12–19. <https://doi.org/10.1016/j.scitotenv.2018.03.312>
- Liu, Z., Ostrenga, D., Teng, W., & Kempler, S. (2012). Tropical rainfall measuring mission (TRMM) precipitation data and services for research and applications. *Bulletin of the American Meteorological Society*, 93(9), 1317–1325. <https://doi.org/10.1175/BAMS-D-11-00152.1>
- Lo Conti, F., Hsu, K. L., Noto, L. V., & Sorooshian, S. (2014). Evaluation and comparison of satellite precipitation estimates with reference to a local area in the Mediterranean Sea. *Atmospheric Research*, 138, 189–204. <https://doi.org/10.1016/j.atmosres.2013.11.011>
- LUAS. (2015). *Sungai Selangor: State of the River Report 2015*.
- M., P., Varanou, E., Baltas, E., Dassaklis, A., & Mimikou, M. (2003). Application of the SWAT model in the Pinios river basin under different land-use scenarios. *Global NEST Journal*, 5(2), 71–79. <https://doi.org/10.30955/gnj.000277>
- M.Z.M. Amin, A.J. Shaaban, Ercan, A., Ishida, K., Kavvas, M. L., Chen, Z. Q., & Jang, S. (2017). Future climate change impact assessment of watershed scale hydrologic processes in Peninsular Malaysia by a regional climate model coupled with a physically-based hydrology model. *Science of the Total Environment*, 575(March), 12–22. <https://doi.org/10.1016/j.scitotenv.2016.10.009>
- Ma, L., Ascough, J. C., Ahuja, L. R., Shaffer, M. J., Hanson, J. D., & Rojas, K. W. (2000). Root Zone Water Quality Model sensitivity analysis using Monte Carlo simulation. *Transactions of the American Society of Agricultural Engineers*, 43(4), 883–895. <https://doi.org/10.13031/2013.2984>
- Mahmood, R. (2013). Assessment of climate change impact on water resources and hydropower in the Helum River Basin, Pakistan [Asian Institute of Technology, School of Engineering and Technology, Thailand]. In *Asian Institute of Technology, School of Engineering and Technology, Thailand*. <https://doi.org/10.1111/j.1600-0870.2010.00480.x>
- Mannschatz, T., Wolf, T., & Hülsmann, S. (2016). Nexus Tools Platform: Web-based comparison of modelling tools for analysis of water-soil-waste nexus. *Environmental Modelling and Software*, 76, 137–153. <https://doi.org/10.1016/j.envsoft.2015.10.031>
- Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock, T., Tignor, M., & Waterfield, T. (2018). IPCC Special Report 1.5 - Summary for Policymakers. In *ipcc*. <https://doi.org/10.1017/CBO9781107415324>
- Mathur, M. B. (1983). A Quasi-Lagrangian Regional Model Designed for Operational Weather Prediction. *Mon. Weather Rev.*, III, 2087–2099.

- Meng, Xian-yong, Wang, H., Cai, S., Zhang, X., Lei, X., Leng, G., Shi, C., & Liu, S. (2016). The China Meteorological Assimilation Driving Datasets for the SWAT Model ( CMADS ) Application in China : A Case Study in Heihe River Basin. *Preprints, February*. <https://doi.org/10.20944/preprints201612.0091.v1>
- Meng, Xianyong, & Wang, H. (2017). Significance of the China meteorological Assimilation Driving Datasets for the SWAT model (CMADS) of East Asia. *Water (Switzerland)*, 9(10). <https://doi.org/10.3390/w9100765>
- Meng, Xianyong, Wang, H., Lei, X., Cai, S., Wu, H., Ji, X., & Wang, J. (2017). Hydrological Modeling in the Manas River Basin using Soil and Water Assessment Tool driven by CMADS. *Tehnicki Vjesnik*, 24(2), 525–534. <https://doi.org/10.17559/TV-20170108133334>
- Meng, Xianyong, Wang, H., Shi, C., Wu, Y., & Ji, X. (2018). Establishment and evaluation of the China meteorological assimilation driving datasets for the SWAT model (CMADS). *Water (Switzerland)*, 10(11), 1–18. <https://doi.org/10.3390/w10111555>
- Merz, R., & Blöschl, G. (2004). Regionalisation of catchment model parameters. *Journal of Hydrology*, 287(1–4), 95–123. <https://doi.org/10.1016/j.jhydrol.2003.09.028>
- Miao, C., Ashouri, H., Hsu, K. L., Sorooshian, S., & Duan, Q. (2015). Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over China. *Journal of Hydrometeorology*, 16(3), 1387–1396. <https://doi.org/10.1175/JHM-D-14-0174.1>
- MNRE. (2012). *National Water Resources Policy* (Ministry of Natural Resources and Environment Malaysia (ed.); 2nd editio). Ministry of Natural Resources and Environment Malaysia. <https://doi.org/10.1038/nmat2778.A>
- Molina-Navarro, E., Andersen, H. E., Nielsen, A., Thodsen, H., & Trolle, D. (2017). The impact of the objective function in multi-site and multi-variable calibration of the SWAT model. *Environmental Modelling and Software*, 93, 255–267. <https://doi.org/10.1016/j.envsoft.2017.03.018>
- Monteiro, J. A. F., Strauch, M., Srinivasan, R., Abbaspour, K., & Gücker, B. (2016). Accuracy of grid precipitation data for Brazil: Application in river discharge modelling of the Tocantins catchment. *Hydrological Processes*, 30(9), 1419–1430. <https://doi.org/10.1002/hyp.10708>
- Mousavi, S. J., Kamali, B., Abbaspour, K. C., Amini, M., & Yang, H. (2012). Uncertainty-based automatic calibration of HEC-HMS model using sequential uncertainty fitting approach. *Journal of Hydroinformatics*, 14(2), 286–309. <https://doi.org/10.2166/hydro.2011.071>
- NAHRIM. (2016). *Vulnerability and Adaptation to Climate Change in Malaysia under the Third National Communication (TNC)* (Issue July).
