IDENTIFICATION OF UNCERTAINTY IN HYDROLOGICAL MODELLING USING SEVERAL MODEL STRUCTURES, OPTIMISATION ALGORITHMS AND OBJECTIVE FUNCTIONS

MSc in Management of Hydrometeorological Hazards

A thesis submitted by

Eleftherios Gkilimanakis

to the University of Thessaly and the Joseph Fourier University

supervised by

Prof. Loukas Athanasios

Dr. Vasiliades Lampros

Laboratory of Hydrology and Aquatic Systems Analysis, Department of Civil Engineering, School of Engineering, University of Thessaly, Pedion Areos, 38333, Volos, Greece.

January 2013

Acknowledgements

I would like to thank my supervisors Prof. Loukas Athanasios and Dr. Vasiliades Lampros for their timely and constructive guidance and encouragement throughout the duration of this study. Also, acknowledgement is made to Papaioannou George who has helped me so freely and supplied me with courage to complete this project.

I wish to thank in general all the teaching staff at UJF and at the University of Thessaly that participated in the MSc of Hydrohazards for the knowledge that I acquired and for having more confidence now to stand in the “world” of hydrometeorology.

ABSTRACT

The aim of this study is to assess and identify different types of hydrological uncertainty in daily runoff simulation at Yermasoyia watershed, Cyprus. This is achieved with the aid of a free online software called 'HYDROMAD' which is based on the unit hydrograph theory of rainfall-runoff component separation. This implies a two-component structure: a soil moisture accounting (SMA) module and a routing or unit hydrograph module. The SMA module converts rainfall and temperature into effective rainfall: The routing module converts effective rainfall into streamflow,

There is a number of different strategies that can be used to calibrate a model. The typical approach is a joint optimization of all parameters or, alternatively, the unit hydrograph could be estimated directly from streamflow data, using inverse filtering or average event unit hydrograph estimation. Several well-known lumped hydrological models (such as the GR4J, the IHACRES models and the AWBM) are applied and tested for accurate runoff modeling using the split-sample test for estimating model error uncertainty. Several local non-linear and global optimization algorithms (i.e. Shuffled Complex Evolution, DiffeRential Evolution Adaptive Metropolis, DiffeRential Evolution, Simulated Annealing, quasi-Newton) have been deployed and compared for calibrating the different hydrological model structures to observed streamflow data for identification of optimization error uncertainty and model parameter uncertainty.

Finally, several objective functions (i.e. Nash-Sutcliffe Efficiency and variations or adaptations) which address different parts of the hydrograph have been used to identify parameter and model uncertainty. The model performance of the above tests was evaluated with the use of fit statistics or metrics for calibration and validation periods. Application of the tests in Yermasoyia watershed showed that the primary source of uncertainty in rainfall-runoff modeling was the choice of the hydrological model (model structure) followed by the parameter uncertainty caused by the optimization algorithm and the choice of objective function.

KEYWORDS: Hydrological models, rainfall-runoff modeling, model structure uncertainty, optimization algorithms, objective functions, parameter uncertainty, streamflow modeling.

ΠΕΡΙΛΗΨΗ

Σκοπός αυτής της εργασίας είναι ο προσδιορισμός και η αξιολόγηση των διαφόρων πηγών υδρολογικής αβεβαιότητας, στην ημερήσια προσομοίωση της απορροής, στη λεκάνη απορροής της Yermasoyia στην Κύπρο. Αυτό επιτυγχάνεται με τη βοήθεια ενός δωρεάν διαδικτυακού λογισμικού που ονομάζεται «HYDROMAD» και το οποίο βασίζεται στη θεωρία του μοναδιαίου υδρογραφήματος για το διαχωρισμό του υδρολογικού ισοζυγίου και της διόδευσης της ροής. Με βάση το διαχωρισμό αυτό, υπονοείται ότι και το υπολογιστικό πλαίσιο (model framework) θα αποτελείται από δύο επιμέρους τμήματα στη δομή του: τη μονάδα λογιστικής εδαφολογικής υγρασίας (SMA) και τη μονάδα διόδευσης ροής (routing) ή τη μονάδα του μοναδιαίου υδρογραφήματος. Στο πρώτο τμήμα γίνεται η μετατροπή της βροχόπτωσης και της θερμοκρασίας σε ενεργή βροχόπτωση, ενώ στο δεύτερο τμήμα γίνεται η μετατροπή της ενεργής βροχόπτωσης σε επιφανειακή απορροή.

Γενικά, υπάρχει ένας μεγάλος αριθμός διαφορετικών μεθοδολογιών που μπορούν να χρησιμοποιηθούν στη βαθμονόμηση ενός μοντέλου. Η τυπική προσέγγιση είναι η από κοινού βελτιστοποίηση όλων των παραμέτρων του μοντέλου ή, εναλλακτικώς, η εκτίμηση του μοναδιαίου υδρογραφήματος μπορεί να γίνει απευθείας από τα δεδομένα της απορροής, χρησιμοποιώντας αντίστροφο φιλτράρισμα (inverse filtering) ή του μέσου όρου των γεγονότων. Διάφορα ημερήσια αδρομερή υδρολογικά μοντέλα βροχόπτωσης – απορροής (όπως το GR4J, το AWBM και παραλλαγές των μοντέλων IHACRES) δοκιμάστηκαν και μελετήθηκαν, ως προς την ακρίβεια της προσομοιωμένης απορροής, χρησιμοποιώντας τη μέθοδο χωριστού δείγματος στην εκτίμηση της υδρολογικής αβεβαιότητας. Διάφοροι μη-γραμμικοί αλγόριθμοι βελτιστοποίησης τοπικής και καθολικής προσέγγισης (όπως για παράδειγμα οι: Shuffled Complex Evolution, DiffeRential Evolution Adaptive Metropolis, DiffeRential Evolution, Simulated Annealing και quasi-Newton) αναπτύχθηκαν και συγκρίθηκαν κατά τη βαθμονόμηση των διαφορετικών δομών των υδρολογικών μοντέλων ως προς τα παρατηρημένα δεδομένα της απορροής. Μέσα από τη σύγκριση αυτή εμφανίζεται η αβεβαιότητα στο σφάλμα της βελτιστοποίησης, στις τιμές των παραμέτρων των μοντέλων αλλά και στην ίδια τη δομή τους.

Τέλος, διάφορες αντικειμενικές συναρτήσεις (όπως η Nash-Sutcliffe Efficiency και παραλλαγές ή προσαρμογές αυτής), οι οποίες απευθύνονται σε διαφορετικά τμήματα του υδρογραφήματος, χρησιμοποιήθηκαν για τον εντοπισμό της αβεβαιότητας των παραμέτρων και της δομής των μοντέλων. Η απόδοση των μοντέλων αξιολογήθηκε με τη χρήση στατιστικών καλής προσαρμογής για τις περιόδους βαθμονόμησης και επαλήθευσης. Η εφαρμογή δοκιμών στη λεκάνη απορροής της Yermasoyia δείχνει πως η βασική πηγή αβεβαιότητας στη μοντελοποίηση βροχόπτωσης – απορροής είναι η επιλογή της δομής του

υδρολογικού μοντέλου, ακολουθούμενη από την αβεβαιότητα των παραμέτρων που προκαλείται από την επιλογή των αλγορίθμων βελτιστοποίησης και των αντικειμενικών συναρτήσεων.

Λέξεις-κλειδιά: Υδρολογικά μοντέλα, μοντέλα βροχόπτωσης – απορροής , αβεβαιότητα δομής μοντέλων, αλγόριθμοι βελτιστοποίησης, αντικειμενικές συναρτήσεις, αβεβαιότητα παραμέτρων, μοντελοποίηση απορροής.

RESUME

Le but de cette étude est d'identifier et d'estimer les différentes sources d'incertitude hydrologique associées aux simulations de ruissellement journalier dans le bassin versant de Yermasoyia à Chypre. Pour ce faire, on utilise le programme 'HYDROMAD', disponible gratuitement en ligne, basé sur la théorie de l'hydrographe unitaire de séparation des composantes de ruissèlement et d'écoulement. Ceci implique une structure de modèle constituée de deux éléments: un module d'évolution de l'humidité du sol (EHS) ainsi qu'un module de routage hydrologique ou un hydrographe unitaire. Le module EHS calcule la pluie nette à partir des données de pluie totale et de température. Le module de routage convertie ensuite cette pluie nette en débit. Il y a de nombreuses méthodes de calibration différentes pour un tel modèle.

L'approche usuelle consiste soit à optimiser simultanément l'ensemble des paramètres soit à estimer directement l'hydrographe unitaire à partir des données de débit en utilisant un filtrage inverse ou une estimation de l'hydrographe unitaire associé a un évènement moyen. Un grand nombre de modèles hydrologiques empiriques (tels les modèles GR4J, IHQCRES et AWBM) sont appliqués et leurs simulations du ruissèlement sont évaluées par le biais du test de validation croisée qui permet d'estimer l'incertitude associée à l'erreur modèle. De nombreux algorithmes d'optimisation non linéaires locaux ou globaux (c.à.d. Shuffled Complex Evolution, DiffeRential Evolution Adaptive Metropolis, DiffeRential Evolution, Simulated Annealing, quasi-Newton) sont déployés et leurs calibrations des différentes structures du modèle hydrologique par rapports aux données de débit observées sont comparées afin d'identifier et de réduire les incertitudes liées à l'erreur et aux paramètres du modèle.

Enfin une série de fonctions objectives (par exemple : Nash-Sutcliffe Efficiency et variantes ou adaptations) focalisées sur différentes caractéristiques de l'hydrographe sont utilisées pour identifier ces sources d'incertitudes. Les performances du modèle par rapport aux tests précédents sont évaluées grâce à des statistiques ou mesures de la qualité des données simulées pour les périodes de calibration et de validation. L'application de cette méthodologie sur le bassin versant de Yermasoyia a prouvé que la première source d'incertitude des simulations pluie-débit est le choix du modèle hydrologique (structure du modèle) suivi par l'incertitude sur les paramètres engendrée par le choix de l'algorithme d'optimisation et de la fonction objective.

MOTS CLES: modèles hydrologiques, simulation pluie-débit, incertitude liée a la structure du modèle, algorithme d'optimisation, fonction objective, incertitude liée aux paramètres, simulation de débit

Contents

CHAPTER 1

1.1 Background and project overview:

Conceptual lumped Rainfall-Runoff models, (R-R models), have been coupled with computing science since the early 1970s and have earned a central spot in hydrological modeling. Among many other models that are referred in the literature, some of the most representative "early" ones are undoubtedly the HBV *(Bergstrom & Forsman, 1973)* and the Sacramento *(Burnash, 1973)* models. Most of such models have surely undergone modifications or refinements over the years, through continuous improvement processes by their developers or others. However, the core physical and functional basis of the models has been kept intact and is still in use nowadays as such.

R-R models usually have a relatively simple structure and use not too complex mathematical equations. They all aim to interpret the physical processes that take place within a catchment, through which rainfall is transformed into runoff. As far as the lumped R-R models are concerned, their parameters cannot be directly measured on site, although they represent physical processes and characteristics of the catchments. Therefore, their values, or ranges of values have to be determined through calibration techniques*, (Wheater et al., 1993).*

Initially, calibration was performed manually by trial and error approaches but this soon proved to be a rather laborious and time consuming work *(Madsen et al., 2001)*. Inevitably scientific research on this area focused on the development of different calibration approaches which were more automated. Such approaches generally involve the selection of either a single "objective function" as a measure of goodness of fit or a multi-objective calibration approach, as proposed by *Gupta et al., (1998)*. Furthermore, the Pareto optimal solution for parameter sets as proposed by *Gharari et al., 2012* give an indication of where research in this field focuses more. The use of a single objective calibration procedure aims to develop models that either focus on a particular characteristic of the hydrograph (peak / low flows) or in case of multi-objective calibration, models are developed to represent an overall behavior for all the parts of the hydrograph. Other research works have been conducted considering multi-criteria in the calibration process, for example tracers' concentrations or remotely sensed evaporation, *(Weiler et al., 2003) and (Winsemius et al., 2008).*

Automated calibration procedures also include the selection of a parameter search procedure, called "optimization algorithm". Until recently, calibration has been performed by local-search optimization algorithms *(Yapo et al., 1996),* i.e. using the *Nelder and Mead (1965)* simplex algorithm and others. In *1992, Duan et al.* addressed the problem of the local search algorithms that likely produce a huge

number of parameter values trapped as local optima within the objective function optimum space. He and others suggested the use of a global search algorithm, the SCE, which proved to be very effective and efficient at the same time, finding consistently the region of global optimum solutions and requiring not many function evaluations, *(Duan et al., 1992, 1993, 1994; Sorooshian et al., 1993).*

This study examines R-R models that predict a modeled streamflow and is then compared to the observed one. The process of prediction requires that the model parameters are estimated (calibrated) in a way that the modeled flows are as closely related as possible to the observed flows. This implies that during the modeling processes, uncertainty that result in poor predictions must be identified, evaluated and reduced.

Uncertainty in R-R modeling can be described as the degree of confidence that resides in the predictions after models have been calibrated. In general, uncertainty is inherent due to the randomness and the variability in nature and cannot be reduced. However, what can be reduced is the "epistemic" uncertainty which is attributed to the choice of model, **the model structure and the model parameters**. Parameter uncertainty is caused by the choice of the calibration procedure.

In this study, model calibration has been implemented by different parameter search methods, called **Optimization Algorithms**, all of which optimize the value of different *"goodness of fit"* measures, called **Objective Functions.** Therefore, the selection of an appropriate optimization algorithm and a suitable objective function is very significant and has been made on the grounds of reduced parameter uncertainty. This type of uncertainty has been assessed by the use of *"fit statistics"* in both calibration and validation periods.

1.2 Methodology:

Four different conceptual and spatially lumped R-R models have been assessed, namely the **GR4J** *(Perrin, 2000)*, the **AWBM** *(Boughton, 2004)* and two versions of the IHACRES model, the **IHACRES_CWI** *(Jakeman and Hornberger, 1993)* and the **IHACRES_CMD** *(Croke and Jakeman, 2004)*. The same dataset consisting of 11 years of daily Precipitation, Evapotranspiration and Observed streamflow was used for the Yermasoyia watershed in Cyprus. Assessment has been performed during two periods of the same length; one wet and one dry.

The modeling framework was exclusively provided by HYDROMAD, an open-source software available in the R-statistical computing environment. Its structure is consists of two components. First, a Soil Moisture Accounting (SMA) converts inputs of rainfall, temperature, evapotranspiration and others into effective rainfall. The

second component, Routing, is based on the unit hydrograph theory and converts the effective rainfall into streamflow.

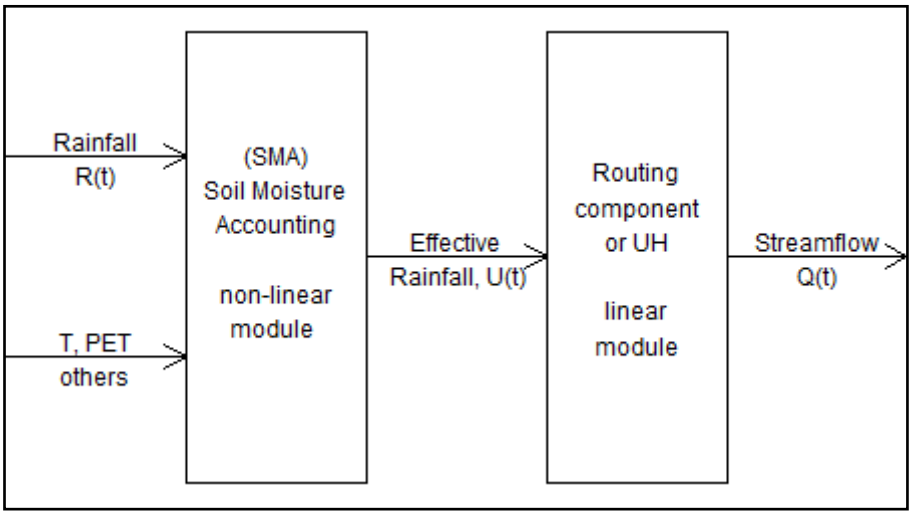


Figure 1. *Rainfall - Runoff model framework*

The objective of this study was to assess the uncertainty and the performance of each model based on different calibration strategies for parameter estimation. Sampling of the parameters was based on both *local search* and non-linear **optimization algorithms** using ***multi-start*** or ***pre-sampling*** (i.e. **Nelder-Mead**, **PORT**: Powerful-Outstanding-Reliable-Tested, **BFGS**: Broyden–Fletcher–Goldfarb–Shanno) and ***global search*** algorithms (i.e. **SANN**: Simulated ANNealing, **SCE**: Shuffled Complex Evolution, **DE**: Differential Evolution and **DREAM**: DiffeRential Evolution Adaptive Metropolis). All algorithms were attempting to optimize four different **objective functions**, namely: the classic **NSE**: Nash-Sutcliffe Efficiency, **NSE^3**: a transformed NSE where the absolute residuals are raised to cubic power and two "bias constraint" objective functions, named **Viney**, as proposed by *Viney et al. (2009)* and the other, named **BL,** proposed by *Bergstrom and Lindstrom, (2002)*.

Finally, several **fit statistics** were used to evaluate model performance in both calibration-validation periods. Indicatively, these include: the **Relative Bias,** the **NSE** as fit statistic too, the **NSE$_{SQRT}$,** and the **AMAFE.** Fit statistics indicate the amount of uncertainty that persists to exist after calibration.

1.3 Study Area:

The watershed of Yermasoyia is a small catchment of approximately 157Km2 area which is located in the mountain of Troodos, near Limassol, Cyprus. The mean annual areal precipitation is about 640mm (450mm at lower elevations up to 850mm at higher elevations) and the mean annual runoff at the catchment outlet is about

150mm or 0.42 m3/s (for the 11years of observations). It was calculated that: for the years '87-'92 runoff is 0.45 m3/s (WET period) and for the years '92-'97 is 0.35 m3/s (DRY period). The catchment's elevation ranges from 70m to 1400m, *(Loukas, et al. 2003)*. The Mediterranean climate of the area produces mild winters and hot and dry summers. Thus, the stream of the watershed is transient and ephemeral with rainfall-induced peak flows being observed during winter months.

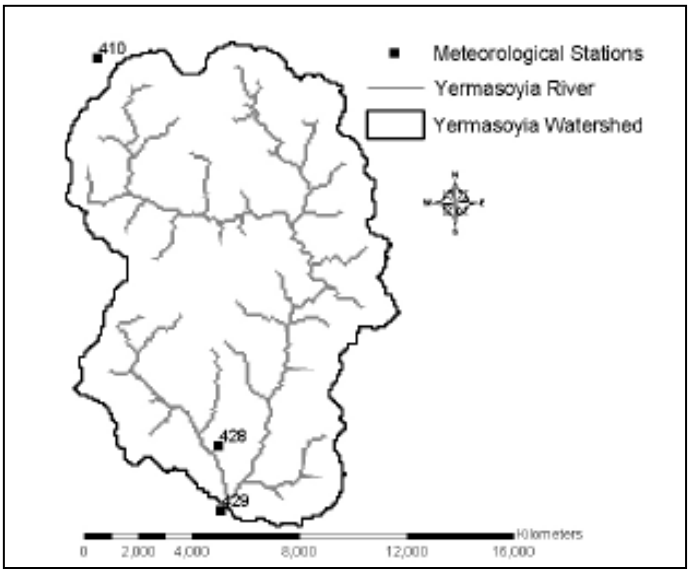


Figure 2. *Yermasoyia watershed, Cyprus (Loukas et al, 2003)*

1.4 Available data and data check:

The available data consists of 11 years of daily precipitation, evapotranspiration and streamflow measurements. Areal precipitation was estimated using the measurements from two stations (at 70m and 995m elevation) by the method of precipitation gradient. Evapotranspiration was estimated by the Hargreaves approach.

Exploratory analysis of the input data is always important since it is a direct input in the models and influences their performance. Although the selected models are, apart from conceptual, also data-based, only an indicative data analysis has been performed. This is because the scope of this study is to explore more the structure of the models and their performance under different optimization methods.

Arguably the first check one should perform is a simple visualization of the available dataset. This may assist in the identification of erroneous data, discontinuities in the time-series, among others, all of which are potential sources of uncertainty in the model assessment. By plotting the raw daily time-series one can assume the

seasonal pattern of ET and a strong correlation between rainfall and streamflow peaks, ***(Figure 3)***.

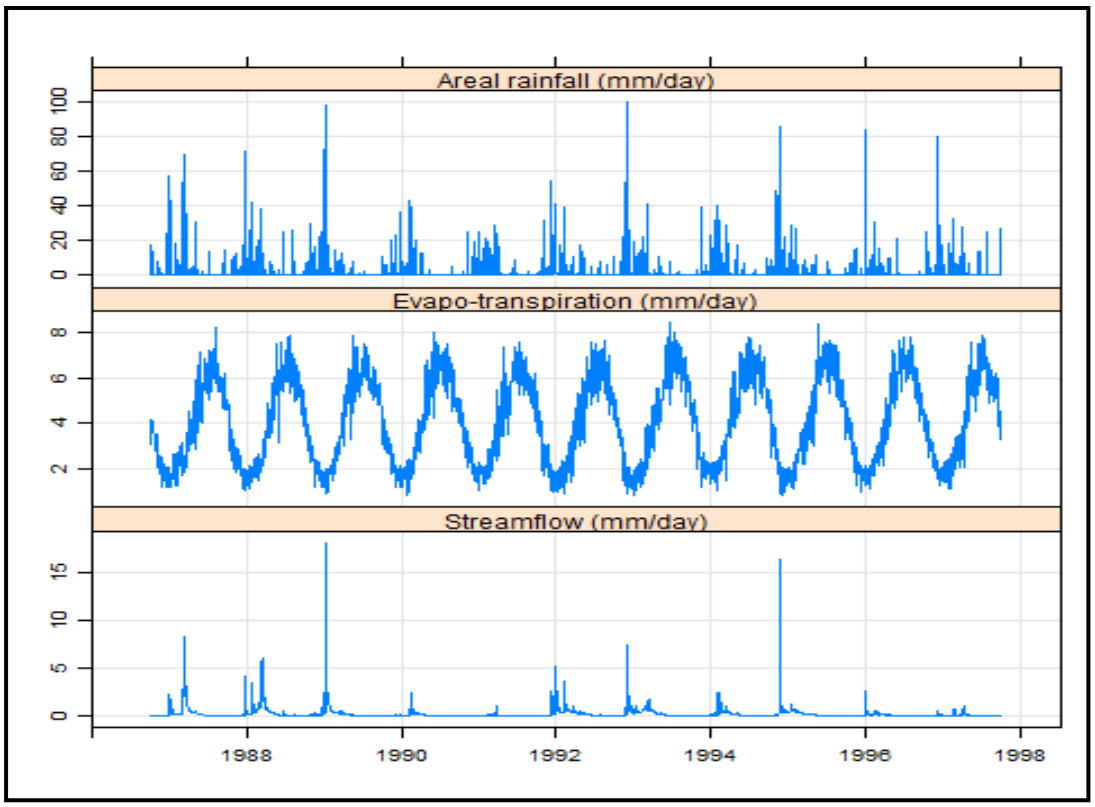


Figure 3. *Daily raw data*

For the purposes of the calibration procedure, a wet period (*form 01-10-1987 to 30-09-1992*) and a dry one (*from 01-10-1992 to 30-09-1997*) have been identified. This is shown more clearly in ***Figure 4*** where the time-series have been aggregated at a monthly scale.

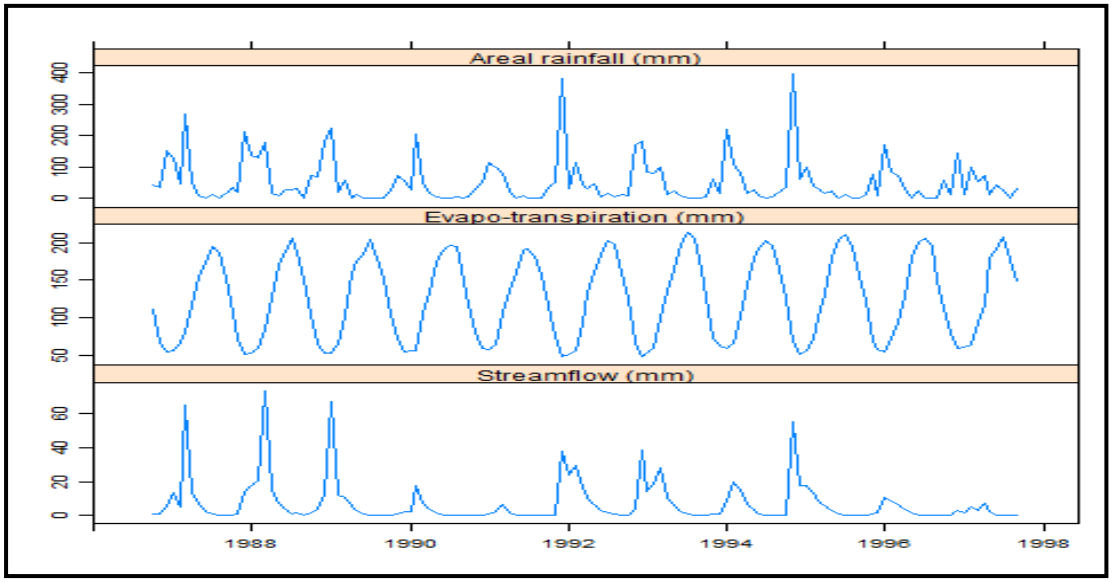


Figure 4. *Monthly aggregated data*

Given that R-R models are being studied here, a simple data check can provide of an estimation of the catchment ***runoff ratio***. This is the percentage of catchment areal precipitation that appears as observed streamflow. Therefore, runoff ratio = sum(Q) / sum(P) = **0.135**. This number appears to be in accordance with what one would expect for such a watershed.

Equally important to the runoff ratio is the estimation of the ***lag-time*** (or ***delay***) between rainfall and rises in the observed streamflow. In this case, the estimate Delay = 1 day.

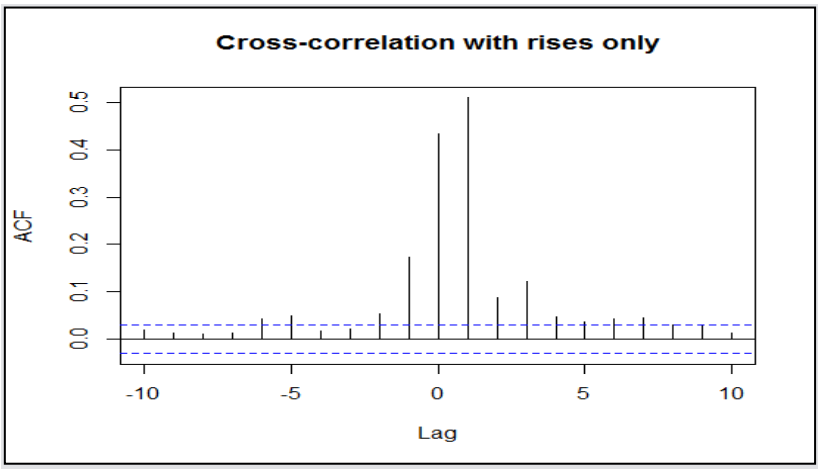


Figure 5. *Lag-time at maximum correlation of P and Q*

The cross-correlation of P and Q can also be examined at given lags (0, 1, 2, etc.) and over different periods (90 days, 365 days, etc).

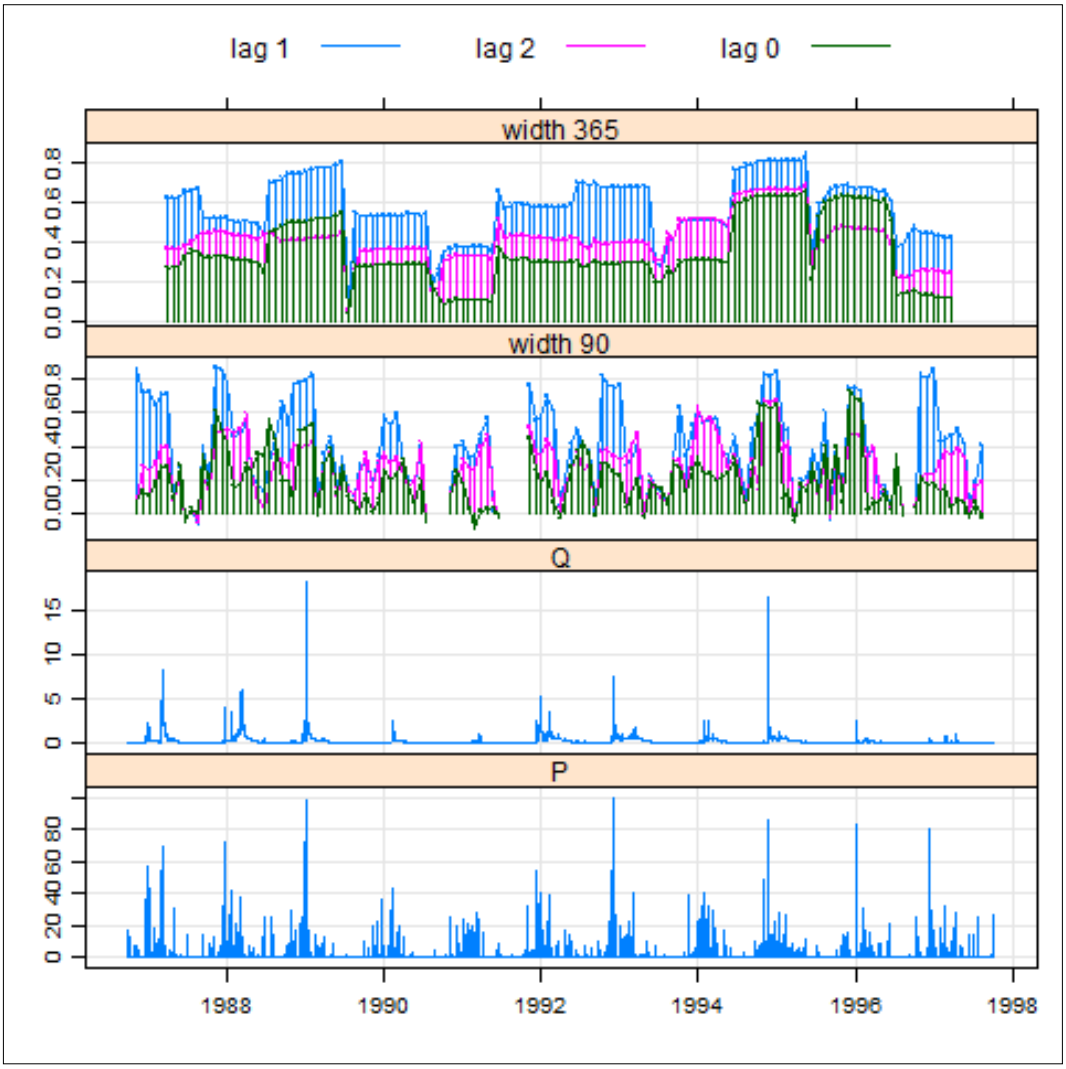


Figure 6. *Cross-correlation of P and Q in different time windows*

This graph also confirms the lag 1 (1 day delay) between rainfall and streamflow where the highest correlations appear. Only in few cases of lag 0, correlation values become higher than those of lag 1, meaning that the response time has changed from 1 to 0 days. Correlation between the input variables was also examined in order to check for non-stationarity issues and if there is linearity in their between relationship.

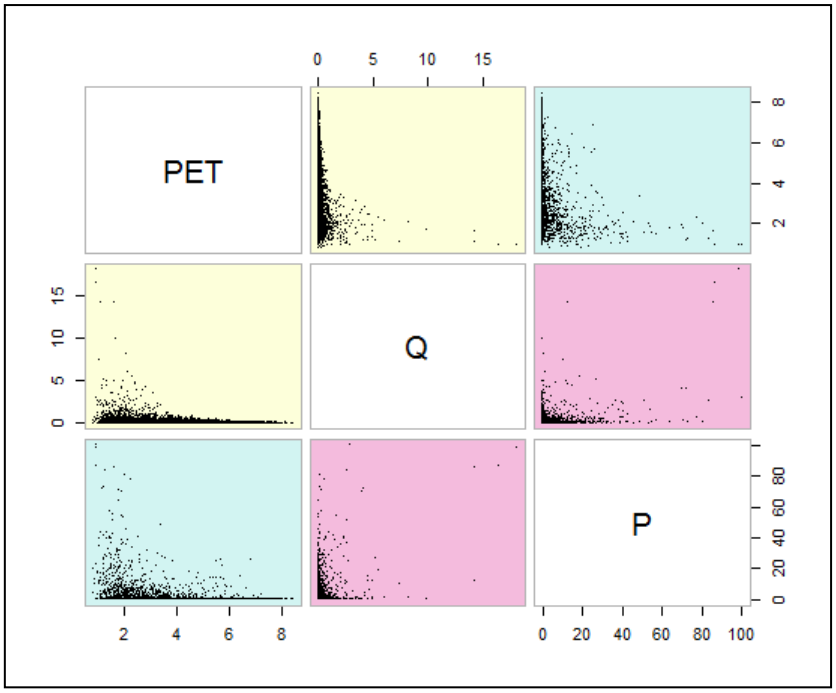


Figure 7. *Correlation matrices of state variable P, Q and PET*

Finally, a seasonal and trend decomposition of rainfall time-series was performed using the STL algorithm (Cleveland et al. 1990).

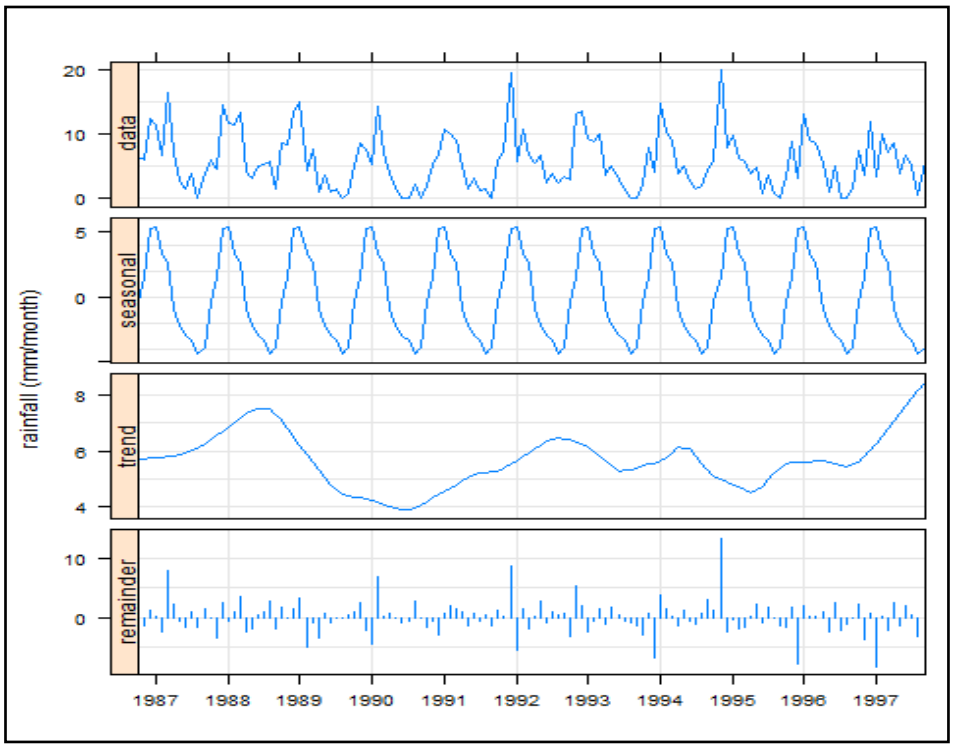


Figure 8. *STL decomposition of time-series*

CHAPTER 2

2.1 General:

Rainfall – Runoff models aim to relate rainfall estimations, (observed or forecast), to the streamflow at a catchment outlet. They can either perform event-based or continuous temporal simulations. Care should be taken into account when selecting the criteria of what defines an event (i.e. thresholds of precipitation / streamflow levels, durations etc). This study will focus only on continuous simulations.

Understanding the physical processes that take place during the rainfall transformation into streamflow is not always easy, considering the variety of hydrological (underground / surface) and climatological profiles that exist. However, with the use of good quality and sufficient quantity of historical records of rainfall, evapotranspiration, temperature and runoff observations, R-R models can produce very satisfactory results requiring a relatively simple structure and only a few parameters. This study will investigate the structure and then the performance of four different R-R models, all of which ran at a continuous daily time-step and are spatially aggregated.

2.2 GR4J model description:

The **GR4J model**, as proposed by *Perrin (2000),* is a daily lumped Rainfall-Runoff model. Daily because the input data required is a daily-step time series of raw areal rainfall and daily estimate of potential evapotranspiration (PET) denoted by **P (mm)** and **E (mm)** respectively. Lumped model, in contrast to a distributed model, because the perceptual basis of the model does not take into account the spatial variability in terms of the hydrological processes and characteristics and treats the catchment as a single unit. (Fig. 9)

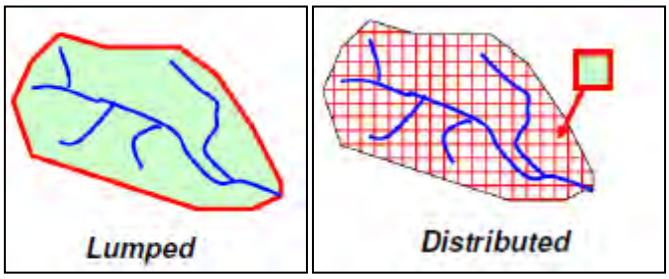


Figure 9. *Spatial scales of R-R models*

Given that this model is a result of a continuous improvement process over many years and acknowledging the findings and suggestions of previous studies such as *Perrin et al. (2001),* the four-parameter version was used with its existing model structure (Fig. 10) as a starting point. The GR4J is an updated version of the GR3J model initially developed by *Edijatno and Michel (1989)* and then further evolved by *Nascimento (1995) and Edijatno et al. (1999)*.

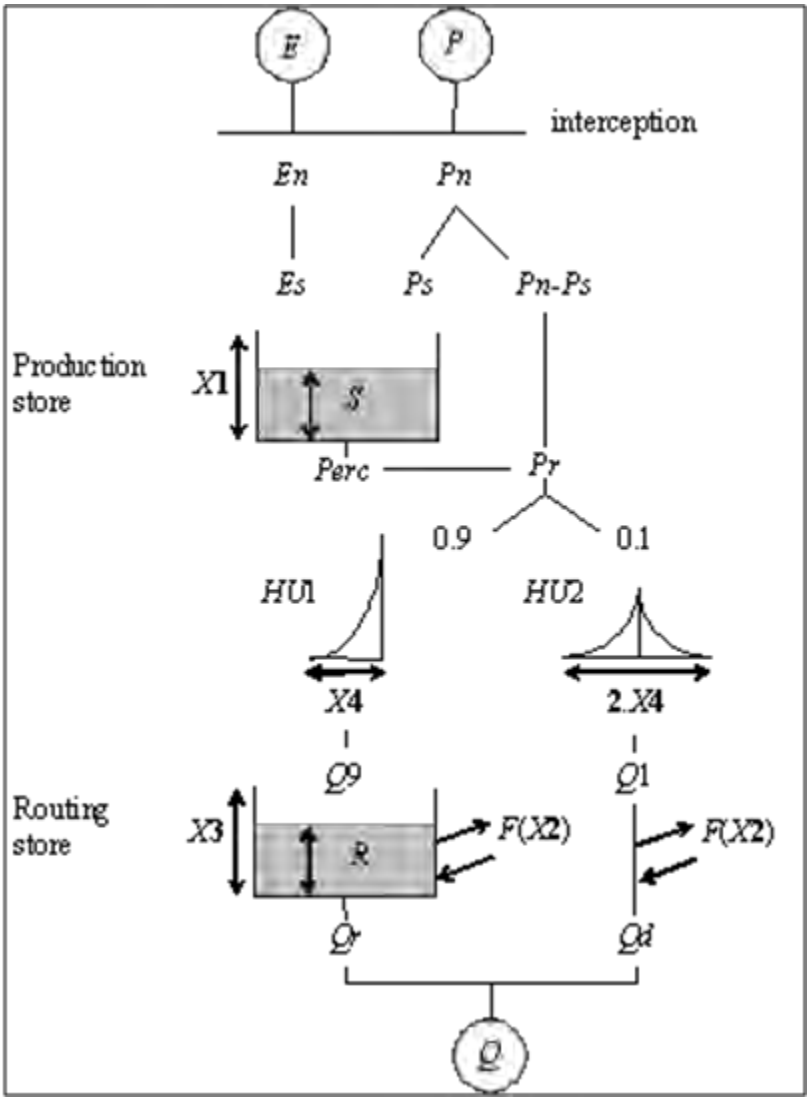


Figure 10. *GR4J model structure (Perrin et al. 2003).*

The four parameters of the model which are calibrated are:

- ***x1***: maximum *capacity of production store* (**mm**)
- ***x2***: *groundwater exchange coefficient*, (**mm**)
- ***x3***: *1-day ahead capacity of routing store*, (**mm**)
- ***x4***: *time base of the unit hydrograph UH1*, (**days**)

The structure of the model consists of two modules. The **SMA** (**Soil Moisture Accounting**) or **Production store** and the **Routing store**.

First, the **SMA** is the part of the model which converts the rainfall and PET into effective rainfall, i.e. rainfall that reaches the outlet of the catchment as streamflow. By subtracting E from P, net values of precipitation ($\mathbf{P_n}$) or evapotranspiration capacity ($\mathbf{E_n}$) are determined, assuming that there is no interception storage capacity. Therefore,

- If P ≥ E, then $P_n = P - E$ and $E_n = 0$ **(1)**
- If P < E, then $E_n = E - P$ and $P_n = 0$ **(2)**

In the case $P_n \neq 0$ a part P_s of P_n fills the Production store given by the function:

$$Ps = \frac{x1\left[1 - \left(\frac{S}{X1}\right)^2\right]\tanh\left(\frac{Pn}{x1}\right)}{1 + \frac{S}{x1}\tanh\left(\frac{Pn}{x1}\right)} \quad \textbf{(3)}$$

where ***x1*** is the maximum capacity of the production store (mm).