- Nakicenovic, N., Alcamo, J., Davis, G., Vries, B. de, Fenhann, J., Gaffin, S., Gregory, K., Griibler, A., Jung, T. Y., Kram, T., Rovere, E. L. La,

- Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H., Price, L., Riahi, K., Roehrl, A., ... Dadi, Z. (2000). *Special Report on Emissions Scenarios*.
- Nash, E., & Sutcliffe, V. (1970). River flow forecasting through conceptual models Part I - A discussion of principles. *Journal of Hydrology*, 10, 282–290.
- Nasrabadi, E., Masoodian, S. A., & Asakereh, H. (2013). Comparison of Gridded Precipitation Time Series Data in APHRODITE and Asfazari Databases within Iran's Territory. *Atmospheric and Climate Sciences*, 03(02), 235–248. <https://doi.org/10.4236/acs.2013.32025>
- Neitsch, S. ., Arnold, J. ., Kiniry, J. ., & Williams, J. . (2011). Soil & Water Assessment Tool Theoretical Documentation Version 2009. *Texas Water Resources Institute*, 1–647. <https://doi.org/10.1016/j.scitotenv.2015.11.063>
- Nguyen, P., Ombadi, M., Sorooshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., Ashouri, H., & Rose Thorstensen, A. (2018). The PERSIANN family of global satellite precipitation data: A review and evaluation of products. *Hydrology and Earth System Sciences*, 22(11), 5801–5816. <https://doi.org/10.5194/hess-22-5801-2018>
- Nhi, P. T. T., Khoi, D. N., & Hoan, N. X. (2019). Evaluation of five gridded rainfall datasets in simulating streamflow in the upper Dong Nai river basin, Vietnam. *International Journal of Digital Earth*, 12(3), 311–327. <https://doi.org/10.1080/17538947.2018.1426647>
- Nijssen, B., & Lettenmaier, D. P. (2004). Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the Global Precipitation Measurement satellites. *Journal of Geophysical Research: Atmospheres*, 109(2), 1–15. <https://doi.org/10.1029/2003jd003497>
- Nkiaka, E., Nawaz, N. R., & Lovett, J. C. (2017). Analysis of rainfall variability in the Logone catchment, Lake Chad basin. *International Journal of Climatology*, 37(9), 3553–3564. <https://doi.org/10.1002/joc.4936>
- NWRS. (2011). Engineering Studies. In *Review of The National Water Resources Study (2000-2050) and Formulation of National Water Resources Policy* (Vol. 1, Issue Grid 2). Ministry of Natural Resources and Environment Malaysia. <https://doi.org/10.1038/nprot.2014.031>
- Oduşanya, A. E., Mehdi, B., Schürz, C., Oke, A. O., Awokola, O. S., Awomeso, J. A., Adejuwon, J. O., & Schulz, K. (2019). Multi-site calibration and validation of SWAT with satellite-based evapotranspiration in a data-sparse catchment in southwestern Nigeria. *Hydrology and Earth System Sciences*, 23(2), 1113–1144. <https://doi.org/10.5194/hess-23-1113-2019>
- Oki, T., & Kanae, S. (2006). Review Global Hydrological Cycles and World Water Resources. *Science*, 313(5790), 1068–1072. <https://doi.org/10.1126/science.1128845>
- Okiria, E., Okazawa, H., Noda, K., Kobayashi, Y., Suzuki, S., & Yamazaki, Y. (2022). *A Comparative Evaluation of Lumped and Semi-Distributed*

*Conceptual Hydrological Models: Does Model Complexity Enhance Hydrograph Prediction ?*

- Ozdemir, A., & Leloglu, U. M. (2019). A fast and automated hydrologic calibration tool for SWAT. *Water and Environment Journal*, 33(4), 488–498. <https://doi.org/10.1111/wej.12419>
- Pandey, V. P., Dhaubanjari, S., Bharati, L., & Thapa, B. R. (2019). Hydrological response of Chamelia watershed in Mahakali Basin to climate change. *Science of the Total Environment*, 650, 365–383. <https://doi.org/10.1016/j.scitotenv.2018.09.053>
- Prudhomme, C., Jakob, D., & Svensson, C. (2003). Uncertainty and climate change impact on the flood regime of small UK catchments. *Journal of Hydrology*, 277(1–2), 1–23. [https://doi.org/10.1016/S0022-1694\(03\)00065-9](https://doi.org/10.1016/S0022-1694(03)00065-9)
- Prudhomme, C., & Nick Reynard, S. C. (2002). Downscaling of global climate models for flood frequency analysis: Where are we now? *Hydrological Processes*, 16(6), 1137–1150. <https://doi.org/10.1002/hyp.1054>
- Raffar, N., Zulkafli, Z., Yiwon, M., Muharam, F. M., Rehan, B. M., & Nurulhuda, K. (2022). Watershed-scale modelling of the irrigated rice farming system at Muda, Malaysia, using the Soil Water Assessment Tool. *Hydrological Sciences Journal*, 67(3), 462–476. <https://doi.org/10.1080/02626667.2021.2022682>
- Razavi, T., & Coulibaly, P. (2013). Streamflow Prediction in Ungauged Basins: Review of Regionalization Methods. *Journal of Hydrologic Engineering*, 18(8), 958–975. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000690](https://doi.org/10.1061/(asce)he.1943-5584.0000690)
- Rivas-Tabares, D., Tarquis, A. M., Willaarts, B., & De Miguel, Á. (2019). An accurate evaluation of water availability in sub-arid Mediterranean watersheds through SWAT: Cega-Eresma-Adaja. *Agricultural Water Management*, 212(August 2018), 211–225. <https://doi.org/10.1016/j.agwat.2018.09.012>
- Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M. A., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., & Schulzweida, U. (2006). Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *Journal of Climate*, 19(16), 3771–3791. <https://doi.org/10.1175/JCLI3824.1>
- Rouholahnejad, E., Abbaspour, K. C., Vejdani, M., Srinivasan, R., Schulin, R., & Lehmann, A. (2012). A parallelization framework for calibration of hydrological models. *Environmental Modelling and Software*, 31, 28–36. <https://doi.org/10.1016/j.envsoft.2011.12.001>