In the case $E_n \neq 0$ the production store loses water due to evaporation at a rate E_s given by the following formula:

$$Es = \frac{S\left[2 - \frac{S}{x1}\right]\tanh\left(\frac{En}{x1}\right)}{1 + \left[1 - \frac{S}{x1}\right]\tanh\left(\frac{En}{x1}\right)} \quad \textbf{(4)}$$

Therefore, the water content in the Production store is given by:

$$S = S - E_s + P_s \quad \textbf{(5)}$$

The GR4J model also considers water losses from the Production store due to Percolation, P_{erc} , as a power function of the store capacity:

$$Perc = S\left\{1 - \left[1 + \left(\frac{4S}{9x1}\right)^4\right]^{-\frac{1}{4}}\right. \quad \textbf{(6)}$$

The above mathematical formula implies that P_{erc} is always < S and its contribution to the final streamflow is not great and only important for low flow simulation, *Perrin et al. (2003)*.

The water content in the Production store then becomes:

$$S = S - P_{erc} \quad \textbf{(7)}$$

Finally, the total quantity of water that continues in the **Routing part** of the model is determined by:

$$P_r = P_{erc} + (P_n - P_s) \quad \textbf{(8)}$$

90% of $\mathbf{P_r}$ is routed by the unit hydrograph **UH1** to a non-linear routing store and **10%** of $\mathbf{P_r}$ is routed by the unit hydrograph **UH2**. Both **UH1** and **UH2** simulate the lag time between rainfall events and peak streamflows and depend on the ***x4*** parameter, *time base of the unit hydrograph UH1*, (days), with **UH1** having ***x4*** days as time base and **UH2** having 2***x4*** days (***x4*** > 0.5 days). The two unit hydrographs have ordinates of ***n*** and ***m*** respectively, meaning that the water is spread over a period of time into ***n*** and ***m*** unit hydrograph inputs for **UH1** and **UH2**. These ordinates are calculated by the related S-curves (cumulative proportion of the input with time) and are denoted by ***SH1(t)*** and ***SH2(t)*** respectively.

Table 1. *Ordinates of UH1 and UH2 for the GR4J routing component*

For the SH1(t):		**For the SH2(t):**	
If $t \le 0$:	SH1(t) = 0	If $t \le 0$:	SH2(t) = 0
If $0 < t < x4$:	$SH1(t) = [\frac{t}{x4}]^{\frac{5}{2}}$	If $0 < t < x4$:	$\mathrm{SH2(t)} = \frac{1}{2}[\frac{t}{x4}]^{\frac{5}{2}}$
If $t > x4$	SH1(t) = 1	*If* $x4 < t < 2x4$:	$\mathrm{SH2(t)} = 1 - \frac{1}{2}[2 - \frac{t}{x4}]^{\frac{5}{2}}$
		If $t \ge 2x4$	SH2(t) = 1

Therefore, the ordinates for **UH1** and **UH2** can be calculated by:

$$\mathbf{UH1}(j) = \mathbf{SH1}(j) - \mathbf{SH1}(j\text{-}1) \quad \textbf{(9)}$$

and

$$\mathbf{UH2}(j) = \mathbf{SH2}(j) - \mathbf{SH2}(j\text{-}1) \quad \textbf{(10)}$$

with j and $(j\text{-}1)$ being integers.

The two unit hydrographs provide two outputs, Q_9 and Q_1 at each time-step i and are calculated by the following formulas:

$$\mathbf{Q9} = \mathbf{0.9}\sum_{k=l}^{l} \boldsymbol{UH1}(\boldsymbol{k}).\mathbf{Pr}(\boldsymbol{i}-\boldsymbol{k}+\mathbf{1}) \quad \textbf{(11)}$$

And

$$\mathbf{Q1} = \mathbf{0.1}\sum_{k=l}^{m} \boldsymbol{UH2}(\boldsymbol{k}).\mathbf{Pr}(\boldsymbol{i}-\boldsymbol{k}+\mathbf{1}) \quad \textbf{(12)}$$

Where *l = int(x4) +1* and *m = int(2x4) + 1*.

Both Q_9 and Q_1 are subjected to a groundwater exchange term **F** which is calculated by:

$$F = [x2\frac{R}{x3}]^{\frac{7}{2}} \quad (13)$$

This term introduces the second and third parameter of the model. The 1-day ahead capacity of the routing store, ***x3*** and ***x2*** the water exchange coefficient, i.e. how much water is entering the deeper aquifer. **R** is the level in the non-linear routing store and is estimated by:

$$R = \max(0 ; R + Q_9 + F) \quad (14)$$

And the outflow Q_r from this routing store is determined by:

$$Qr = R\{1 - [1 + (\frac{R}{x3})^4]^{\frac{-1}{4}}\} \quad (15)$$

With the level in the reservoir being:

$$R = R - Q_r \quad (16)$$

Similarly, for the Q_1 output of the **UH2** being subjected to the exchange term **F**, a direct flow component Q_d is added to the total streamflow. This is estimated by:

$$Q_d = \max(0 ; Q_1 + F) \quad (17)$$

Finally, the total streamflow Q_{total} that the model calculates is:

$$Q_{total} = Q_r + Q_d \quad (18)$$

2.3 AWBM description:

As its name indicates, the **AWBM** (Australian Water Balance Model) as proposed by *Boughton, (2004)* is a catchment water balance model that simulates runoff from rainfall (Rainfall – Runoff model) at a daily or hourly time-step. The main concept of the model is the runoff generation based on the *"hortonian"* saturation overland flow. All excess rainfall becomes runoff after the catchment surface capacity has been reached. This simple idea though generates two main issues: Firtly, the initial

soil moisture conditions within a catchment influence the amount of rainfall which is abstracted and, therefore, should be considered. Secondly, the spatial variability within catchments influence both the rainfall abstraction and the runoff generation. Thus, such a variable should also be incorporated to the model structure.

From simple to more complex model structure:

The development of the **AWBM** started from the simple structure of the ***1-bucket model***, Fig. 11(a) to the more complex structures of ***multi-capacity Water Balance (WB) models***, Fig. 11(b)-(c).

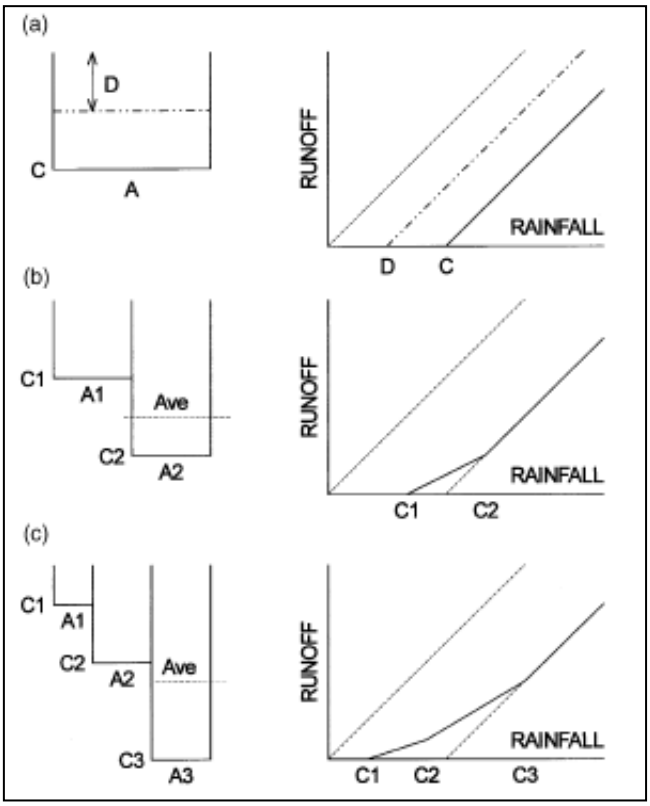


Figure 11. *Single and multi-capacity WB models (Boughton, 2004)*

In the ***1-bucket model***, the catchment surface storage capacity is signified by the model storage capacity **C** and no spatial variability is considered. Any rainfall is abstracted and runoff is generated only after the bucket is filled. This means that if there were no initial soil moisture ammount at the start of a rainfall event, the bucket would be completely empty and all rainfall would be consumed to fill it, as line **C** indicates in the Runoff-Rainfall scheme. In the case of presence of initial soil wetness at the start of a rainfall event, the amount of rainfall abstraction which would be required to fill the bucket could be represented by the line **D** in the same scheme. Finally, in the case of a fully saturated catchment, all excess rainfall would

become runoff and, this is shown by the the 45^0 line from the origin of the axes of the Runoff-Rainfall scheme.

The ***multi-capacity models*** aim to incorporate the spatial variability to their strucuture by taking into account two, three or more different catchment capacities, $\mathbf{C_1, C_2, C_3,\ldots,C_i}$, (say $C_1<C_2<C_3<\ldots<C_i$) and their corresponding partial areas $\mathbf{A_1, A_2, A_3,\ldots, A_i}$. The sum of all areas: $\mathbf{A_1+A_2+A_3+\ldots+A_i = 100\%A_{catchment}}$.

At a rainfall event, runoff occurs when the (smallest) storage capacity $\mathbf{C_1}$ has been reached. When the next larger capacity, $\mathbf{C_2}$ has been reached too, all excess rainfall becomes runoff. The Runoff-Rainfall relationship is again a 45^0 line and, if this line is projected backwards to start at the origins of the axes, it would represent the average storage capacity:
$\mathbf{C_{ave} = C_1A_1 + C_2A_2}$. **(Fig.11(b))**

The partitioning of the catchment into more partial areas and, therefore, storage capacities whould have, as an effect, a smoother curve representing the runoff-rainfall relationship as shown in Fig.11(c), Shifting this curve closer to the origins of the axes would simply mean higher initial soil moisture conditions.

So far, the conventional multi-capacity accounting models were in fact trying to estimate runoff by taking into account the precipitation and the soil moisture deficiencies as weighted indices, considering neither the partial areas of the catchment nor the saturation overland flow, *(Boughton, 2003)*. Also, these models were assuming that all runoff is surface runoff and that there was no contribution from the baseflow. This, of course, is not always true in nature. For the aforementioned hydrological issues, model structures, such as the ones developed in the **AWBM** and the **IHACRES** *(Evans and Jakeman, 1998)* models, take account of the baseflow contribution to runoff, the rational partitioning of the catchment and evapotranspiration. Figure 12 illustrates the structure of the AWMB that has been adopted for the purposes of this project.

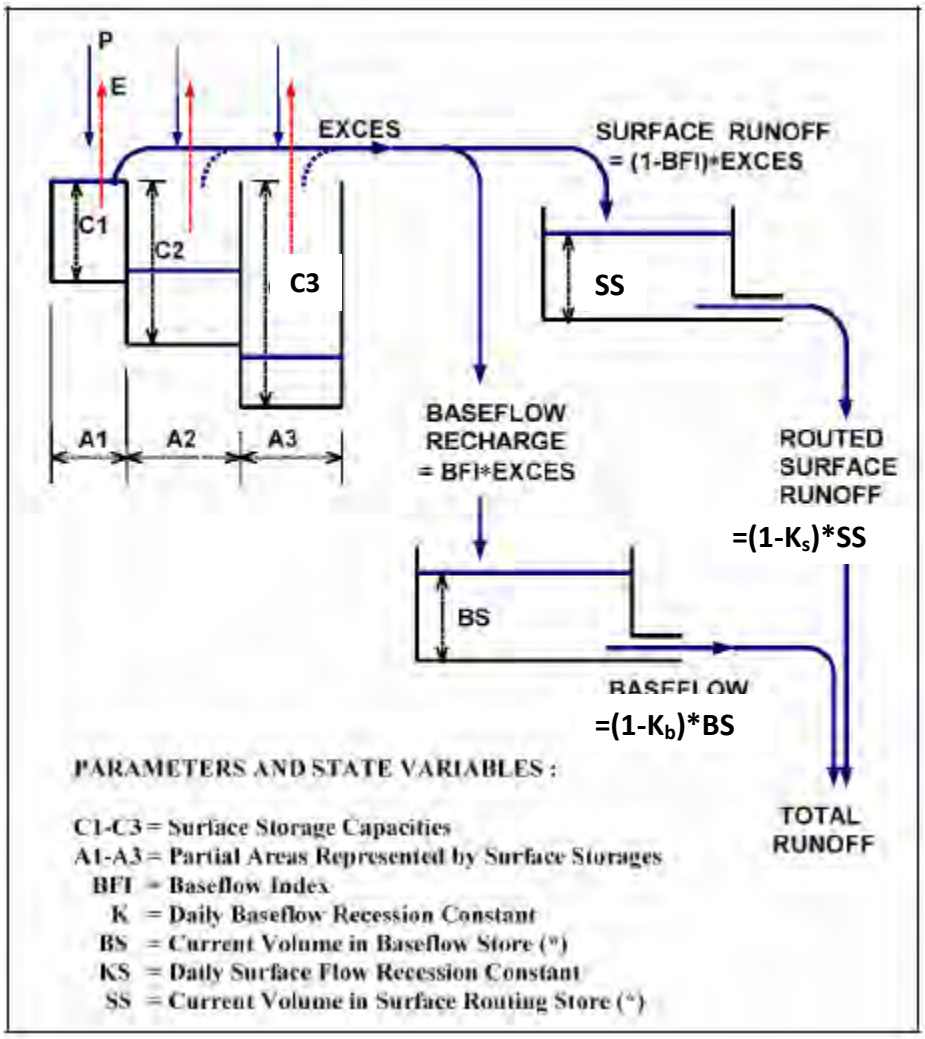


***Figure 12**. AWBM structure (Podger, 2004).*

Similarly to a multi-capacity accounting model, the **AWBM** requires as inputs precipitation and evapotranspiration at a daily or hourly time-step. Precipitation **P (mm)** fills the three surface stores $\mathbf{C_1}$, $\mathbf{C_2}$ and **C3** whereas evapotranspiration **E (mm)** is reduced out of them. Once any of the three stores reaches its capacity, the excess water is added together and separated into two stores.

A fraction of the excess denoted as **Baseflow Recharge = BFI* Excess** is transferred to the **Baseflow store** with current capacity **BS**. BFI is the fraction of total flow that appears as baseflow. The recession of the baseflow at any day is a fixed fraction $\mathbf{K_b}$ of flow at the previous day, where $\mathbf{K_b}$ is the *"recession constant"*. Therefore, at each time step, the amount of baseflow that is reduced from the baseflow store is: **Baseflow = $(1-K_b)$*BS**

The residual runoff volume that is the **Surface Runoff = (1 – BFI)*Excess** is transferred to the **Surface Store** with current capacity **SS**. Similarly, the amount of surface runoff that is reduced from the surface store is: **Routed Surface Runoff = $(1 – K_s)$*SS,** where $\mathbf{K_s}$ is the recession constant of the surface runoff. Both baseflow and

surface recession constants K_b and K_s can be determined directly from the streamflow record, *(Klaassen and Pilgrim, 1975).*

2.4 IHACRES model description:

The **IHACRES** model as initially proposed by Jakeman et al. (1993) consists of two modules. A ***non-linear*** or ***rainfall loss module*** through which Rainfall $\mathbf{R_{(t)}}$ is transformed into "Excess" or Effective Rainfall $\mathbf{U_{(t)}}$**.** The ***linear module*** converts / routes the excess rainfall $\mathbf{U_{(t)}}$ into streamflow $\mathbf{Q_{(t)}}$**.** A generic configuration of this model structure is shown in Fig.13 below:

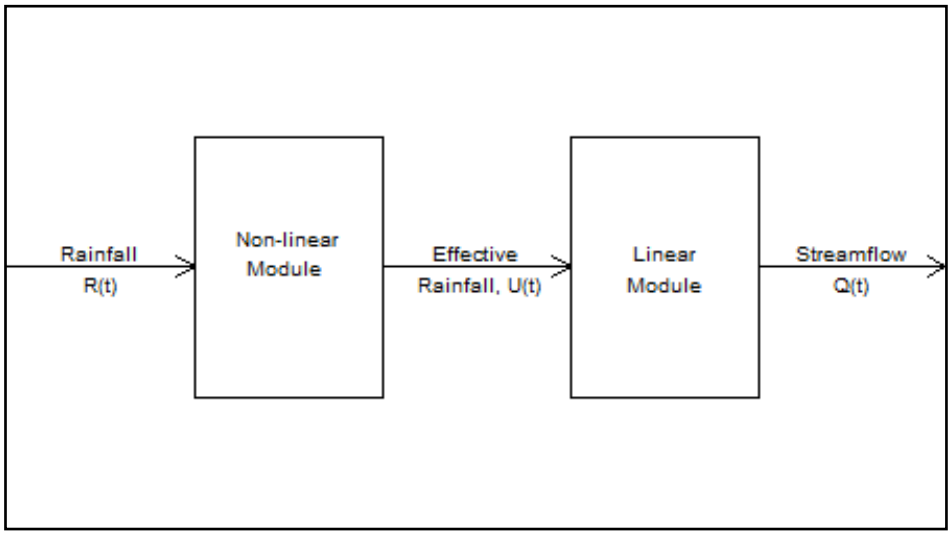


Figure.13 *IHACRES model generic structure (Jakeman and Hornberger, 1993).*

Developments and advancements of the original IHACRES model include different approaches to deal with the non-linear part of the model, rendering it more flexible to explain the hydrological processes or even the climate and land use changes. This study explores two different versions of IHACRES, the more physically-based **CWI** [**Catchment Wetness Index**, by *Jakeman and Hornberger (1993)*] and the **CMD** [**Catchment Moisture Deficit**, by *Croke and Jakeman, (2004)*]. As far as the linear part of IHACRES is concerned, different configurations of routing stores in parallel and/or in series can be deployed depending on the catchment characteristics.

2.5 IHACRES model - CWI version:

In the original IHACRES model by *Jakeman and Hornberger (1993),* at a time step (t), the excess or effective rainfall $\mathbf{U_{(t)}}$ is proportional to rainfall $\mathbf{P_{(t)}}$ and scaled by a soil moisture index $\mathbf{s_{(t)}}$:

- $\mathbf{U_{(t)} = c * s_{(t)} * P_{(t)}}$ **(1a)**

Where $P_{(t)}$ is the observed rainfall,

- $s_{(t)} = P_{(t)} + (1 - 1/\tau_{d(t)}) * s_{(t-1)}$ **(2)**

- $s_{(t)}$ is the catchment wetness index which decays exponentially backwards in time
- **c** is a coefficient for the maintenance of the mass balance (Volume of excess rainfall = total streamflow in calibration period) and
- $\tau_{d(t)}$ is the drying rate of the catchment, given by:

- $\tau_{d(t)} = \tau_w * \exp(0.062 * E_{(t)} * f$ **(3),**

Where τ_w is the drying rate at reference temperature (i.e. number of time steps to reduce by 1/e or 37%), $E_{(t)} = (T_{REF} - T_{(t)})$ with T_{REF} usually 20^0**C**, $T_{(t)}$ is the temperature at time step (t) and **f** is a temperature modulation (i.e. how $\tau_{d(t)}$ changes with temperature *(Jakeman and Hornberger, (1993).*

This version of IHACRES has been further extended to function for more ephemeral catchments by *Ye et al. (1997).* This involved the introduction of a ***moisture threshold*** parameter for producing flow, denoted by ***l*** and a non-linear relationship between the catchment wetness index and the fraction of rainfall that becomes effective rainfall, referred as ***power law with exponent parameter, p***. Therefore, the previous equation **(1)** can be revised to:

- $U_{(t)} = [c * (s_{(t)} - l)]^p * P_{(t)}$ **(1b)***

* NOTE: in case ***l*** is set to zero and ***p*** to one, equation (1b) describes the original *Jakeman and Hornberger, (1993)* IHACRES model.

2.6 IHACRES model - CMD version:

This version of IHACRES is a conceptual model and the main characteristic is the portioning of the input rainfall into drainage (or effective rainfall), evapo-transpiration and changes in catchment moisture *(Croke and Jakeman, 2004)*.

The drainage (effective rainfall) $U_{(t)}$ at a time-step **(t)** is given by the following equation:

- $U_{(t)} = M_{(t)} - M_{(t-1)} - ET_{(t)} + P_{(t)}$ **(1)**

Where $M_{(t)}$ and $M_{(t-1)}$ are the CDM at time-steps (t) and (t-1), $P_{(t)}$ is the catchment areal rainfall and $ET_{(t)}$ is the evapo-transpiration, all in (mm).

The ratio **dU/dP** represents the rainfall effectiveness (or the drainage proportion). According to *Croke and Jakeman, (2004),* this ratio is a function of CDM, with a threshold of CDM for producing flow, ***d*** and is described according to the **dU/dP** relationship by the following equation:

$$\mathbf{U/dP} = \mathbf{1} - \boldsymbol{min}(\mathbf{1}, \frac{\boldsymbol{M}}{\boldsymbol{d}})\mathbf{d}$$

for linear relationship, **(2a)**

$$\mathbf{dU/dP} = \mathbf{1} - \boldsymbol{min}[\mathbf{1}, \boldsymbol{sin}^2\left(\boldsymbol{\pi} * \frac{\boldsymbol{M}}{\mathbf{2}\boldsymbol{d}}\right)]$$

for trigonometric, **(2b)**

$$\mathbf{dU/dP} = \mathbf{1} - \boldsymbol{min}[\mathbf{1}, (\frac{\boldsymbol{M}}{\boldsymbol{d}})^{\boldsymbol{b}}]$$

for a power form **(2c)**

The integration of these relations provides the actual drainage (effective rainfall) at each time-step.