- Rusli, N., Majid, M. R., Yusop, Z., Tan, M. L., Hashim, S., & Bohari, S. N. (2017). Integrating manual calibration and auto-calibration of SWAT model in Muar Watershed, Johor. *2016 7th IEEE Control and System Graduate Research Colloquium, ICSGRC 2016 - Proceeding, August*, 197–202. <https://doi.org/10.1109/ICSGRC.2016.7813327>
- Sa'adin, S. L., Kaewunruen, S., & Jaroszweski, D. (2016). Risks of Climate

- Change with Respect to the Singapore-Malaysia High Speed Rail System. *Climate*, 4(4), 65. <https://doi.org/10.3390/cli4040065>
- Saeidifarzad, B., Nourani, V., Aalami, M. T., & Chau, K. W. (2014). Multi-site calibration of linear reservoir based geomorphologic rainfall-runoff models. *Water (Switzerland)*, 6(9), 2690–2716. <https://doi.org/10.3390/w6092690>
- Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., Van den Dool, H. M., Pan, H. L., Moorthi, S., Behringer, D., Stokes, D., Peña, M., Lord, S., White, G., Ebisuzaki, W., Peng, P., & Xie, P. (2006). The NCEP Climate Forecast System. *Journal of Climate*, 19(15), 3483–3517. <https://doi.org/10.1175/JCLI3812.1>
- Santhi, C., Arnold, J. G., Williams, J. R., Dugas, W. A., Srinivasan, R., Hauck, L. M., The, V. O. F., & Point, B. W. (2002). VALIDATION OF THE SWAT MODEL ON A LARGE RWER tributing point and nonpoint sources so that the River Watershed in North Central Texas , where cy ( USEPA ) reported nutrient enrichment as the TMDL development through simulating loads to ates on a daily time. *October*, 37(5), 1169–1188.
- Sanusi, W., Jemain, A. A., Zin, W. Z. W., & Zahari, M. (2015). The Drought Characteristics Using the First-Order Homogeneous Markov Chain of Monthly Rainfall Data in Peninsular Malaysia. *Water Resources Management*, 29(5), 1523–1539. <https://doi.org/10.1007/s11269-014-0892-8>
- Searcy, J. K., & Hardison, C. H. (1960). Double-Mass Curves. In *Water Supply Paper 1541B*. <https://doi.org/http://udspace.udel.edu/handle/19716/1592>
- Seyyedi, H., Anagnostou, E. N., Beighley, E., & McCollum, J. (2015). Hydrologic evaluation of satellite and reanalysis precipitation datasets over a mid-latitude basin. *Atmospheric Research*, 164–165, 37–48. <https://doi.org/10.1016/j.atmosres.2015.03.019>
- Shaaban, A. J., M. Amin, M. Z., Q. Chen, Z., ASCE, M., Ohara, N., & ASCE, A. M. (2011). Regional Modeling of Climate Change Impact on Peninsular Malaysia Water Resources. *Journal of Hydrologic Engineering*, 16(12), 1040–1049. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000305](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000305)
- Shao, G., Guan, Y., Zhang, D., Yu, B., & Zhu, J. (2018). The impacts of climate variability and land use change on streamflow in the Hailiutu River Basin. *Water (Switzerland)*, 10(6). <https://doi.org/10.3390/w10060814>
- Shaw, R., Luo, Y., Cheong, T. S., Halim, S. A., Chaturvedi, S., Hashizume, M., Insarov, G. E., Ishikawa, Y., Jafari, M., Kitoh, A., Pulhin, J., Singh, C., Vasant, K., & Zhang, Z. (2022). *Asia. In: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. <https://doi.org/10.1017/9781009325844.012.1457>
- Shrestha, M. K., Recknagel, F., Frizenschaf, J., & Meyer, W. (2016). Assessing SWAT models based on single and multi-site calibration for the simulation of flow and nutrient loads in the semi-arid Onkaparinga

catchment in South Australia. *Agricultural Water Management*, 175, 61–71. <https://doi.org/10.1016/j.agwat.2016.02.009>

- Sitterson, J., Knightes, C., Parmar, R., Wolfe, K., Muche, M., & Avant, Brian. (2018). An Overview of Rainfall-Runoff Model Types. *International Congress on Environmental Modelling and Software, September*, 30. <https://scholarsarchive.byu.edu/iemssconference><https://scholarsarchive.byu.edu/iemssconference/2018/Stream-C/41Thisoralpresentation>
- Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., & Zehe, E. (2003). IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal*, 48(6), 857–880. <https://doi.org/10.1623/hysj.48.6.857.51421>
- Song, Y., Zhang, J., & Lai, Y. (2021). Influence of multisite calibration on streamflow estimation based on the hydrological model with CMADS inputs. *Journal of Water and Climate Change*, 00(0), 1–18. <https://doi.org/10.2166/wcc.2021.115>
- Stahl, K., Hisdal, H., Tallaksen, L., Lanen, H. A. J., Hannaford, J., & Sauquet, E. (2008). *Trends in low flows and streamflow droughts across Europe. March*. <http://library.wur.nl/WebQuery/wurpubs/374150%0Ahttp://en.scientificcommons.org/39629850>
- Suhaila, J., & Jemain, A. A. (2009). Investigating the impacts of adjoining wet days on the distribution of daily rainfall amounts in Peninsular Malaysia. *Journal of Hydrology*, 368(1–4), 17–25. <https://doi.org/10.1016/j.jhydrol.2009.01.022>
- Swain, J. B., & Patra, K. C. (2017). Streamflow estimation in ungauged catchments using regionalization techniques. *Journal of Hydrology*, 554, 420–433. <https://doi.org/10.1016/j.jhydrol.2017.08.054>
- T. Tangang, F., Juneng, L., Salimun, E., Meng Sei, K., Jui Le, L., Muhamad, H., Tangang, F. T., Juneng, L., Salimun, E., Sei, K. M., Le, L. J., & Muhamad, H. (2012). Climate change and variability over Malaysia: Gaps in science and research information. *Sains Malaysiana*, 41(11), 1355–1366.