Finally, the evapotranspiration term $\mathbf{ET_{(t)}}$ is a proportion of the potential rate $\mathbf{PE_{(t)}}$ and is also a function of CMD, with a threshold of ***M' =f* d*** and is calculated by the following equation:

$$\mathbf{ET_{(t)} = e^* PE_{(t)}^*/min(exp(2(1-M''/M')))} \quad \mathbf{(3)}$$

where ***e*** is a temperature to PET conversion factor, ***f*** is a CMD stress threshold as a proportion to the ***d*** threshold and $\mathbf{M''}$ is the CMD after precipitation and drainage have been accounted for.

CHAPTER 3

3.1 General:

Before testing the model performance, the selected Rainfall-Runoff models had to be specified. This is a crucial first step in hydrological modeling where important decisions have to be taken regarding the following:

- The ***conceptual model formulation / structure*** (choice from a variety of model versions)
- The way the models handle the ***physical processes***
- The ***mathematical equations*** that resolve these natural processes
- The ***use of the available data***
- The ***number***, ***range*** or ***fixed parameter values*** that generate consistently optimal model performance
- The ***scale of the model parts and parameters*** that are valid and finally,
- The ***application of the model*** and the ***use of the results.***
-

A prudent "rule of thumb" when specifying models is to keep them not too simple but not too complex as well. Oversimplified models obviously are not flexible enough to explain adequately the physical processes involved. On the other hand, with too complex models (over-parameterized, models with too many mathematical functions or models trying to answer too complex questions etc.) comes the cost of increasing uncertainty in the model predictions *(Croke and Jakeman, 2004)*.

Therefore, ideal rainfall-runoff models are those encompassing a structure that best explain the physical processes in catchment hydrology. Such model structure should not be fixed but rather be flexible to include catchment behavior that varies. At the same time, maintaining parsimony in the model parameters is also desirable. The less parameters that suffice to describe the rainfall-runoff relationship the more appealing the model is.

In this study, the same dataset of areal precipitation, evapotranspiration and observed streamflow, for the Yermasoyia watershed, was used in all the models. All inputs were in mm/day.

A general two-component structure was deployed in all models. Namely, an SMA module that converts input data (P and E) into effective rainfall and a Routing module that converts effective rainfall into streamflow. . All models were set to function with a warmup period of 365days.

Where necessary, (see AWBM, IHACRES – CWI/CMD model specification) and based on the unit hydrograph theory, the routing parameters were determined by fitting a linear transfer function. In other words, an ARMAX-like model was specified, (Autoregressive, Moving Average), with autoregressive terms =n and moving average terms = m, i.e. the order of the transfer function. To ensure a good choice for the

order of the ARMAX model, HYDROMAD provides a built-in command called "tryModelOrders", which tests different model structures (n, m ,delay between effective rainfall and runoff). This command enables the fitting of the ARMAX models with the SRIV (Simple Refined Instrumental Variable) algorithm as described by *Young et al. (2008)*. Trials in different model structures produce various values of R^2, criterion of determination, and %ARPE, (Average Relative Parameter Error). Best values of these statistics indicate which order of the transfer function is best.

3.2 GR4J model specifications:

This model has a definite structure regarding its SMA component (as described previously) and the default parameter ranges were taken from the 80% confidence intervals *(Perrin et al., 2003)*. The proposed "fixed" split of 10% and 90% of effective rainfall was also selected. Initially, the choice for the model structure, the number of parameters and their value ranges was founded on the empirical developments (more than 200 different model versions) and the results of model performance in a large sample of catchments (more than 400) with different climatic conditions *(Perrin et al., 2003)*.
However, during the initial attempts to calibrate the parameters, most of the algorithms were converging towards the default lower bound of -5mm regarding the X2 parameter (groundwater exchange coefficient), which is the amount of water that enters the deeper aquifer. This possibly means that the algorithms were "entrapped" to converge to a local minimum. Therefore, the bounds of x2 parameter were extended from (-5, 3) to (-25, 5) mm, in order to overcome this constraint and remove this type of possible uncertainty. (This is shown in the results section: 6.1.1).

The selected model specification is summarized in the next page:

Default model specification by *(Perrin et al., 2003)*:

```
Hydromad model with "gr4j" SMA and "gr4jrouting" routing:
Simulation Start = 1986-10-01, End = 1992-09-30, warmup= 365 days

SMA Parameter Ranges:
       lower upper
x1       100  1200  (mm) – SMA max. capacity
etmult     1     1   (multiplier for the E data)

Routing Parameters:
       lower upper
x2    -25.0   3.0  (mm) – groundwater exchange coefficient.
x3     20.0 300.0  (mm) - Routing store capacity
x4     1.1   2.9   (days) – UH time base

S0 = 0 (initial soil moisture level as fraction of x1)
R0 = 0 (initial groundwater reservoir level as fraction of x3)
```

3.3 AWBM model specifications:

The 3-bucket structure of the AWBM has been used in the model specification, exactly as described previously. The excess water from three different capacities, with weighted areas, is added together and then is divided into two stores. The Routing component of the model has three parameters: **BFI**, $\mathbf{K_b}$ and $\mathbf{K_s}$. These parameters are either directly provided to the model based on streamflow records, or in this case, they were calculated using a transfer function (effective rainfall transformed into component runoff) with exponentially decaying components.

This transfer function is denoted by "expuh", i.e. exponential unit hydrograph and, its decaying runoff components are described by a recession rate α and a peak response β. In Hydromad, these characteristics are better explained by time constants τ (time-steps required to reduce to 1/e or 37%) and volume fractions v respectively, where:

$$\tau=-1/\log(\alpha) \quad \textbf{(1)}$$

$$v=\beta/(1-\alpha) \quad \textbf{(2)}$$

The routing runoff components are usually two (one quick and one slow component) having time constants τ _q and τ _s and fractional volumes v_q and v_s. They can be arranged in parallel or in series configuration*. In this study the model has two stores in parallel and their sum is the total simulated runoff, calculated by following transfer functions:

$$\mathbf{Q\_s(t) = \alpha\_s\, Q\_s(t-1) + \beta\_s\, U(t)} \quad \textbf{(3)}$$

$$\mathbf{Q\_q(t) = \alpha\_q\, Q\_q(t-1) + \beta\_q\, U(t)} \quad \textbf{(4)}$$

These transfer functions of order (2,1) were fitted by the Simple Refined Instrumental Variable Method, SRIV *(Young, 2008).* Other methods include the Least Squares or the Inverse Filtering fittings which were also tried but did not function as good as the SRIV.

***Note :** the AWBM can also be arranged to have three routing stores. In this case their configuration can have four possible types: 3 parallel, 3 in series, 2 parallel and 1 in series, 1 parallel and 2 in series, *(Jakeman et al. 1990).*

The model specification used in this study is based on the self-calibrating version of the original model, the AWBM2002. According to *Boughton et al. (2003),* findings from high quality datasets that demonstrated very high correlation between actual and simulated runoff, reinforced the importance of the average surface storage capacity, $\mathbf{C_{ave} = C_1A_1 + C_2A_2 + C_3A_3}$. Their study suggests an average pattern that relates acceptably the three capacities and the three partial areas as followes:

```
A1 = 0.134: Partial area of smallest store
A2 = 0.433: Partial area of middle store
A3 = 0.433: Partial area of largest store
C1 = 0.01*Ave/A1 = 0.075*Ave: Capacity of smallest store
C2 = 0.33*Ave/A2 = 0.762*Ave: Capacity of middle store
C3 = 0.66*Ave/A3 = 1.524*Ave: Capacity of largest store
```

The rest of the specifications given to the model in this study are summarized as follows:

```
Hydromad model with "awbm" SMA and "expuh" routing:
Start = 1986-10-01, End = 1992-09-30, warmup = 365 days


SMA Parameters:
        lower upper
cap.ave  1.00  1000  (mm)
etmult   0.01     1  (multiplier for the E data)


Routing Parameters:
NULL i.e (t_q, t_s and v_q, v_s) will be calculated by:

Routing fit spec.: list("sriv", order = c(2, 1))
```

The above order (2, 1) of the transfer function was calculated by fitting the ARMAX model with the SRIV algorithm. The results are shown below:

***Table 2**. AWBM model, specification of routing structure. ARPE and fit statistics calculated by fitting unit hydrograph transfer functions of different orders*

	ARPE	r.squared	r.sq.log
(n=0, m=0, d=0)	0.000	0.277	-1.072
(n=1, m=0, d=0)	0.000	0.689	-0.131
(n=1, m=1, d=0)	0.006	0.483	-0.185
(n=2, m=0, d=0)	NaN	0.501	0.724
(n=2, m=1, d=0)	0.000	0.756	0.854
(n=2, m=2, d=0)	0.012	-0.684	0.740
(n=3, m=0, d=0)	NaN	0.499	0.750
(n=3, m=1, d=0)	NaN	0.709	0.843
(n=3, m=2, d=0)	NaN	-26783.899	-0.238
(n=3, m=3, d=0)	55.692	0.325	0.587

Therefore, a structure of (2, 1) was selected because it appears to yield better statistics.

3.4 IHACRES model - CWI version specifications:

Similar to the AWBM, a structure of two runoff components in parallel has been used in this study. The model specification is summarized below:

```
Hydromad model with "cwi" SMA and "expuh" routing:
Start = 1986-10-01, End = 1992-09-30

SMA Parameters:
      lower upper
tw        0   100      (drying rate at 20°C)
f         0     8      (temperature modulation)
scale    NA    NA
l         0     0 (==) (As in the original IHACRES model)
p         1     1 (==) (As in the original IHACRES model)
t_ref    20    20 (==)

Routing Parameters:
NULL i.e (t_q, t_s and v_q, v_s) will be calculated by:

Routing fit spec.: list("sriv", order = c(2, 1))
```

The above order (2, 1) of the transfer function was calculated by fitting the AMAX model with the SRIV algorithm. The results are shown below:

***Table 3**. IHACRES model –CWI version, specification of routing structure. ARPE and fit statistics calculated by fitting unit hydrograph transfer functions of different orders*

	ARPE	r.squared	r.sq.log
(n=0, m=0, d=0)	0.000	0.258	-0.703
(n=1, m=0, d=0)	0.000	0.664	0.411
(n=1, m=1, d=0)	0.002	0.227	0.230
(n=2, m=0, d=0)	NaN	0.613	0.589
(n=2, m=1, d=0)	0.001	0.690	0.754
(n=2, m=2, d=0)	0.006	-0.393	0.666
(n=3, m=0, d=0)	NaN	0.610	0.596
(n=3, m=1, d=0)	NaN	0.647	0.797
(n=3, m=2, d=0)	NaN	-0.165	0.672
(n=3, m=3, d=0)	52.759	0.354	0.530

Therefore, a structure of (2, 1) was selected because it appears to yield better statistics.

3.5 IHACRES model - CMD version specifications:

Similarly to the previous models, a two runoff component in parallel was used. The rest of the model specification is summarized below:

```
Hydromad model with "cmd" SMA and "expuh" routing:
Start = 1986-10-01, End = 1992-09-30

SMA Parameters:
      lower  upper
f      0.01    3.0       (CMD stress threshold as a proportion of d)
e      0.01    1.5       (temperature to PET conversion factor)
d     50.00  550.0       (CMD threshold for producing flow)
shape  0.00    0.0 (==)  (Linear dU/dP relationship )

Routing Parameters:
NULL i.e (t_q, t_s and v_q, v_s) will be calculated by:

Routing fit spec.: list("sriv", order = c(2, 1))
```

The above order (2, 1) of the transfer function was calculated by fitting the AMAX model with the SRIV algorithm. The results are shown below:

***Table 4**. IHACRES model –CMD version, specification of routing structure. ARPE and fit statistics calculated by fitting unit hydrograph transfer functions of different orders*

	ARPE	r.squared	r.sq.log
(n=0, m=0, d=0)	0.000	0.343	-0.910
(n=1, m=0, d=0)	0.000	0.706	-0.278
(n=1, m=1, d=0)	0.001	-0.690	-0.343
(n=2, m=0, d=0)	NaN	0.627	-0.040
(n=2, m=1, d=0)	0.000	0.706	0.563
(n=2, m=2, d=0)	2.160	0.569	0.523
(n=3, m=0, d=0)	NaN	0.628	-0.046
(n=3, m=1, d=0)	NaN	0.688	0.566
(n=3, m=2, d=0)	NaN	-16.299	-0.080
(n=3, m=3, d=0)	NaN	0.161	0.417

CHAPTER 4

4.1 Calibration of Models:

Calibration is the next step in the assessment of hydrological models. The selected Rainfall-Runoff (R-R) models are spatially-aggregated (lumped) and their mathematical structures are more conceptual-based. With these R-R models streamflow was predicted from input time-series of P, PET and $Q_{obs.}$ The model parameters that define the R-R relationship were estimated or adjusted in such a way that the modeled runoff would match as much as possible the observed streamflow and also, were assumed to be stationary during the calibration period.

The available input data consists of daily measurements for 11 hydrological years between 01-10-1986 and 30-09-1997. Based on the streamflow observations, two temporal different periods were identified. A **wet period** *from 01-10-1987 to 30-09-1992* and a **dry period** *from 01-10-1992 to 30-09-1997*. Using a split-sample test based on the wet/dry periods, the models were calibrated for half of the years during the wet and the dry periods leaving the rest half of the years for validation. A warmup period of 365 days was used in all cases. This allowed the models to acquire information of the rainfall and therefore, for the initial soil moisture conditions, prior to the start. Therefore, two sets of Calibration-Validation periods were defined:

Set A:

Calibration period:

TS1 <- window(dataz, start="1986-10-01", end="1992-09-30")

Validation period:

TS2 <- window(dataz, start="1991-10-01", end="1997-09-30")

Set A:

Calibration period:

TS3 <- window(dataz, start="1991-10-01", end="1997-09-30")

Validation period:

TS4 <- window(dataz, start="1986-10-01", end="1992-09-30")

The performance of all four models had to be assessed both individually and in comparison with each other, using seven different optimization techniques all of which would try to optimize the value of four different objective functions. Calculation of several fit statistics (which will be presented in the Validation chapter) was the typical approach to assess the models during the specified (wet and dry)

calibration periods. This rendered the automation of the calibration process imperative.

4.2 Optimization Algorithms:

In general, the algorithms plot the objective function values against the model parameters corresponding to the *response surface*. There are two main categories of optimization algorithms: **Local Search Algorithms**, (**LSA**, finding a lowest value of OF in the near vicinity) and **Global Search Algorithms** (**GSA**, finding the lowest value Of the OF in the whole response area).

Seven different optimization algorithms were used at each calibration test for all four R-R models. Using the command ***"fitByOptim"*** in R, one can select:

- the method/algorithm for the calibration,
- the objective function to be minimized,
- the sample/number of parameters to be tested (here, **samples=100**) and,
- the maximum evaluations/iterations to be performed (here, **maxeval=1000**)

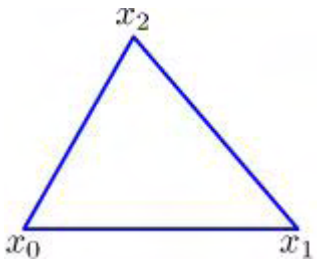
This command also allows the choice between a single and a multi-start mode of sampling the initial parameters. By trial, **multi-start mode** required a lot more time for the algorithms to converge and at the same time it did not improve significantly their performance in general and hence, **single-mode sampling** was selected. Another important setting in the calibration process was to maintain the same ***"seeding"*** by selecting set.seed(0) in R. This means that algorithms iterate in the same way each time they run and give the same result.

The assessment and evaluation of each algorithm's performance is performed by plotting their optimization traces in a diagram of "Objective Function Value versus Number of Function Evaluations", (presented in the Results section).

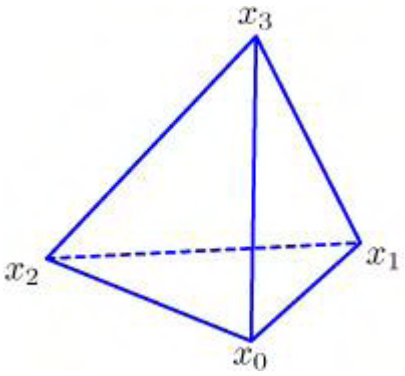
A brief description of the all the optimization algorithms deployed and how they function is summarized below:

- **The "Nelder-Mead":** *(source of images: www.scholarpedia.org)*

NM is a ***non-linear*** optimization method that identifies points (simplices) in the parameter space that produce an OF with one ***local minimum***. It uses only the OF values (direct search method) to the decision process. A simplex S is defined as the set of n+1 points, $x_0,\ldots,x_n \in R^n$ (R^n is the parameter space). This simply means that if n=2 or n=3, then a simplex in R^2 or in R^3 would be a set of 3 points (x_0, x_1, x_2, forming a triangle) or 4 points (x_0, x_1, x_2, x_3 forming a tetrahedron) respectively.



3-point Simplex



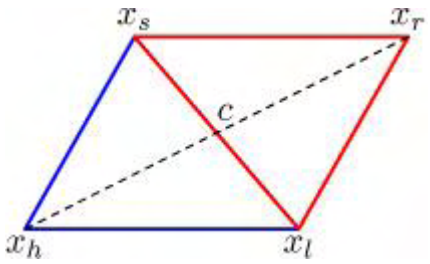
4-point Simplex

The x points are assigned with values of objective function, f_l, f_s, f_h, so that:

$x_0 < x_1 < x_2 (< x_3)$ and can be re-written as: x_l, x_s, x_h

The NM algorithm performs transformations (iteratively) to the initial size and shape of the simplex by:

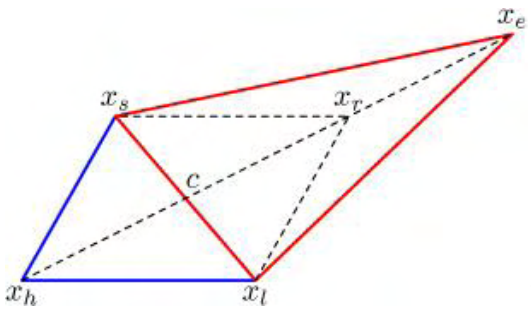
-***Reflection***: of the point with the highest OF value, x_h, through the centroid c formed by the rest best points, at the opposite side of x_h. The new point is called the reflection point x_r as shown below



, if $f_l < f_r < f_s$ then x_r is accepted and the iteration can be terminated, otherwise:

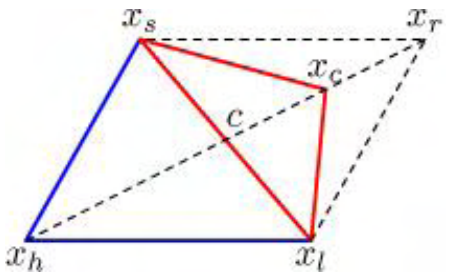
If $f_r < f_l$ then the simplex is transformed by either:

- ***Expansion***: of the reflection point x_r on the c-x_r line creating a new point x_e as shown below:

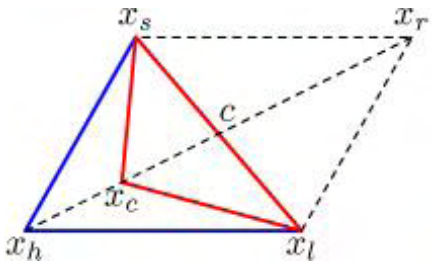


if $f_e < f_r$ then x_e is accepted and iteration terminates, otherwise if $f_r < f_e$ then x_r is accepted and the iteration is terminated or, alternatively, the algorithm can change direction in its search for a minimum value of the OF by:

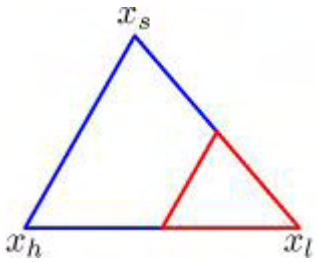
- ***Contraction***: of the x_h or the x_r to create a new point x_c, depending on which one has bigger value of OF. If $f_h < f_r < f_s$ (outside contraction), then the contracted simplex shape becomes as the one shown below:



, if $f_c < f_r$ then x_c is accepted and the iteration terminates. In the opposite case, a **shrinkage** transformation in the simplex will be performed. If $f_r > f_h > f_s$ inside contraction), then the contracted simplex shape becomes:



similarly, if $f_c < f_r$ then x_c is accepted and the iteration terminates. Otherwise again, a **shrinkage** transformation should be done. Shrinkage of the simplex is performed by moving each point (x_h and x_s) except x_l by half way towards x_l as shown in the picture below:



and the algorithm follows the same procedures as before.

- The **"PORT"** functions describe a ***gradient search*** method for finding a ***local minimum***. It uses the OF f(x), the gradient G(x) (vector of 1^{ST} partial derivatives of f(x), and the Hessian of the OF (matrix of 2^{nd} partial derivatives of f(x).

- The **"BFGS"** is a *quasi-Newton non-linear* optimization method that uses both OF values and ***gradients*** to identify **local minimum**. It builds a *quadratic* model of the OF where $G(x_{min})=0$ at the optimum value and requires the calculation of the inverse Hessian matrix.

The next four algorithms are Global Search

- The **"SANN"** algorithm (Simulated Annealing) belongs to the **probabilistic** methods of optimization of a given objective function and returns a **global minimum** solution.