- Tan, M. L. (2014). Free internet datasets for streamflow modelling using SWAT in the Johor river basin, Malaysia. *IOP Conference Series: Earth and Environmental Science*, 18(1). <https://doi.org/10.1088/1755-1315/18/1/012193>
- Tan, Mou Leong, Gassman, P. W., & Cracknell, A. P. (2017). Assessment of three long-term gridded climate products for hydro-climatic simulations in tropical river basins. *Water (Switzerland)*, 9(3). <https://doi.org/10.3390/w9030229>
- Tan, Mou Leong, Gassman, P. W., Liang, J., & Haywood, J. M. (2021). A

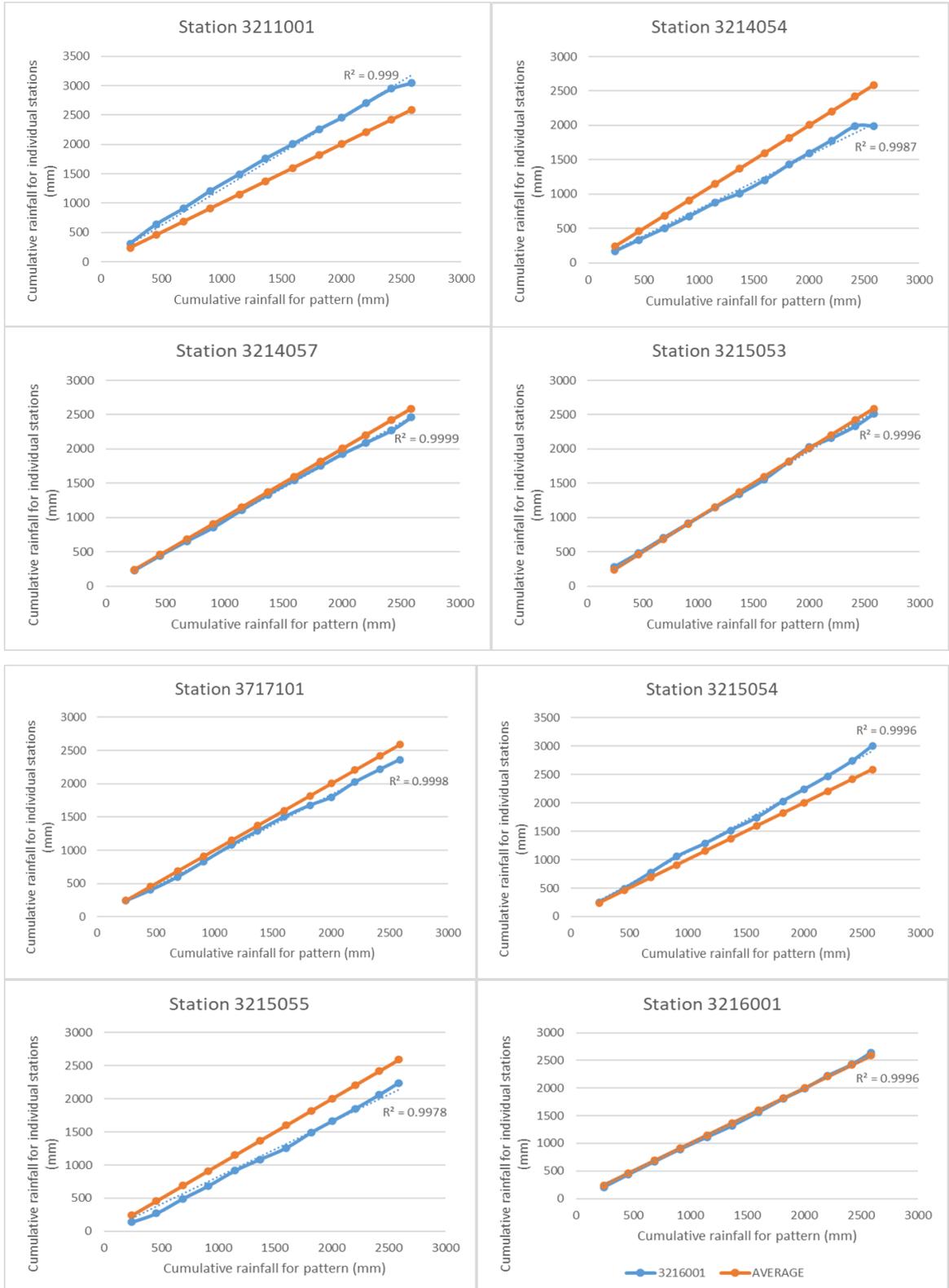
- review of alternative climate products for SWAT modelling: Sources, assessment and future directions. *Science of the Total Environment*, 795, 148915. <https://doi.org/10.1016/j.scitotenv.2021.148915>
- Tan, Mou Leong, Gassman, P. W., Srinivasan, R., Arnold, J. G., & Yang, X. Y. (2019). A review of SWAT studies in Southeast Asia: Applications, challenges and future directions. *Water (Switzerland)*, 11(5), 1–25. <https://doi.org/10.3390/w11050914>
- Tan, Mou Leong, Ibrahim, A. L., Yusop, Z., Chua, V. P., & Chan, N. W. (2017). Climate change impacts under CMIP5 RCP scenarios on water resources of the Kelantan River Basin, Malaysia. *Atmospheric Research*, 189, 1–10. <https://doi.org/10.1016/j.atmosres.2017.01.008>
- Tan, Mou Leong, Samat, N., Chan, N. W., & Roy, R. (2018). Hydro-meteorological assessment of three GPM Satellite Precipitation Products in the Kelantan River Basin, Malaysia. *Remote Sensing*, 10(7). <https://doi.org/10.3390/rs10071011>
- Tian, Y., Zhang, K., Xu, Y. P., Gao, X., & Wang, J. (2018). Evaluation of potential evapotranspiration based on CMADS reanalysis dataset over China. *Water (Switzerland)*, 10(9), 1–17. <https://doi.org/10.3390/w10091126>
- Tsanis, I. K., Koutroulis, A. G., Daliakopoulos, I. N., & Jacob, D. (2011). Severe climate-induced water shortage and extremes in Crete. *Climatic Change*, 106(4), 667–677. <https://doi.org/10.1007/s10584-011-0048-2>
- Vasiliades, L., Sarailidis, G., & Loukas, A. (2017). Hydrological Modelling of Low Flows for Operational Water Resources Management. *Proceedings of the 10th WORLD CONGRESS on Water Resources and Environment Panta Rhei, d*, 223–229.