- The **SCE** (Shuffled Complex Evolution) is a probabilistic population-sampling evolutionary algorithm. Initially, the algorithm samples populations at random and then divides them into complexes based on their objective function value. Then, these complexes evolve and are being optimized using techniques similar to the Nelder-Mead simplex method. When this procedure stops, the partitioned populations are combined back together again, sharing information and then they are divided similarly to the previous step. The procedure is repeated until the algorithm converges to a global optimum.

- The **DE** (Differential Evolution) is a stochastic global optimization algorithm. Similar to other evolutionary algorithms, DE optimization is suitable for functions that are either continuous or differentiable.

- The **DREAM** is a multiple Markov Chains Monte Carlo method and searches for global solution.

4.3 Objective Functions:

Objective functions are measures of how much model output and observed output (in this study streamflows) differ. Optimization algorithms *adjust the model parameter values iteratively* until they *converge* (i.e until optimum value of Objective Function has been discovered). Four different Objective Functions were applied in the calibration process.

- **The Nash-Sutcliffe Efficiency, NSE:**

This Objective Function (developed by Nash and Sutcliffe, 1970) is a traditional and straight-forward method for model assessment. It measures the degree of agreement between modeled and observed streamflow by the following formula:

$$NSE = 1 - \frac{\sum_{i=1}^{n} \left(Qobs_i - Qsim_i \right)^2}{\sum_{i=1}^{n} \left(Qobs_i - \overline{Qobs} \right)^2}$$

The numerator of the fraction is the residual variance or "Noise" and the denominator is the variance of the flows from the mean observed. Calibration methods, based on the maximization of the NSE criterion, aim to reduce the variance error. Usually, a residual analysis should be performed to identify heteroscedastic errors.

- **The "Viney" Objective Function:**

This OFs, proposed by *Viney et al. (2009)*, include bias constraint terms in the NSE. These are simply "penalty" terms subtracted from the NSE in a similar way to many other modelers. In their work, Viney et al, tested three different OF in the model calibration. The classic NSE, a biased constraint OF that assigns a penalty (proportional to the % of the **Bias, B = $\sum (X - Q)$**) to any prediction that has overall bias (total model error / total observed flow) greater than 5% (called "bucket constraint") and a "log-bias"-penalty OF defined by:

$F = NSE - 5 * [\ln(1+B)]^{2.5}$

They have found out that the log-bias constraint produced, in most model assessments, better results that the bucket constraint OF. The difference between them is that the later assigns an additively symmetrical penalty to predictions of the same % of overestimation and underestimation (i.e. the same penalty will be assigned to a 20% over-estimation and to 20% under-estimation of volume). On the other hand, the log-bias constraint is a multiplicatively symmetrical penalty (i.e. the same penalty will be applied to a prediction that is twice or half of the observation volume). The following figure shows that the bucket constraint, as used by Chiew et al. 2009 is much more severe than the log-bias. *(Viney et al. 2009).*

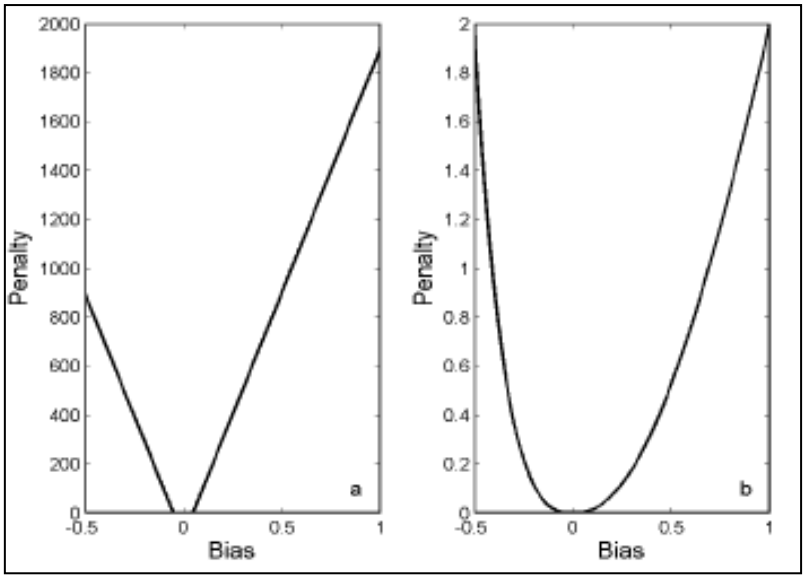


Figure 14. *Penalties induced by the bucket constraint (a) and the log-bias constraint (b), (Viney et al. 2009).*

For example, if the Bias is -0.5 the penalty in case (a) is 850 while in case (b) only 2. This makes the log-bias OF a very challenging measure for this study's model assessment too.

- **The "BL" Objective Function:**

This is another multi-variable OF proposed by *Bergstrom and Lindstrom, (2002)* that combines two classic OF, the **NSE** and the **relative Bias** with the following formula:

$BL_{,V} = NSE - w*abs(V_E)$

NSE is the Nat-Sutcliffe efficiency, $\mathbf{V_E = \sum (X - Q) / \sum Q)}$ is the relative volume error and w is usually 0.1

This OF produces an optimal R^2 and practically no volume error, *(Lindstrom, 1997).*

- **The "NSE^3" Objective Function:**

This is a transformation to the NSE where the absolute residuals are raised to cubic power.

$$NSE^3 = 1 - \frac{\sum_{i=1}^{n}\left(Qobs_i - Qsim_i\right)^3}{\sum_{i=1}^{n}\left(Qobs_i - \overline{Qobs}\right)^3}$$

CHAPTER 5

5.1 Validation

Validation or verification of the model is the next step in the model performance assessment. The models were tested for the specified validation periods with the same parameter estimates obtained from the calibration periods. This assumes stationarity of the state variables over the entire time-series. Calibration and validation periods are of approximately the same size and include both wet and dry periods. Assessing the model performance outside the calibration period, as it is the case in this study and, over periods with different climatic patterns, may indicate how robust or not the models are.

5.2 Evaluation Metrics:

Several *"goodness of fit"* measures and *fit statistics* were calculated for the validation periods (same as in calibration period). These criteria determine the amount of uncertainty that remains in the models after they have been calibrated. These include the following:

- **Relative Bias**: (rel.bias) :is the Bias as a fraction of the total observed flow,
  (+/- values indicate over/under-estimation).

$$\sum (Q_{sim} - Q_{obs}) / \sum Q_{obs}$$

- **NSE**: (r.squared): Nash-Sutcliffe Efficiency, (more weight on peak flows)

$$NSE = 1 - \frac{\sum_{i=1}^{n} \left(Qobs_i - Qsim_i \right)^2}{\sum_{i=1}^{n} \left(Qobs_i - \overline{Qobs} \right)^2}$$

- **NSE_{SQRT}**: (r.sq.sqrt): Nash-Sutcliffe Efficiency using square-root transformed data (less weight on peak flows),

$$NSE_{\sqrt{Q}} = 1 - \frac{\sum_{i=1}^{n} \left(\sqrt{Qobs_i} - \sqrt{Qsim_i} \right)^2}{\sum_{i=1}^{n} \left(\sqrt{Qobs_i} - \sqrt{\overline{Qobs}} \right)^2}$$

Another important statistic that has been calculated is the:

- **AMAFE**: *Average* percent error of the *Maximum Annual Flows*,

  where:

  $MaxQsim_j$ is the simulated maximum annual flow of year *j*, $MaxQobs_j$ is the observed maximum annual flow of year *j*, and *k* is the number of hydrological years of the simulation period.

$$\%AMAFE = \frac{1}{k} \cdot \sum_{j=1}^{k} \left(\frac{MaxQsim_j - MaxQobs_j}{MaxQobs_j} \times 100 \right)$$

The evaluation of the models has been done in a comparative way of different model structures and different modelling approaches. As mentioned before, the split sample test aims to demonstrate the versatility of the models while, the different evaluation metrics aim to address the different parts of the hydrograph (peak or low flows).

The next chapter contains a summary of the most important results and findings from the calibration and validation of the models. For the complete set of results one should refer to the appropriate Appendices. (APPENDIX A: GR4J model, APPENDIX B: AWBM model, APPENDIX C: IHACRES model – CWI version and APPENDIX D: IHACRES model – CMD version). For a better understanding and interpretation of the results, it would be useful if Appendices and Results chapter are viewed together.

CHAPTER 6

6.1 Procedure

A standard procedure has been followed for the performance assessment of all four models. This includes:

(1) The split of the test period into two, for calibration-validation:

(a) "**Wet period**" *from 01-10-1987 to 30-09-1992*

(b) "**Dry period**" *from 01-10-1992 to 30-09-1997*

(2) The check of model performance based on four different objective functions:

(a) **NSE**

(b) **Viney**

(c) **BL**

(d) **NSE^3**

For every simulation based on the abovementioned criteria, the following aspects have been examined:

- Performance of Optimization Algorithm
- Model performance and uncertainty caused by the model structure
- Parameter Values, Ranges and Stability
- Objective Function evaluation based on fit statistics

Performance of the algorithms was assessed based on the amount of iterations that they perform (evaluations) until they converge. Plots of algorithm optimization traces illustrate the effort required to optimize a given objective function.

Parameter Stability is shown in tables containing their calibrated values for every optimization method.

Finally, several objective functions addressing the different parts of the hydrograph are used to identify parameter and model uncertainty.

The following results, figures and findings refer to the most representative cases and effort has been made to include the different behavior of all models tested under different optimization functions and objective functions. The full set of results can be viewed in the appendices.

6.2 Model performance and model structure uncertainty:

In general, the modeled streamflows (in log-scale) produced by the different model structures, appear to be closely related to the observed runoff in most of the cases. The figures below demonstrate this finding by comparing some of the models that were calibrated by different objective functions over the two different periods (wet / dry).

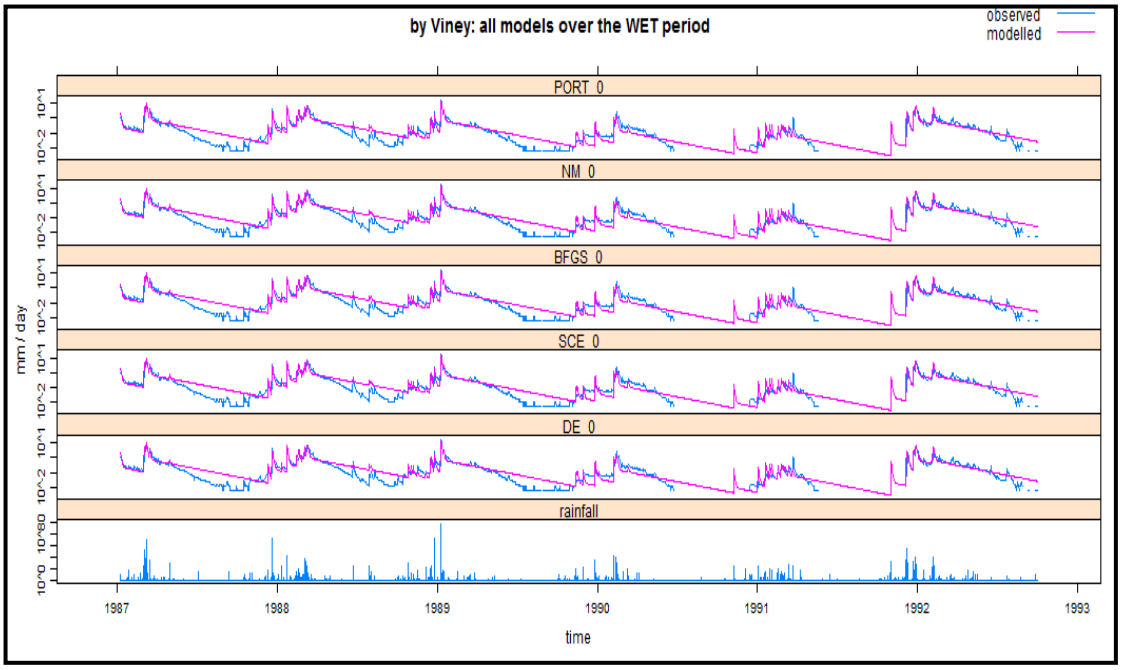


Figure 15. AWBM calibration results for the wet period (objective function used: Viney).

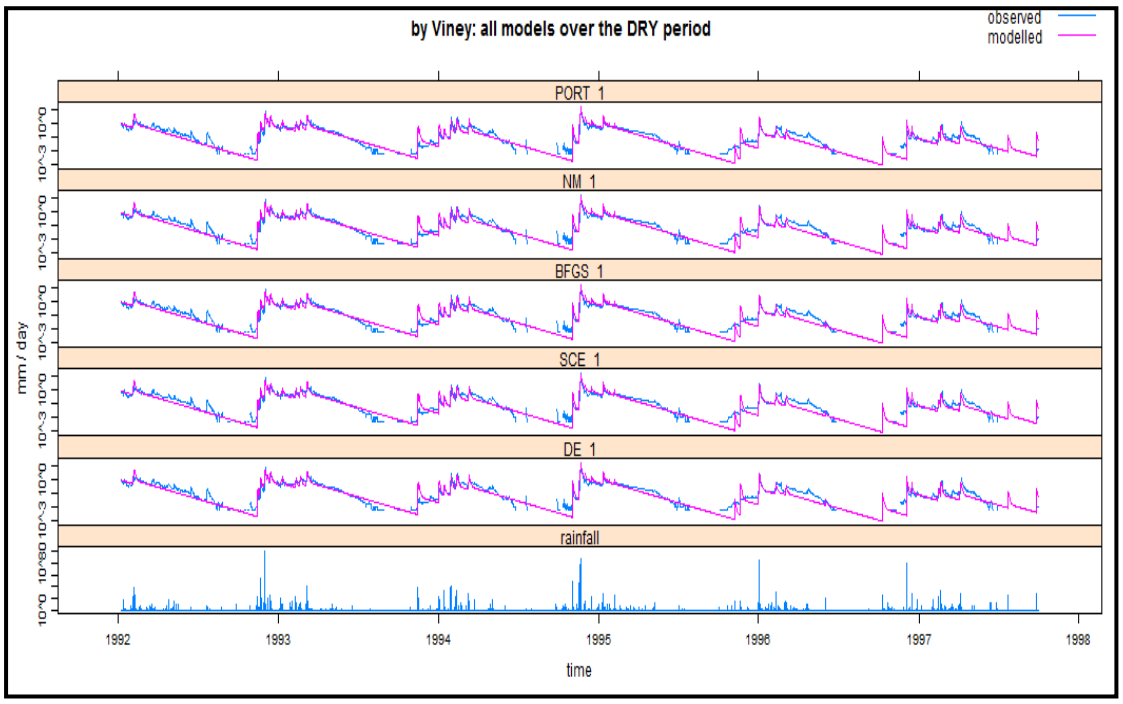


Figure 16. AWBM calibration results for the dry period (objective function used: Viney).

Similar patterns for the other models can be seen in the appendices. In general, the relative bias produced during the wet period is always smaller than the one during the dry period (see fit statistics in appendices). This indicates that models perform better during wet periods. To assess the quality of the fitted models, the Normal Probability Quantile-Quantile plots have been produced. As the pictures below indicate, the fitted models appear to follow very similar behavior in terms of distribution shape, scale and location.

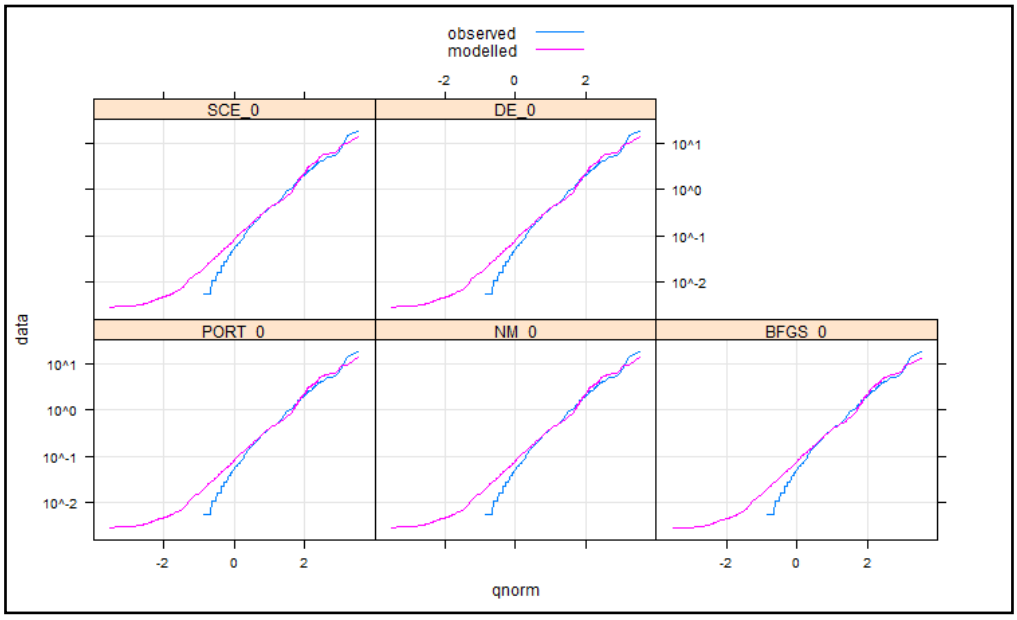


Figure 17. *AWBM Q-Q plot (normal distribution) for the dry period,* objective function used*: Viney*

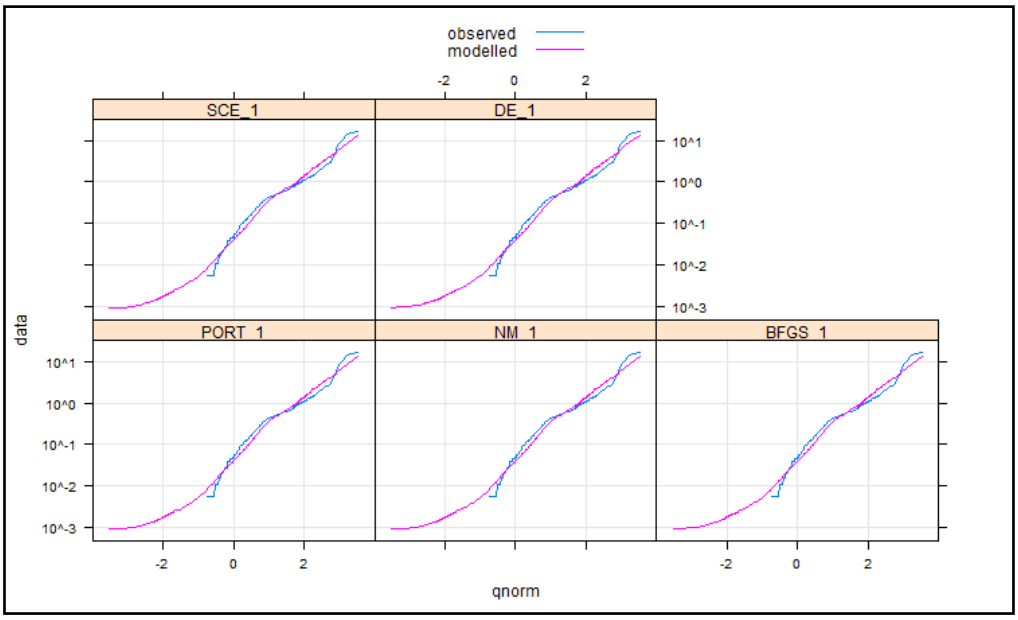


Figure 18. *AWBM Q-Q plot (normal distribution) for the wet period,* objective function used*: Viney*

A visual inspection of the Q-Q plots indicates that there is a small advantage in the streamflow prediction when the model is calibrated over the Wet period. Another important finding is that in all the cases, low flows are more difficult to be "captured" by the models, in contrast to the peak flows that are always predicted close to the observed ones.

Looking at the scatter plots again for the AWBM, calibrated by the Viney objective function, all simulations underestimate the observed streamflow. Best results were obtained for the wet period with the use of global search algorithms, although local search algorithms also yielded similar results as in this case.