- Villarini, G., Krajewski, W. F., & Smith, J. A. (2009). New paradigm for statistical validation of satellite precipitation estimates: Application to a large sample of the TMPA 0.25° 3-hourly estimates over Oklahoma. *Journal of Geophysical Research Atmospheres*, 114(12), 1–12. <https://doi.org/10.1029/2008JD011475>
- Voon, K. L., Tan, K. W., & Chin, K. S. (2022). Assessment of the Flood and Drought Occurrence Using Statistically Downscaled Local Climate Models: a Case Study in Langat River Basin, Malaysia. *Civil and Environmental Engineering*, 18(1), 221–233. <https://doi.org/10.2478/cee-2022-0021>
- Vu, T. T., Li, L., & Jun, K. S. (2018). Evaluation of multi-satellite precipitation products for streamflow simulations: A case study for the Han River Basin in the Korean Peninsula, East Asia. *Water (Switzerland)*, 10(5). <https://doi.org/10.3390/w10050642>
- Wang, Q., Xia, J., Zhang, X., She, D., Liu, J., & Li, P. (2020). *Multi-Scenario Integration Comparison of CMADS and TMPA Datasets for Hydro-Climatic Simulation over Ganjiang River Basin, China*. 1–22.
- Wang, S., Zhang, Z., Sun, G., Strauss, P., Guo, J., Tang, Y., & Yao, A. (2012).

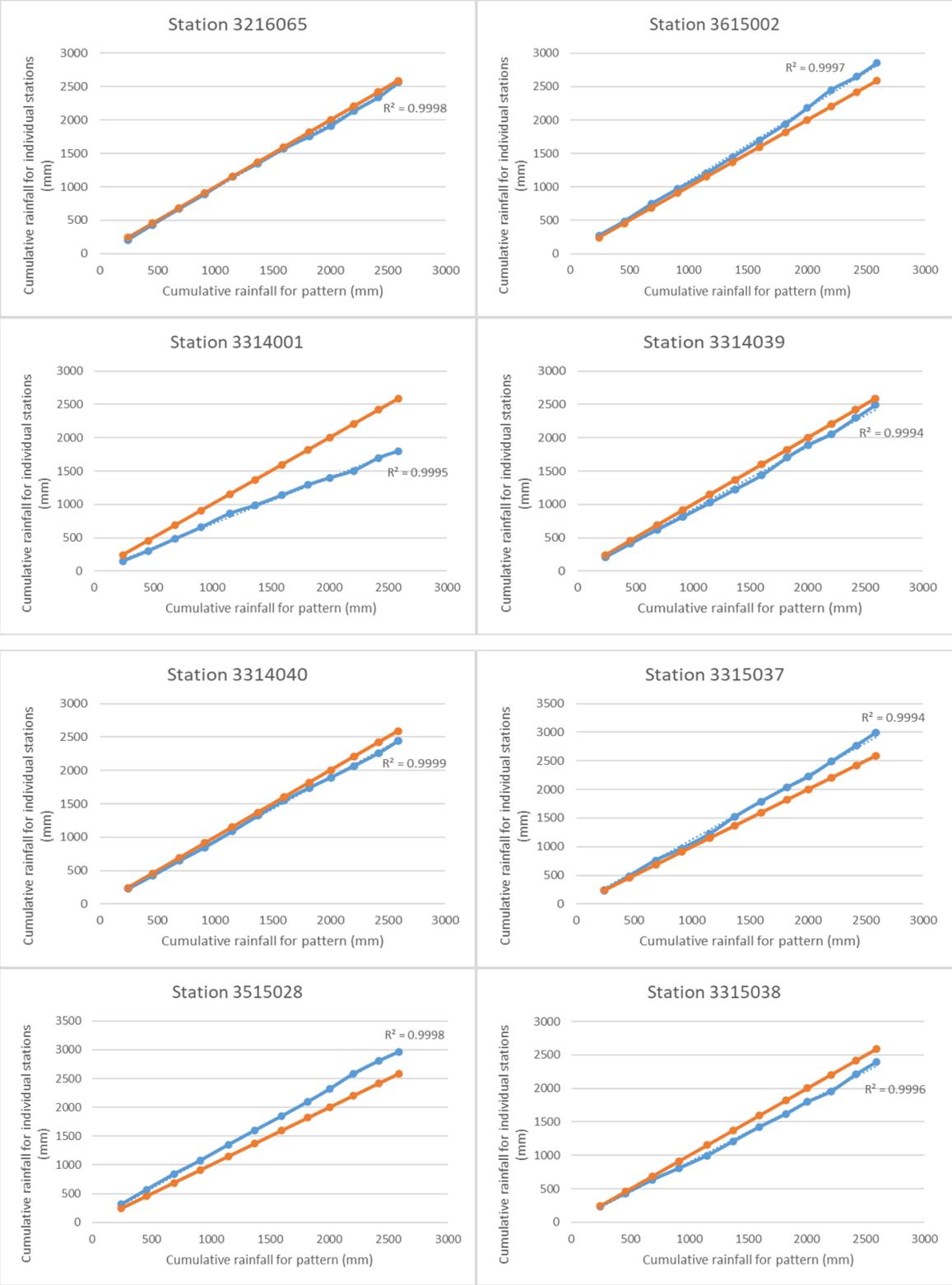
- Multi-site calibration, validation, and sensitivity analysis of the MIKE SHE Model for a large watershed in northern China. *Hydrology and Earth System Sciences*, 16(12), 4621–4632. <https://doi.org/10.5194/hess-16-4621-2012>
- Wang, W., Xie, P., Yoo, S. H., Xue, Y., Kumar, A., & Wu, X. (2011). An assessment of the surface climate in the NCEP climate forecast system reanalysis. *Climate Dynamics*, 37(7–8), 1601–1620. <https://doi.org/10.1007/s00382-010-0935-7>
- Wilby, R L, Charles, S. P., Zorita, E., Timbal, B., Whetton, P., & Mearns, L. O. (2004). *Guidelines for use of climate scenarios developed from statistical downscaling methods. Supporting material of the Intergovernmental Panel on Climate Change, prepared on behalf of Task Group on Data and Scenario Support for Impacts and Climate Analysis. August*, 1–27.
- Wilby, Robert L., Hassan, H., & Hanaki, K. (1998). Statistical downscaling of hydrometeorological variables using general circulation model output. *Journal of Hydrology*, 205(1–2), 1–19. [https://doi.org/10.1016/S0022-1694\(97\)00130-3](https://doi.org/10.1016/S0022-1694(97)00130-3)
- Worqlul, A. W., Maathuis, B., Adem, A. A., Demissie, S. S., Langan, S., & Steenhuis, T. S. (2014). Comparison of rainfall estimations by TRMM 3B42, MPEG and CFSR with ground-observed data for the Lake Tana basin in Ethiopia. *Hydrology and Earth System Sciences*, 18(12), 4871–4881. <https://doi.org/10.5194/hess-18-4871-2014>
- Xian-Yong Meng, Dan-Lin, Y., & Zhi-Hui, L. (2015). Energy Balance-Based SWAT Model to Simulate the Mountain Snowmelt and Runoff – Taking the Application in Juntanghu Watershed (China) as an Example. *Journal of Mountain Science*, 12(2). <https://doi.org/10.1007/s11629-014-3081-6>
- Yang, J., Reichert, P., Abbaspour, K. C., Xia, J., & Yang, H. (2008). Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *Journal of Hydrology*, 358(1–2), 1–23. <https://doi.org/10.1016/j.jhydrol.2008.05.012>
- Yukimoto, S., Noda, A., Kitoh, A., Sugi, M., Kitamura, Y., Hosaka, M., Shibata, K., Maeda, S., & Uchiyama, T. (2001). The New Meteorological Research Institute Coupled GCM (MRI-CGCM2) - Model Climate and Variability. In *Meteorology and Geophysics* (Vol. 51, Issue 2, pp. 47–88).
- Zhang, D., Tan, M. L., Dawood, S. R. S., Samat, N., Chang, C. K., Roy, R., Tew, Y. L., & Mahamud, M. A. (2020). Comparison of ncep-cfsr and cmads for hydrological modelling using swat in the muda river basin, malaysia. *Water (Switzerland)*, 12(11). <https://doi.org/10.3390/w12113288>
- Zhang, L., Meng, X., Wang, H., Yang, M., & Cai, S. (2020). Investigate the applicability of CMADS and CFSR reanalysis in Northeast China. *Water (Switzerland)*, 12(4). <https://doi.org/10.3390/W12040996>
- Zhang, Q., Körnich, H., & Holmgren, K. (2013). How well do reanalyses represent the southern African precipitation? *Climate Dynamics*, 40(3–4),

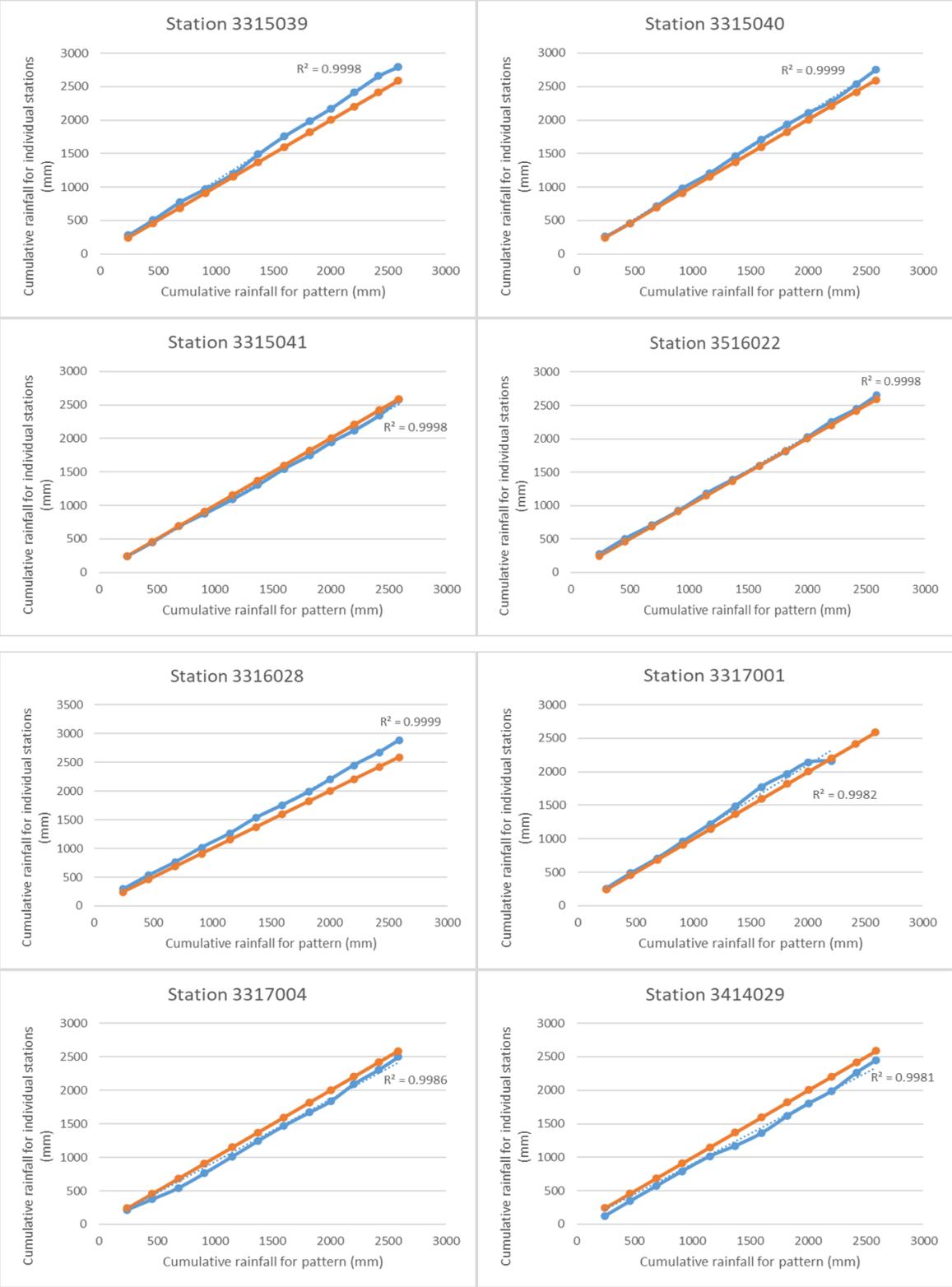
951–962. <https://doi.org/10.1007/s00382-012-1423-z>

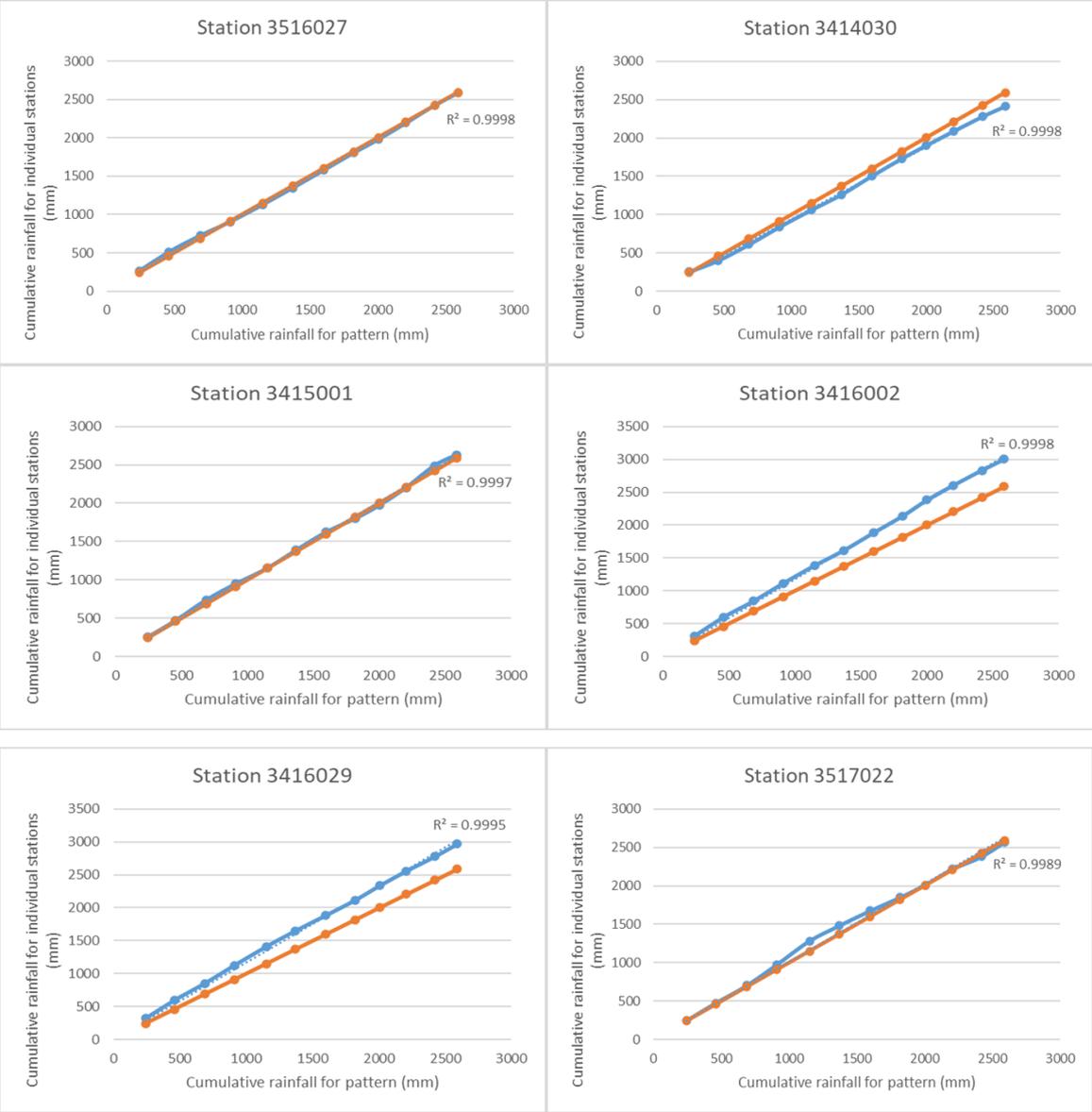