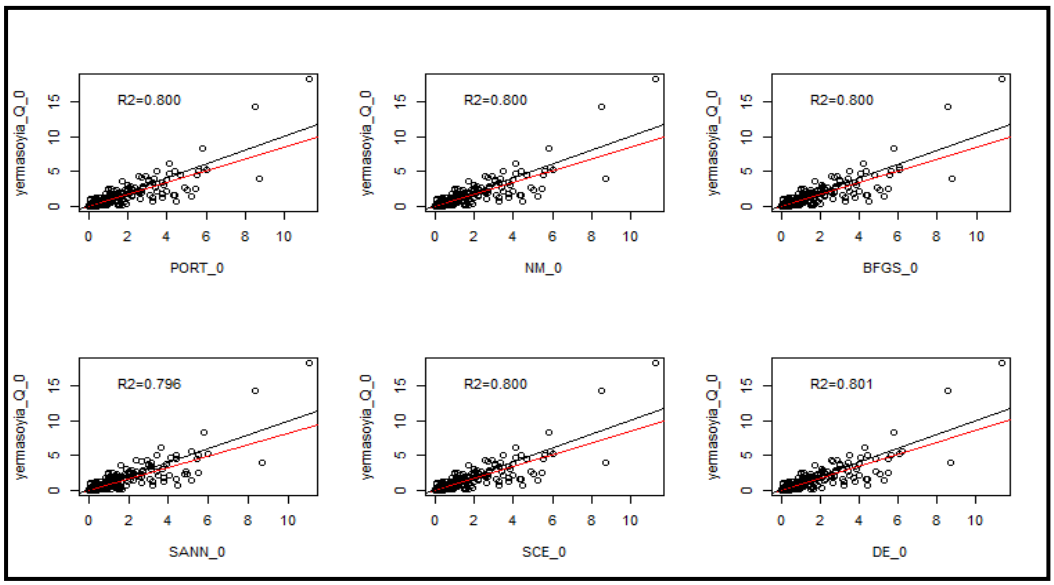


Figure 19. *AWBM scatter plot of simulated streamflows for the dry period, objective function used: Viney*

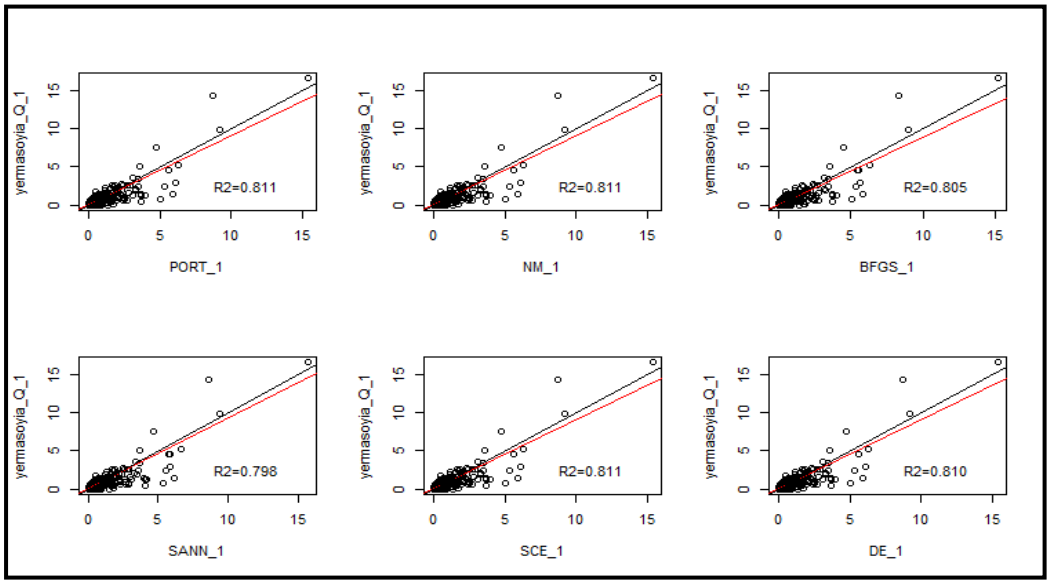


Figure 20. *AWBM scatter plot of simulated streamflows for the wet period, objective function used: Viney*

To assess the uncertainty caused by the structure of the model, the following graphs were produced. These graphs simply represent the uncertainty bounds (upper – lower) of the higher – lower value of the predicted (by any optimization algorithm) streamflow for each objective function. For example, in the IHACRES – CMD the uncertainty bounds are very narrow for both the wet and the dry periods, indicating that this model predicts flows that are not affected by the different time periods or their characteristics. However, again this model does not predict the low flows so well, especially during the dry period.

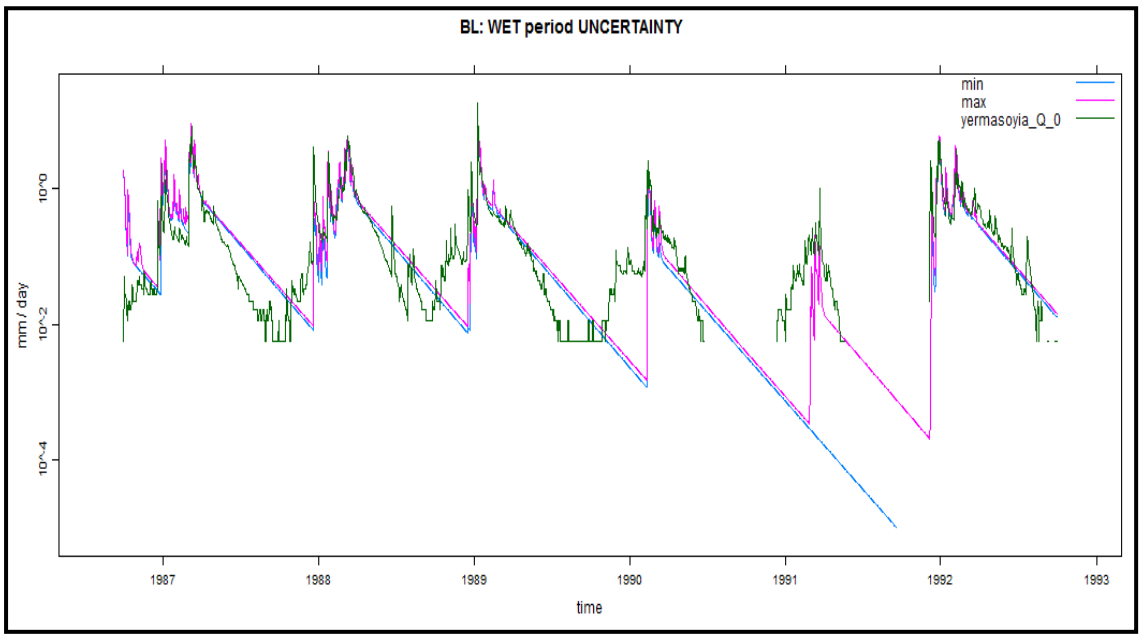


Figure 21. *IHACRES model – CMD version, Wet period uncertainty bounds, objective function used: BL*

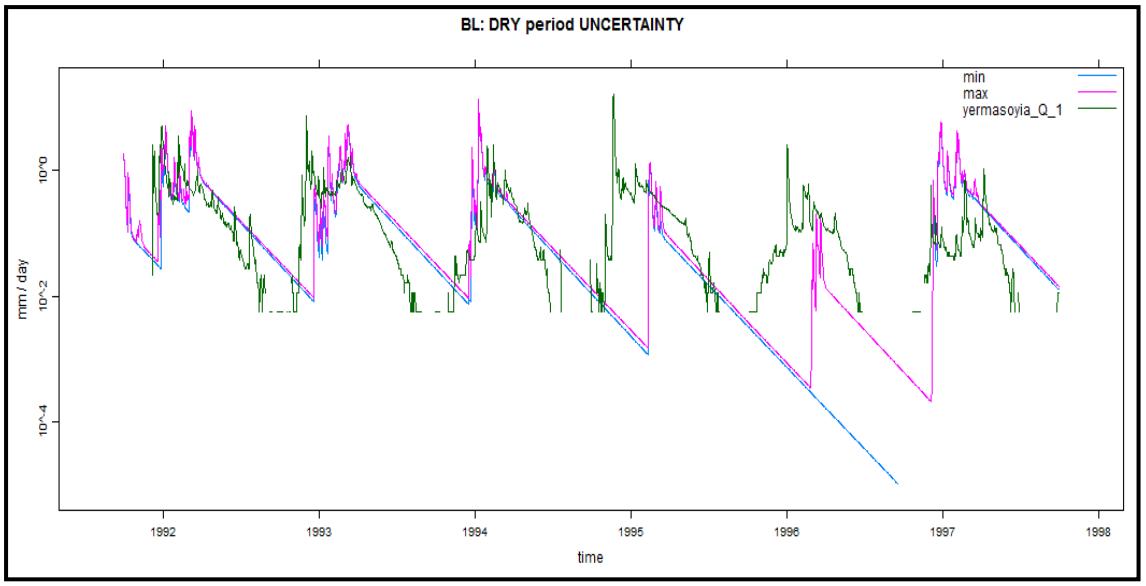


Figure 22. *IHACRES model – CMD version, Dry period uncertainty bounds, objective function used: BL*

Looking at the models of IHCRES and AWBM (see appendices) one can observe a similar behavior. Uncertainty bounds are narrow and models perform better for a wet period rather than a dry period, based on the Objective Functions that have been used in this study. Interestingly, for the GR4J model, uncertainty bounds are larger in most cases, meaning that this model structure is more susceptible to the choice of the optimization method and the period of calibration.

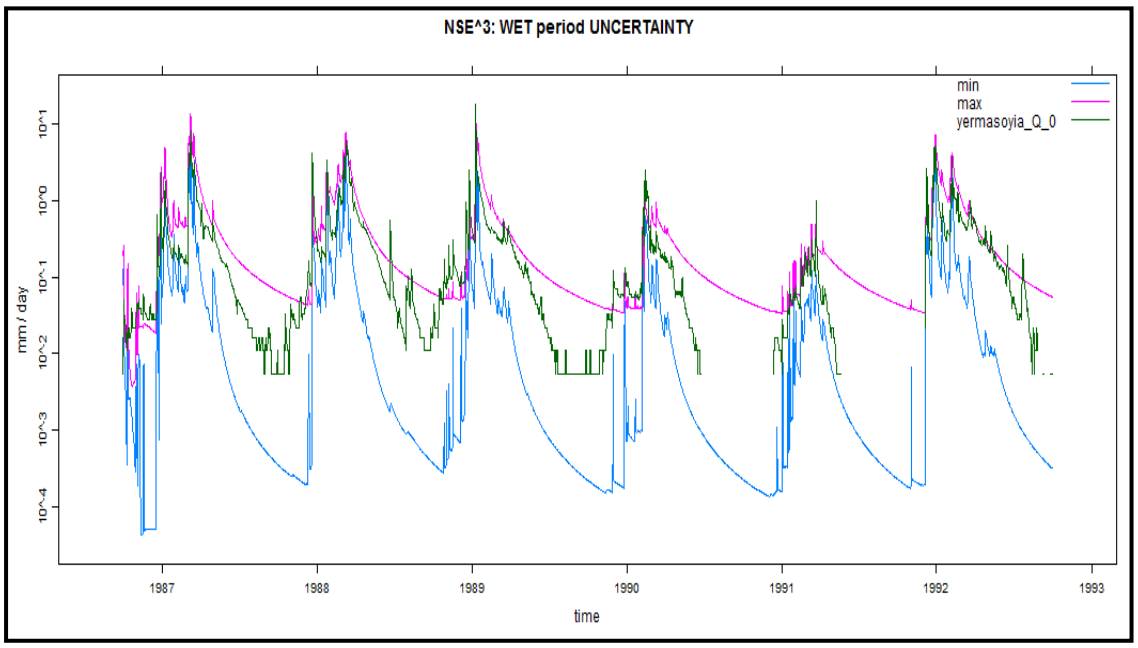


Figure 23. *GR4J model, Wet period uncertainty bounds, objective function used: NSE^3*

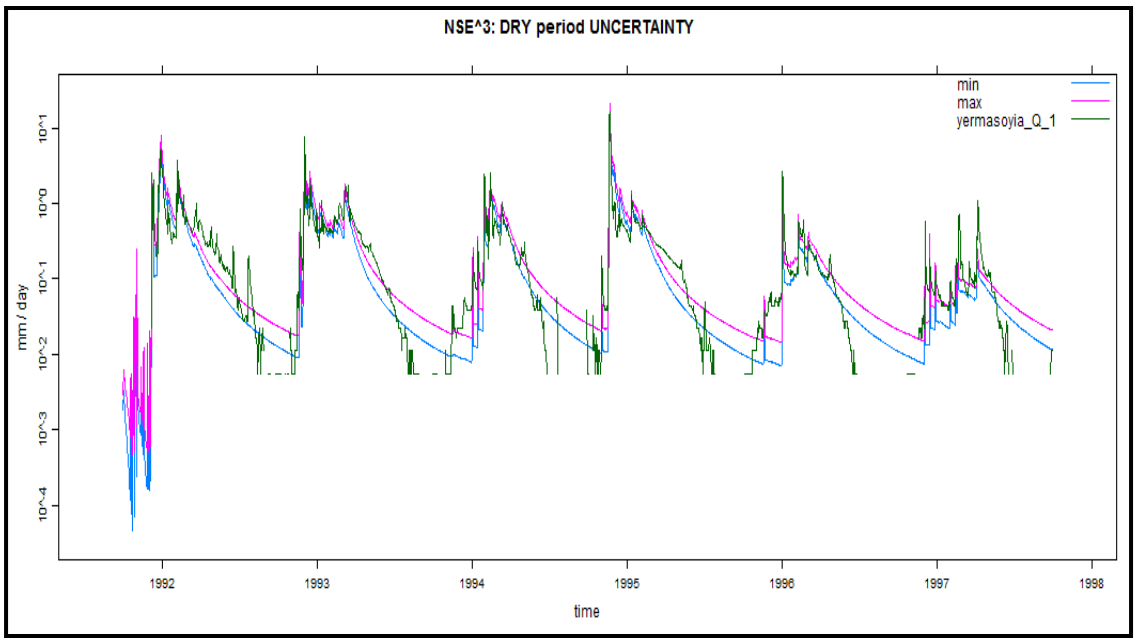


Figure 24. *GR4J model, Dry period uncertainty bounds, objective function used: NSE^3*

This finding illustrates how important the model structure is in terms of uncertainty in the models predictions. But selecting an appropriate objective function can

provide improvements in modelling. For example, again fir the GR4J models, simulations using the Viney (or the NSE) as objective functions reduce significantly the amount of uncertainty in the predictions. Again the wet period yields better results.

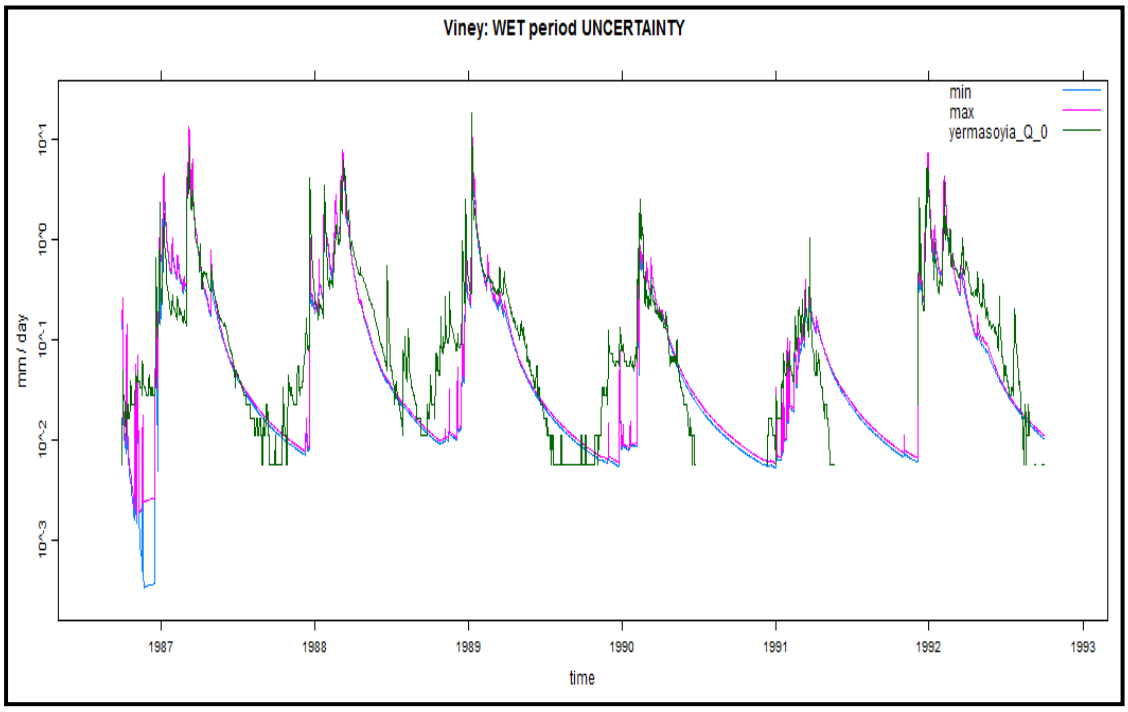


Figure 25. *GR4J model, Wet period uncertainty bounds, objective function used: Viney*

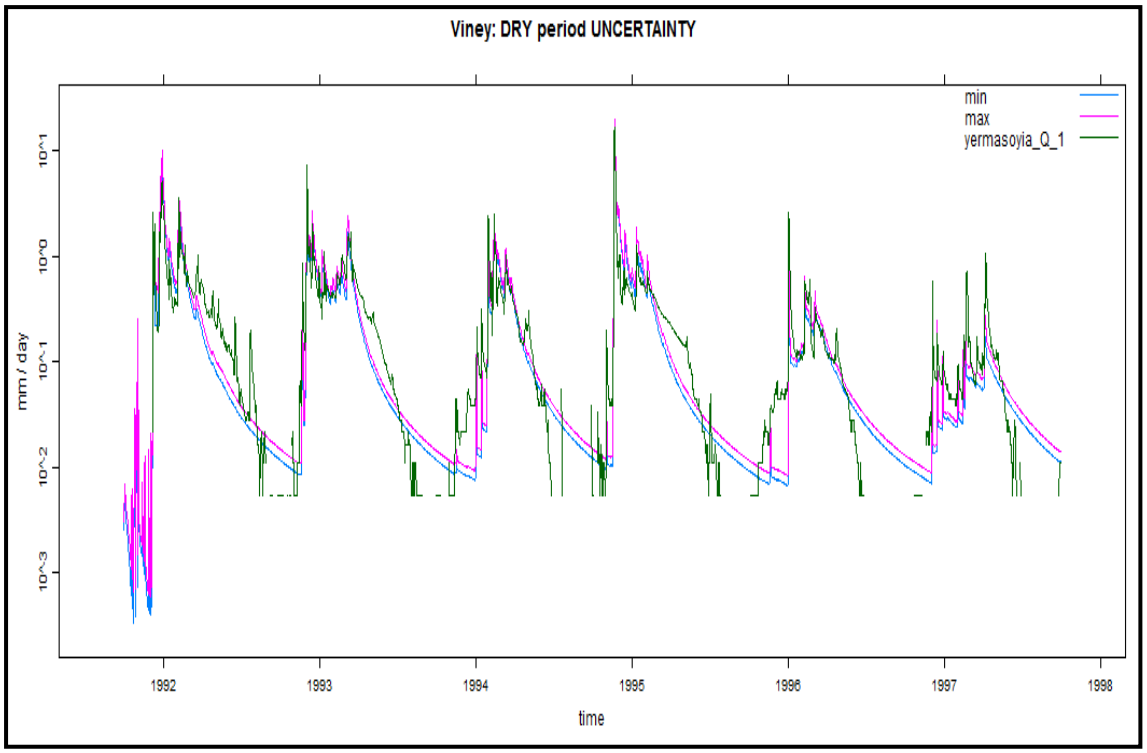


Figure 26. *GR4J model, Dry period uncertainty bounds, objective function used: Viney*

6.3 Assessment of Algorithm choice:

The optimization traces during the **WET period** suggest that all algorithms (except SANN in all cases and DREAM when calibration is done by Viney), perform satisfactorily. Global search algorithms SCE and DE appear to be the most insensitive to the choice of the objective function used and always outperform, even slightly in some cases, the local search algorithms PORT, NM and BFGS. (see also Calibration statistics – Appendix A).

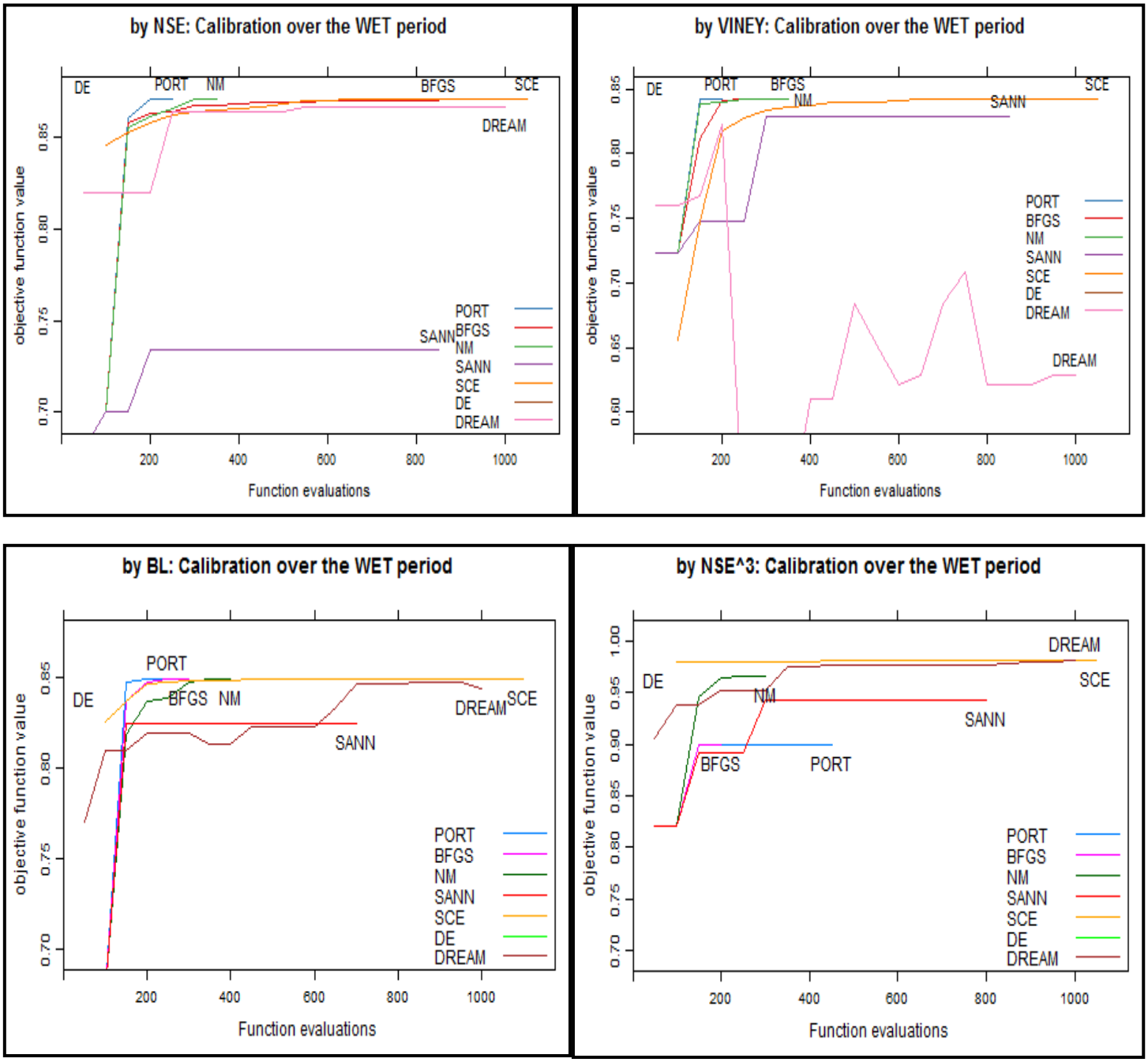


Figure 27. GR4J model. Optimization traces for the wet period

The term "insensitive" to the choice of the Objective Function may also translate that the algorithm is robust for operational use for both low and peak flow simulations. A very similar pattern appears also when calibration is performed during the **DRY period** (Appendix A). However, in absolute values of Objective function, during calibration, this model does not function so efficiently during the DRY period. For example, comparing the NSE of SCE algorithm achieved in the wet period with that obtained in the dry period is about 87% and 79% respectively. Moreover, when the Validation period is dry, SCE achieves 78% in NSE and when the validation is a wet period, the same algorithm achieves 84.3%. The use of the other algorithms also result in similar fluctuations in the values of the Objective functions between wet and dry periods, always having higher values in the wet.

***Table 5**. GR4J model Fit Statistics in Calibration: WET period*

	rel.bias	r.sq.sqrt	NSE	viney	BL	NSE^3
PORT	-0.270	0.794	0.871	0.594	0.844	0.981
BFGS	-0.269	0.797	0.871	0.596	0.844	0.980
NM	-0.268	0.795	0.871	0.598	0.845	0.981
SANN	-0.568	0.662	0.734	-2.498	0.677	0.921
SCE	-0.271	0.793	0.871	0.590	0.844	0.981
DE	-0.257	0.796	0.870	0.630	0.844	0.980
DREAM	-0.207	0.817	0.867	0.738	0.846	0.977

***Table 6**. GR4J model Fit Statistics in Calibration: DRY period*

	rel.bias	r.sq.sqrt	NSE	viney	BL	NSE^3
PORT	-0.225	0.813	0.797	0.632	0.775	0.899
BFGS	0.077	0.787	0.731	0.724	0.723	0.873
NM	0.107	0.794	0.761	0.744	0.750	0.909
SANN	-0.408	0.735	0.728	-0.269	0.687	0.819
SCE	-0.140	0.822	0.791	0.746	0.777	0.897
DE	-0.176	0.819	0.785	0.701	0.767	0.886
DREAM	-0.194	0.818	0.795	0.687	0.776	0.899

***Table 7**. GR4J model Fit Statistics in Validation: DRY period*

	rel.bias	r.sq.sqrt	NSE	viney	BL	NSE^3
PORT	-0.303	0.791	0.777	0.386	0.747	0.873
BFGS	-0.309	0.788	0.768	0.353	0.737	0.853
NM	-0.301	0.792	0.778	0.393	0.748	0.874
SANN	-0.622	0.568	0.542	-4.114	0.480	0.683
SCE	-0.303	0.791	0.778	0.385	0.748	0.875
DE	-0.281	0.796	0.780	0.468	0.752	0.887
DREAM	-0.248	0.801	0.754	0.537	0.729	0.833

***Table 8**. GR4J model Fit Statistics in Validation: WET period*

	rel.bias	r.sq.sqrt	NSE	viney	BL	NSE^3
PORT	-0.206	0.814	0.862	0.734	0.842	0.977
BFGS	0.079	0.817	0.791	0.787	0.790	0.946
SANN	-0.367	0.756	0.849	0.138	0.813	0.973
SCE	-0.127	0.828	0.843	0.809	0.830	0.966
DE	-0.159	0.826	0.845	0.782	0.829	0.965
DREAM	-0.177	0.820	0.859	0.775	0.841	0.975

6.4 Model Parameters Stability:

Looking at the model parameter stability below, for the GR4J model, one can assume that x4 (time base of the UH) is almost invariable (approx. 1.1 days) and insensitive to the choice of the algorithm and the choice of the Objective function. The other parameters i.e. x2, (groundwater exchange coeff.), x1 (max. capacity of production store) and x3 (1-day ahead routing capacity) are sensitive both to the optimization method that is used and the objective function, especially in the wet period.

***Table 9**. GR4J model parameters for the WET period (Calibration)*

Objective Function used = **NSE**

	x2	x3	x4	x1	etmult
PORT	-14.770	116.100	1.130	274.000	1
BFGS	-11.430	97.100	1.150	307.000	1
NM	-14.880	117.200	1.130	273.000	1
SANN	-2.570	20.100	1.320	642.000	1
SCE	-15.170	118.200	1.130	271.000	1
DE	-18.070	136.800	1.140	243.000	1
DREAM	-7.820	86.000	1.150	339.000	1

Objective Function used = **Viney**

	x2	x3	x4	x1	etmult
PORT	-5.000	92.300	1.110	357.000	1
BFGS	-5.000	93.500	1.110	356.000	1
NM	-5.000	91.500	1.110	357.000	1
SANN	-2.470	62.000	1.190	413.000	1
SCE	-5.000	92.300	1.110	357.000	1
DE	-3.710	75.800	1.130	383.000	1
DREAM	-4.920	59.000	1.180	406.000	1

***Table 10**. GR4J model parameters for the DRY period (Calibration)*

Objective Function used = **NSE**

	x2	x3	x4	x1	etmult
PORT	-15.000	158.635	1.056	252.297	1
BFGS	-14.998	142.246	1.074	262.663	1
NM	-14.999	165.698	1.049	248.320	1
SANN	-13.712	137.058	1.030	282.453	1
SCE	-14.999	158.633	1.056	252.439	1
DE	-13.529	154.495	1.041	262.426	1
DREAM	-9.724	118.458	1.026	301.581	1

Objective Function used = **Viney**

	x2	x3	x4	x1	etmult
PORT	-15.000	199.000	1.020	231.000	1
BFGS	-15.000	195.000	1.030	231.000	1
NM	-15.000	200.000	1.010	232.000	1
SANN	-7.270	140.000	1.080	274.000	1
SCE	-15.000	198.000	1.030	230.000	1
DE	-13.000	173.000	1.070	236.000	1
DREAM	-9.720	118.000	1.030	302.000	1

An example of relatively more stable model parameters is demonstrated with the IHACRES model – CWI version, in the table below.

***Table 11**. IHACRES model – CWI version parameters for the DRY period (Calibration)*

Objective Function used = **NSE**

	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	delay
PORT	0.015	8.000	0.001	0	1	20	13.424	0.689	0.619	0.381	1
BFGS	0.121	6.133	0.001	0	1	20	13.064	0.684	0.622	0.378	1
NM	0.870	4.298	0.001	0	1	20	13.059	0.677	0.624	0.376	1
SANN	0.318	4.988	0.001	0	1	20	14.113	0.681	0.619	0.381	1
SCE	0.015	8.000	0.001	0	1	20	13.425	0.689	0.619	0.381	1
DE	1.067	4.156	0.001	0	1	20	12.914	0.677	0.625	0.375	1

Objective Function used = **Viney**

	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	delay
PORT	0.014	8.000	0.001	0	1	20	13.718	0.689	0.618	0.382	1
BFGS	1.841	3.357	0.001	0	1	20	14.148	0.677	0.621	0.379	1
NM	0.013	8.000	0.001	0	1	20	13.951	0.689	0.617	0.383	1
SANN	0.358	5.309	0.001	0	1	20	12.554	0.682	0.626	0.374	1
SCE	0.014	8.000	0.001	0	1	20	13.717	0.689	0.618	0.382	1
DE	2.084	3.343	0.001	0	1	20	13.630	0.675	0.623	0.377	1

6.5 Effect of the objective function:

Some models appear to perform in a similar way regardless the objective function that has been used, and also, the optimization algorithms perform almost in the same way. An example is given below, with the IHACRES – CWI model calibrated in a dry period by NSE and by Viney as objective functions.

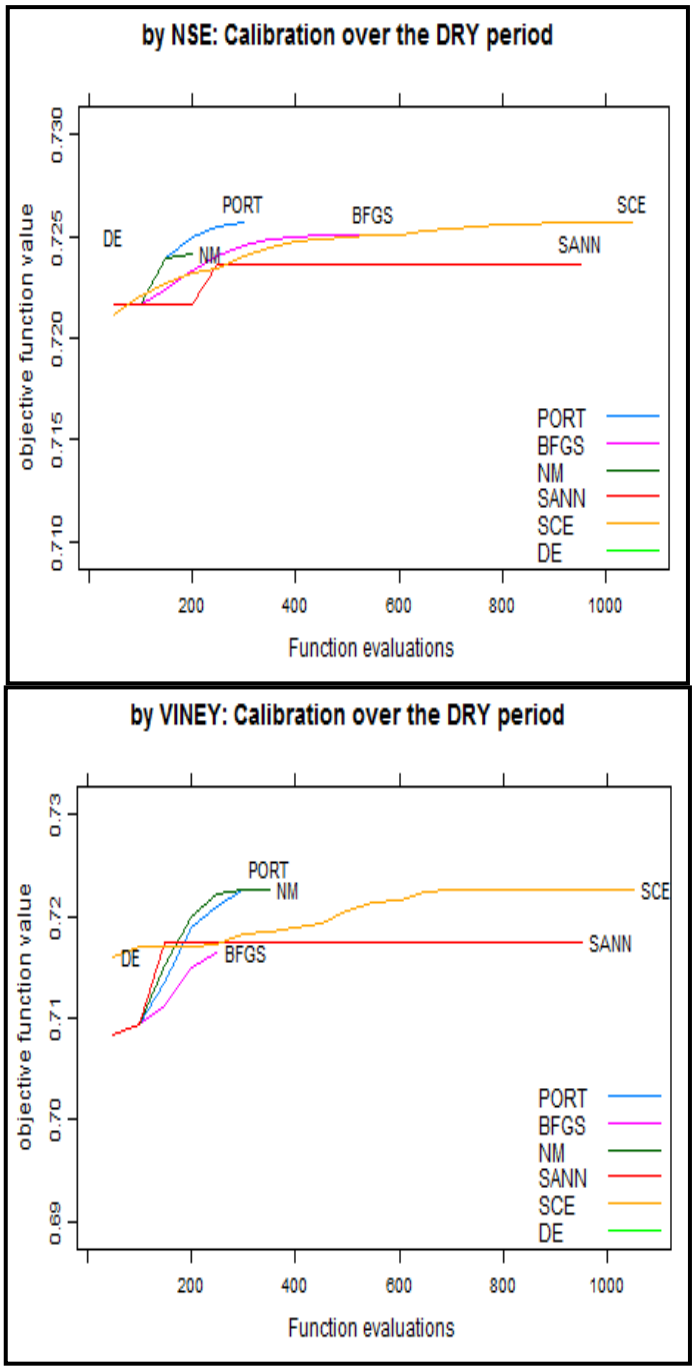


Figure 28. *IHACRES model calibrated by Viney & NSE (DRY period)*

On the other hand, the GR4J model performs better when calibrated by Viney than by NSE, for the same dry period. This is an indication that probably such models are

not so reliable and versatile to cover hydrological variability and uncertainty. This is demonstrated in Figure 16 in the next page.

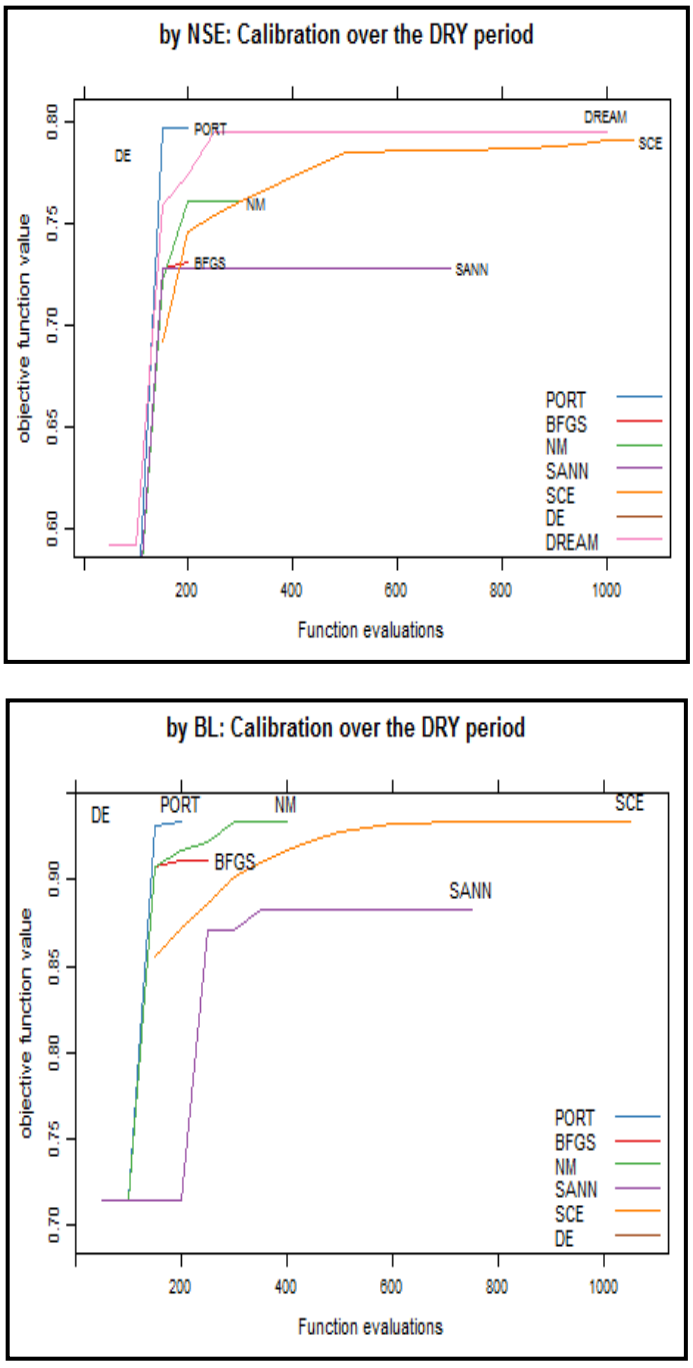


Figure 29. *GR4J model calibrated by Viney & NSE (DRY period)*

6.6 Peak flow assessment:

The AMAFE criterion is an indication of how well a model describes the peak flows. The following table demonstrates the best AMAFE results that have been obtained for every model, based on any objective function and any optimization algorithm.

***Table 12**. Lowest AMAFE values obtained for all models*

Model	Objective Function	Period	Optimization algorithm					
			PORT	NM	BFGS	SANN	SCE	DE
GR4J	NSE	calibration 87-92	-0.111	-0.111	-0.113	0.258	-0.111	-0.067
		validation 92-97	0.040	0.051	0.057	0.470	0.039	0.074
AWBM	NSE^3	calibration 92-97	0.461	0.461	0.462	0.513	0.461	0.470
		validation 87-92	-0.003	-0.003	-0.003	0.032	-0.003	0.002
IHACRES /CWI	Viney	calibration 92-97	-0.091	-0.091	-0.096	-0.100	-0.091	-0.097
		validation 87-92	-0.051	-0.048	-0.052	-0.074	-0.051	-0.058
IHACRES /CMD	NSE^3	calibration 87-92	-0.118	-0.117	-0.005	-0.002	-0.117	-0.064
		validation 92-97	-0.088	-0.087	0.181	0.266	-0.089	-0.008

For the GR4J and the IHACRES / CMD models, the AMAFE criterion gets its optimum value when the models are calibrated during a wet period and then validated for the dry period. In both cases, the global search algorithms have produced the best fits of simulated peak flows that are closer to the peak observed flows. Also, the objective functions that yielded these results are the NSE and the NSE^3, as one would probably expect. On the other hand, for the AWBM and the IHACRES / CWI models, the best results are obtained when the models are calibrated for the dry period and then validated for a wet. For the AWBM the AMAFE criterion is unacceptable for the wet period of calibration (46%).

CHAPTER 7

7.1 Conclusions:

For conceptual R-R models, the most widely used approaches to estimate model parameters are based on the calibration of the models during a specific period. The models are then verified for another period with the use of fit statistics. The following figure illustrates graphically the main steps and methods that were followed for the calibration of four different conceptual lumped R-R models.

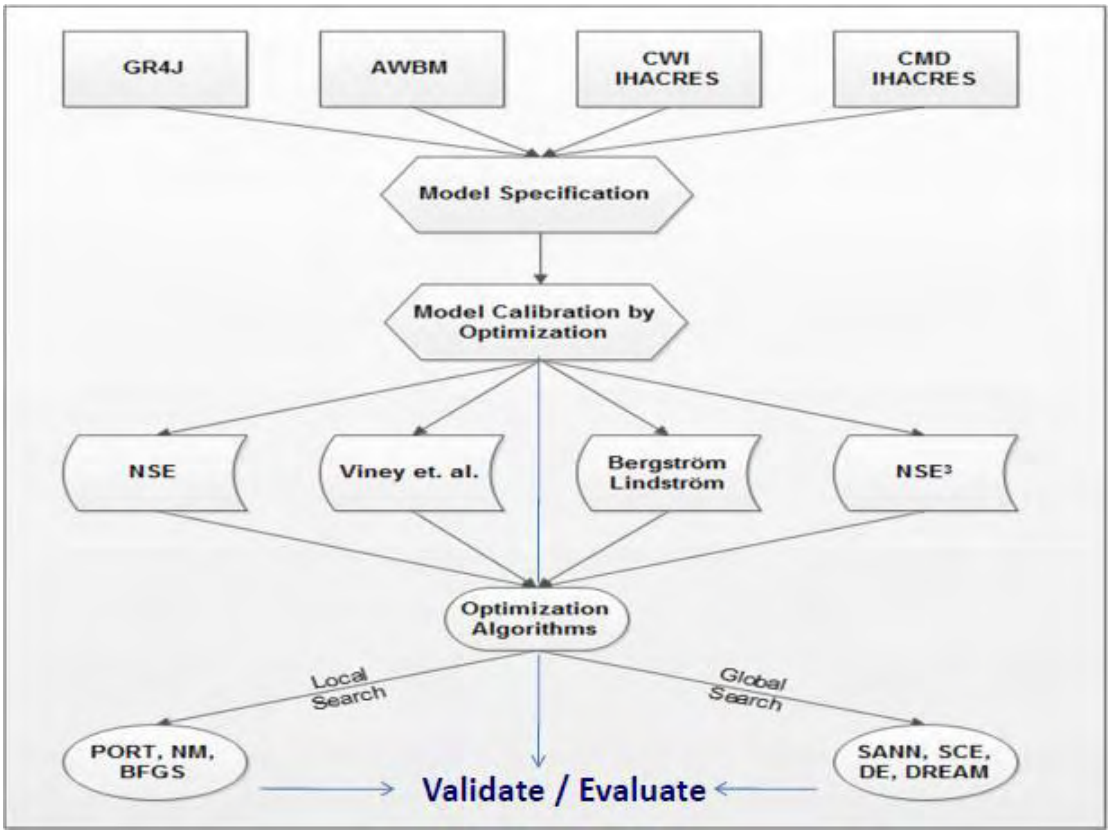


Figure 30. *Project Outline and main calibration strategies*

Any calibration strategy aims to create models that best reflect the physical processes that take place in catchments. This implies that the model parameters must be such so that they ensure a model behavior that is realistic. At the same time, the number and the physical meaning behind the parameters must guarantee optimal model performance. Reduced complexity in the models is indeed an advantage while too many parameters would lead to overparametrization problems and ultimately reduced model performance.

Therefore, the choice of the model structure plays a dominant role and, as it has been shown, it is the fundamental aspect in the elimination of model uncertainty. As

demonstrated in the Results section, different models, although similar in the concept and the approach to explain the physical processes, perform differently under the same conditions (i.e. optimization techniques and choice of objective function).

Different efficiency criteria have been utilized to assess how well a model simulation fits the available observations *(Beven, 2001)*. Depending on the purpose or the target of the study, different efficiency criteria place emphasis on different hydrologic behaviors. For example, the higher portions of a hydrograph (peak flows) can be better simulated at the expense of the lower portions (low flows), *(Krause et al. 2005).* In this study, the NSE and the NSE^3 place emphasis to the high peaks at the expense of the low flows, while the Viney and the BL objective functions demonstrate a finer model behavior in all the segments of the hydrograph.