## APPENDIX A – DOUBLE MASS CURVE ANALYSIS

- Double mass curve analysis result on 30 hydrological stations









- Table of Double Mass Curve analysis on 30 hydrological stations

| Station / Year | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | r     | r <sup>2</sup> |
|----------------|------|------|------|------|------|------|------|------|------|------|------|------|-------|----------------|
| 3211001        | 310  | 636  | 909  | 1207 | 1492 | 1760 | 2006 | 2262 | 2455 | 2702 | 2947 | 3044 | 0.999 | 0.998          |
| 3214054        | 170  | 332  | 499  | 674  | 872  | 1011 | 1202 | 1428 | 1598 | 1780 | 1985 | 1985 | 0.998 | 0.996          |
| 3214057        | 231  | 442  | 656  | 854  | 1109 | 1332 | 1548 | 1753 | 1926 | 2091 | 2268 | 2461 | 1.000 | 0.999          |
| 3215053        | 282  | 479  | 705  | 915  | 1144 | 1341 | 1554 | 1817 | 2025 | 2158 | 2330 | 2512 | 0.999 | 0.999          |
| 3215054        | 254  | 492  | 776  | 1057 | 1284 | 1516 | 1742 | 2030 | 2238 | 2470 | 2743 | 3004 | 0.999 | 0.999          |
| 3215055        | 139  | 270  | 489  | 681  | 915  | 1086 | 1256 | 1490 | 1661 | 1846 | 2060 | 2237 | 0.999 | 0.999          |
| 3216001        | 206  | 433  | 668  | 890  | 1113 | 1320 | 1560 | 1805 | 1987 | 2219 | 2428 | 2644 | 1.000 | 0.999          |
| 3216065        | 206  | 433  | 668  | 890  | 1150 | 1349 | 1571 | 1752 | 1913 | 2132 | 2340 | 2559 | 1.000 | 0.999          |
| 3314001        | 152  | 304  | 484  | 658  | 862  | 987  | 1139 | 1293 | 1399 | 1506 | 1697 | 1800 | 0.999 | 0.998          |
| 3314039        | 213  | 411  | 620  | 813  | 1026 | 1228 | 1437 | 1704 | 1893 | 2057 | 2298 | 2490 | 0.999 | 0.998          |
| 3314040        | 228  | 427  | 646  | 849  | 1090 | 1321 | 1548 | 1741 | 1894 | 2065 | 2258 | 2438 | 1.000 | 0.999          |
| 3315037        | 232  | 482  | 761  | 964  | 1220 | 1523 | 1790 | 2034 | 2224 | 2489 | 2770 | 2997 | 0.999 | 0.999          |
| 3315038        | 238  | 431  | 633  | 805  | 995  | 1206 | 1424 | 1617 | 1797 | 1955 | 2211 | 2391 | 0.999 | 0.998          |
| 3315039        | 280  | 509  | 772  | 971  | 1196 | 1490 | 1760 | 1986 | 2175 | 2412 | 2661 | 2793 | 1.000 | 0.999          |
| 3315040        | 261  | 460  | 717  | 981  | 1210 | 1464 | 1711 | 1931 | 2108 | 2275 | 2537 | 2749 | 1.000 | 0.999          |
| 3315041        | 246  | 448  | 689  | 875  | 1090 | 1303 | 1546 | 1746 | 1941 | 2122 | 2342 | 2580 | 0.999 | 0.999          |
| 3316028        | 292  | 533  | 765  | 1018 | 1265 | 1534 | 1757 | 1986 | 2203 | 2449 | 2675 | 2885 | 1.000 | 1.000          |
| 3317001        | 259  | 490  | 708  | 960  | 1222 | 1486 | 1775 | 1969 | 2139 | 2162 | 2162 | 2162 | 0.976 | 0.952          |
| 3317004        | 218  | 374  | 545  | 763  | 1012 | 1243 | 1472 | 1672 | 1839 | 2096 | 2309 | 2501 | 0.999 | 0.998          |
| 3414029        | 123  | 343  | 570  | 793  | 1015 | 1171 | 1361 | 1616 | 1803 | 1990 | 2265 | 2451 | 0.999 | 0.998          |
| 3414030        | 256  | 404  | 613  | 838  | 1063 | 1260 | 1503 | 1730 | 1904 | 2086 | 2277 | 2411 | 1.000 | 0.999          |
| 3415001        | 254  | 475  | 739  | 948  | 1151 | 1390 | 1622 | 1801 | 1975 | 2203 | 2489 | 2629 | 0.999 | 0.999          |
| 3416002        | 317  | 597  | 844  | 1113 | 1384 | 1614 | 1883 | 2140 | 2382 | 2607 | 2834 | 3002 | 1.000 | 1.000          |
| 3416029        | 318  | 594  | 852  | 1125 | 1410 | 1653 | 1884 | 2117 | 2336 | 2558 | 2779 | 2970 | 1.000 | 0.999          |
| 3515028        | 311  | 571  | 839  | 1078 | 1354 | 1601 | 1851 | 2102 | 2324 | 2587 | 2811 | 2969 | 1.000 | 1.000          |
| 3516022        | 274  | 506  | 710  | 922  | 1187 | 1390 | 1598 | 1812 | 2022 | 2256 | 2449 | 2650 | 1.000 | 0.999          |
| 3516027        | 266  | 510  | 724  | 906  | 1130 | 1348 | 1577 | 1805 | 1983 | 2194 | 2429 | 2588 | 1.000 | 0.999          |
| 3517022        | 246  | 475  | 703  | 977  | 1284 | 1483 | 1676 | 1851 | 2011 | 2217 | 2380 | 2567 | 0.998 | 0.996          |
| 3615002        | 273  | 487  | 746  | 971  | 1204 | 1445 | 1691 | 1943 | 2179 | 2452 | 2650 | 2853 | 0.999 | 0.999          |
| 3717101        | 244  | 406  | 601  | 831  | 1082 | 1291 | 1502 | 1677 | 1799 | 2023 | 2215 | 2362 | 0.999 | 0.999          |
| Average        | 243  | 458  | 688  | 911  | 1151 | 1372 | 1598 | 1820 | 2004 | 2205 | 2420 | 2589 | 0.999 | 0.997          |