The choice of an optimization technique also affects the uncertainty in the model parameters, but in a lower degree than the model structure itself. For example, global search algorithms appear to outperform the local search algorithms in most cases, with the SCE and the DE algorithms being almost in all cases the most effective. Local search algorithms may improve the value of the objective function given that they start at different parameter sets and not from a single sample. This enables the algorithms to avoid being entrapped to a local minimum solution. However, this has a significant cost in the computational resources required to perform the model calibration. In general, local search algorithms also performed in most of the cases, in a very satisfactory way.

Using the split sample test of different calibration / validation periods, in this case a wet / dry period split, it was found that hydrologic variability affects significantly the model performance. Different calibration periods produced different parameters estimates. The more unstable and disperse the parameter values are, the worse the model performs. For a model to be considered reliable and robust, it should perform consistently during the different calibration periods. This is termed as parameter transposability in time *(Gharari et al., 2012)* and is viewed as one of the most important elements in R-R modeling. It has been demonstrated that hydrologically different temporal periods (dry / wet) also increase the instability in some models (i.e. the GR4J) and for others, such as the two IHACRES models and the AWBM, this variability did not affect their performance significantly.

A good initial model specification is also crucial if the target is to reduce model uncertainty. Trusting previous research and studies should always be the starting point in model specification. However, as most models provide some flexibility in their structure and parameter ranges, they should be tested more thoroughly in

order to increase confidence and to select a model specification that will ultimately function efficiently. It is also important to examine the scale of parameter ranges that have a valid physical significance.

In many studies including this one, the search for optimum model parameters using efficiency criteria may often lead to equifinality issues. This means that in the search for optimum parameter values there are multiple optimum model predictions of streamflow close to the observed flow. This does not necessarily mean that the model is over-parameterized but it is due to the model itself and its properties, or due to the characteristics of the catchment and the climate, *(Hreiche et al.)*. Given as an example below, the 3 parameters x1, x2, x3 of the GR4J model found by all methods, (the fourth parameter x4 remains constant), were plotted against each other in a 3D scatter plot. The plot gives an indication of the parameter value ranges in the feasible parameter space

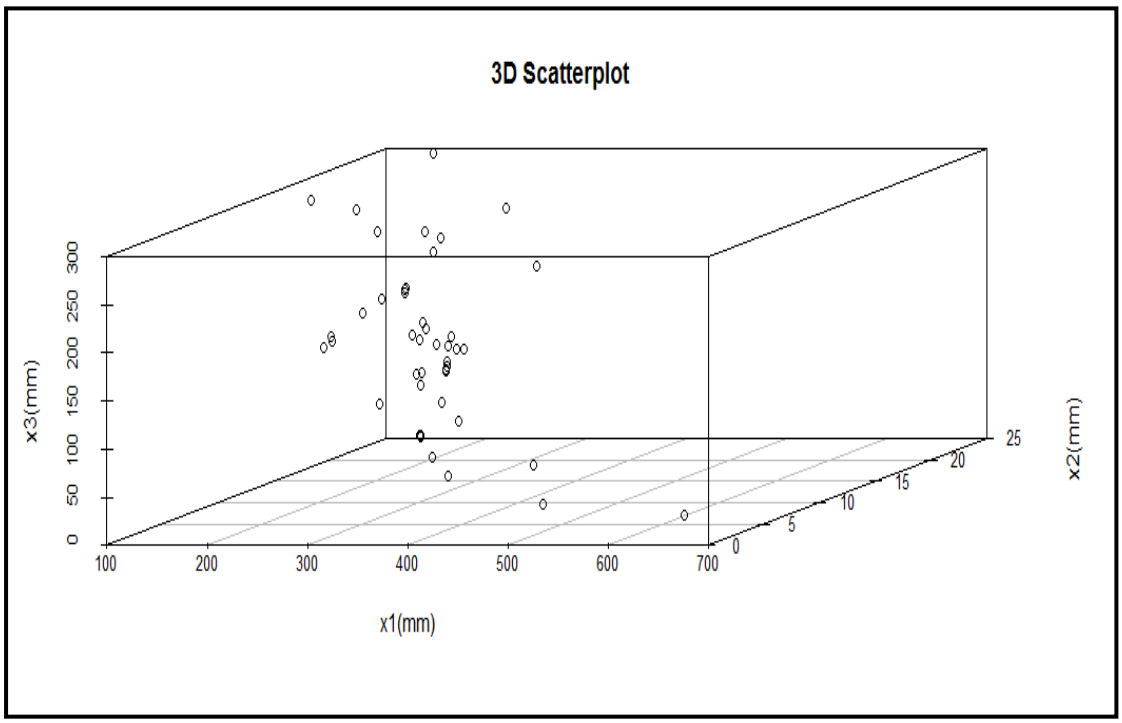


Figure 31. *The issue of equifinality, for the GR4J model parameters.*

7.2 Future work:

In this study, all model simulations have been performed using the observed raw data of streamflow. *Andrews et al. (2011)* suggest that in similar R-R models assessment and identification of model uncertainty, raw data can be transformed to reduce possible data errors. "*This is because streamflow data often tends to be highly skewed and this may result in a large weighting put on large observations"*. For demonstration purposes only, the following figure shows different data transformations and how these relate to the Normal distribution. In this case, a simple log transformation (or a Box-Cox transformation) may result in less inherent uncertainty in the dataset.

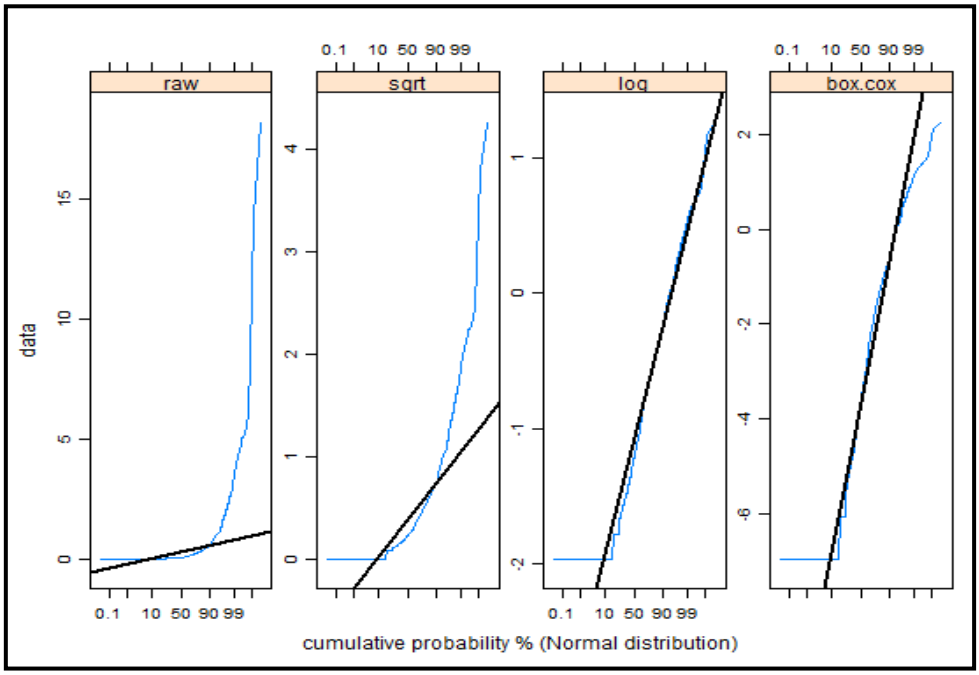


Figure 32. *Different data transformations showing deviations from normality*

The same research study by *Andrews et al., 2011*, suggest also to perform simulations based on rainfall-runoff events only, instead of using the whole raw rata series. This of course has been widely used in the past when extreme value analysis is the key issue, but it may well be used in general. One of the main issues here, is the careful definition of what makes an event. For example, the minimum thresholds of streamflow and rainfall values must be defined, as well as the durations that limit one event. As an example, the next figure illustrates an event-based approach in hydrological modelling. Rainfall events are defined when rainfall exceeds 5mm per day until it remains below 1mm for 4 days. (evp_5). Similarly, the streamflow events (evq_90 or evq_50) are defined when the observed streamflow exceeds the 90 or 50 percentile level for at least 2 time-steps.

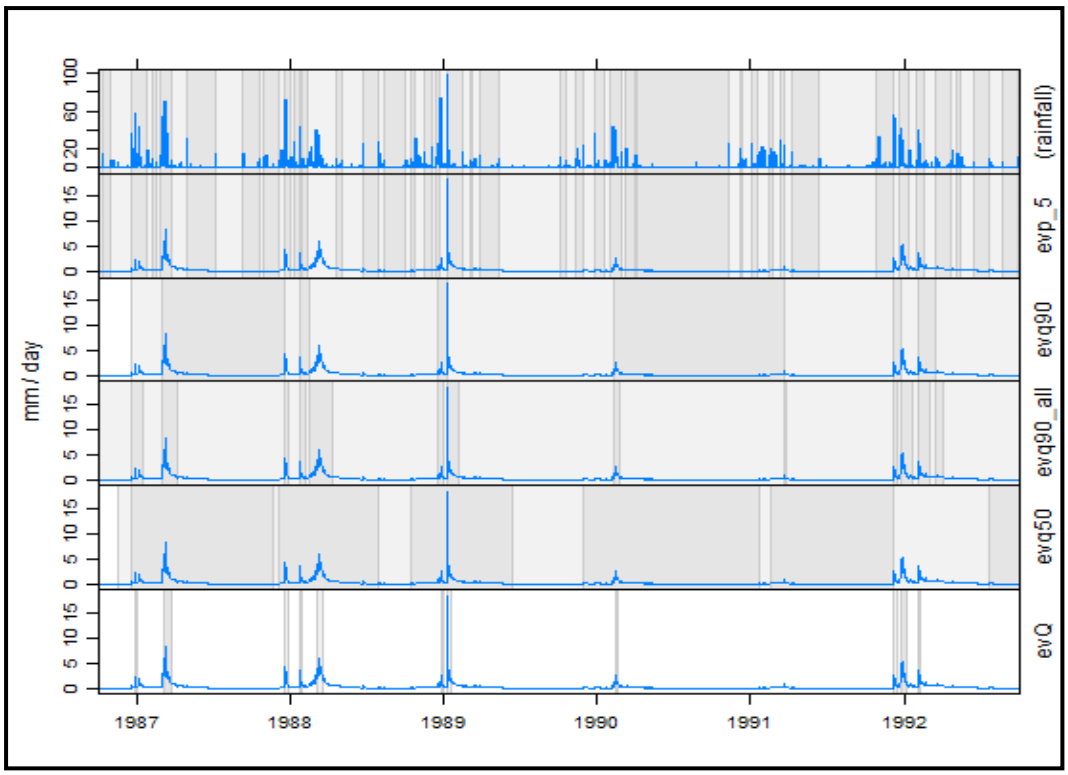


Figure 33. *A typical event-based approach*

Such an approach is likely to reduce the auto-correlation that is inherent in the raw data, as shown in Fig. 30, (aggregation function used for the data: mean).

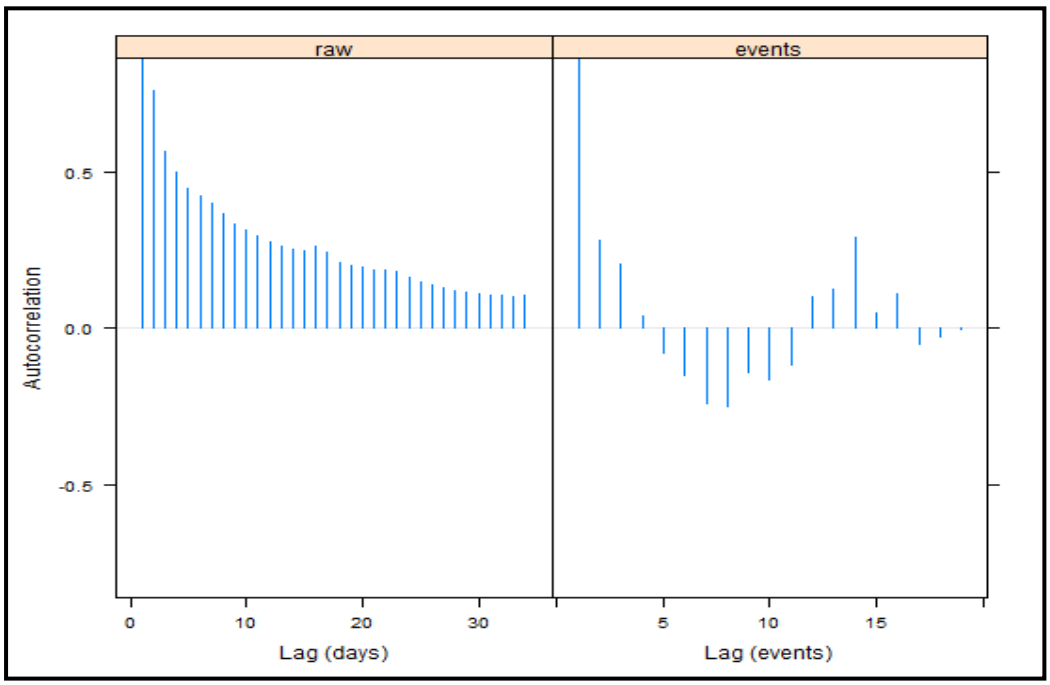


Figure 34. *Event-based / raw data approach and auto-correlation in the time series*

This work can be also used in Frequency Analysis and Extremes in order to estimate, for example, the return levels of peak flows. This work also provides the basis for application to multiple catchments for inter-comparison. Finally, to address the issues of climatic variability, the developed models can be applied to identify the impact of climate change on the hydrological cycle components and in particular streamflow.

Bibliography

Andrews, F.T., Croke, B.F.W., Jakeman, A.J., (2011). An open software environment for hydrological model assessment and development. Environmental Modelling and Software 26, 1171-1185.

Ardia, D., Mullen, K., (2010). The 'DEoptim' Package: Differential Evolution Optimization in 'R'. R package version 2.0-5.

Belisle, C. J. P. (1992). Convergence theorems for a class of simulated annealing algorithms on *Rd*. *J. Applied Probability*, 29, 885–895.

Bergstrom, S., Forsman, A., (1973). Development of a conceptual deterministic rainfall-runoff model. Nordic Hydrol. 4, 147-170.

Bergstrom, S., Lindstrom G. and Pettersson A., (2002). Multi-variable parameter estimation to increase confidence in hydrological modeling. *Hydrol. Process.* 16, 413–421

Beven, J. K.: Rainfall-Runoff Modeling (2001) – The Primer, John Wiley & Sons Ltd., Chichester, 319

Boughton, W., Chiew, F. (2003). Calibrations of the AWBM for use on ungauged catchments. Technical Report 03/15, Cooperative Research center for Catchment Hydrology.

Boughton, W. (2004). The Australian water balance model. Environmental Modeling & Software 19 (10), 943-956

Broyden, C. G. (1970), "The convergence of a class of double-rank minimization algorithms", *Journal of the Institute of Mathematics and Its Applications* 6: 76–90,

Burnash, R.J.C., Ferral, R.L., McGuire, R.A., (1973). A generalized streamflow simulation system – conceptual modeling for digital computers, US Department of Commerce, National Weather Service and State of California, Department of Water Resources.

Chapman, T., (1999). A comparison of algorithms for stream flow recession and baseflow separation. Hydrol. Proces. 13, 701–704.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., Terpenning, I., 1990. STL: A seasonal-trend decomposition procedure based on Loess. Journal of Social Statistics 6, 3{73.

Croke, B.F.W. and A.J. Jakeman (2004), A Catchment Moisture Deficit module for the IHACRES rainfall-runoff model, *Environmental Modeling and Software*, 19(1): 1-5.

Dennis J. E and. Schnabel R. B, Numerical Methods for Unconstrained Optimization and Nonlinear Equations (1983)

Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resources Research 28 (4), 1015{1031.

Duan, Q., Gupta, V.K. and Sorooshian, S., 1993. A shu8led complex evolution approach for effective and efficient global minimization. J. Optimization Theory Appl., 76(3): 501-521.

Duan, Q., Sorooshian, S. and Gupta, V.K., 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. J. Hydrol., 158: 265-284.

Edijatno, Michel, C., 1989. Un mode`le pluie-de´bit journalier a` trois paramètres. La Houille Blanche (2), 113–121.

Edijatno, Nascimento, N.O., Yang, X., Makhloof, Z. et Michel, C. (1999). GR3J : a daily watershed model with three free parameters. Hydrological Sciences Journal, 44(2), 263-278.

Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H. G., (2012): An approach to identify time consistent model parameters: sub period calibration. DELFT University of Technology, the Netherlands. Hydrol. Earth Syst. Sci., 17, 149–161.

Gupta, H. V., Sorooshian, S., and Yapo, P. O., (1998).: Toward improved calibration of hydrologic models: Multiple and non commensurable measures of information, Water Resources, 34, 751–763.

Hreiche, A., Bocquillon, C., Najem, W. Parameter Estimation of Conceptual Rainfall-Runoff model and application to Mediterranean catchments. *Ecole Supérieure d'Ingénieurs de Beyrouth, Lebanon.*

Jakeman, A.J., I.G. Littlewood, and P.G. Whitehead (1990), Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments, *Journal of Hydrology*, 117: 275-300.

Jakeman, A. J., and G. M. Hornberger (1993), How much complexity is warranted in a rainfall-runoff model?, *Water Resources Research*, 29: 2637-2649.

Klaassen, B., Pilgrim, D.H., 1975. Hydrograph recession constants
for New South Wales streams. Civ. Eng. Trans. I.E.Aust. CE17, 43–49.

Krause, P., Boyle, D.P., Base, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences.

Loukas, A., Vasiliadis, L. (2003). Peak Flow Estimation in Ungauged Watersheds, Bologna, Italy

Madsen, H., Wilson, G., Ammentorp H.C. (2001). Comparison of different automated strategies for calibration of rainfall-runoff models. Journal of Hydrology 261, 48-59.

Matthies, H., Strang, G. (1979). ”The solution of non linear finite element equations.” International Journal for Numerical Methods in Engineering 14 (11): 16131626. doi:10.1002/nme.1620141104

Nascimento, N.O., (1995). Appreciation a` l’aide d’un mode`le empirique des effets d’action anthropiques sur la relation pluie-de´bit a` l’e´chelle du bassin versant. PhD Thesis, CERGRENE/ENPC, Paris, France, 550 pp.

Nash, J.E., Sutcliffe, J.V., (1970). River flow forecasting through conceptual models. Part I—A discussion of principles. Journal of Hydrology 27 (3), 282–290.

Nelder, J. A. and Mead, R. (1965) A simplex algorithm for function minimization. *Computer Journal* 7, 308–313.

Nocedal J. and. Wright S. J, Numerical Optimization (2006), chapters 6 and 7

Nocedal, J. (1980). ”Updating Quasi-Newton Matrices with Limited Storage.” Mathematics of Computation 35

Perrin, C., (2000). Vers une amelioration d’un mode`le global pluie-de´bit au travers d’une approche comparative. PhD Thesis, INPG (Grenoble)/Cemagref (Antony), France, 530 pp.

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

Perrin, C., (2002). Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. La Houille Blanche, n°6/7 : 84-91.

Perrin, C., Michel, C., Andreassian, V. (2003). Improvement of a Parsimonious Model for Stream flow simulation. Journal of Hydrology,

Podger, G. (2004). Rainfall-Runoff Library – rrl – User Guide

Price, K., Storn, R. M., Lampinen, J. A., (2005). Differential Evolution: A Practical Approach to Global Optimization (Natural Computing Series), 1st Edition. Springer.

Qingyun, Duan, Soroosh Sorooshian and Vijai Gupta (1992). Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models, *Water Resources Research* 28(4), pp. 1015-1031.

Qingyun Duan, Soroosh Sorooshian and Vijai Gupta (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models, *Journal of Hydrology* 158, pp. 265-284.

Tolson, B.A. and Shoemaker, C.A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. WATER RESOURCES RESEARCH, VOL. 43, W01413

Viney, N.R, Vaze, J., Chiew, F.H.S., Perraud J, Post ., D.A 1 and Teng J. (2009). Comparison of multi-model and multi-donor ensembles for regionalization of runoff generation using five lumped rainfall-runoff models. Proceedings, MODSIM 2009

Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M., Higdon, D., 2009. Accelerating markov chain monte carlo simulation by differential evolution with self-adaptive randomized subspace sampling. International Journal of Nonlinear Sciences and Numerical Simulation 10 (3), 273{290.

Weiler, M., McGlynn, B., McGuire, K., and McDonnell, J., (2003): How does rainfall become runoff? A combined tracer and runoff transfer function approach, Water Resour. Res., 39, 1315.

Wheater, H. S., Jakeman, A. J., and Beven, K. J., (1993): Progress and directions in rainfall-runoff modeling, John Wiley & Sons.

Winsemius, H. C., Savenije, H. H. G., and Bastiaanssen, W. G., (2008): Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins, Hydrol. Earth Syst. Sci., 12, 1403–1413.

Yapo, P.O., Gupta, H.V., Sorooshian, S. (1996). Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. Journal of Hydrology 181, 23-48

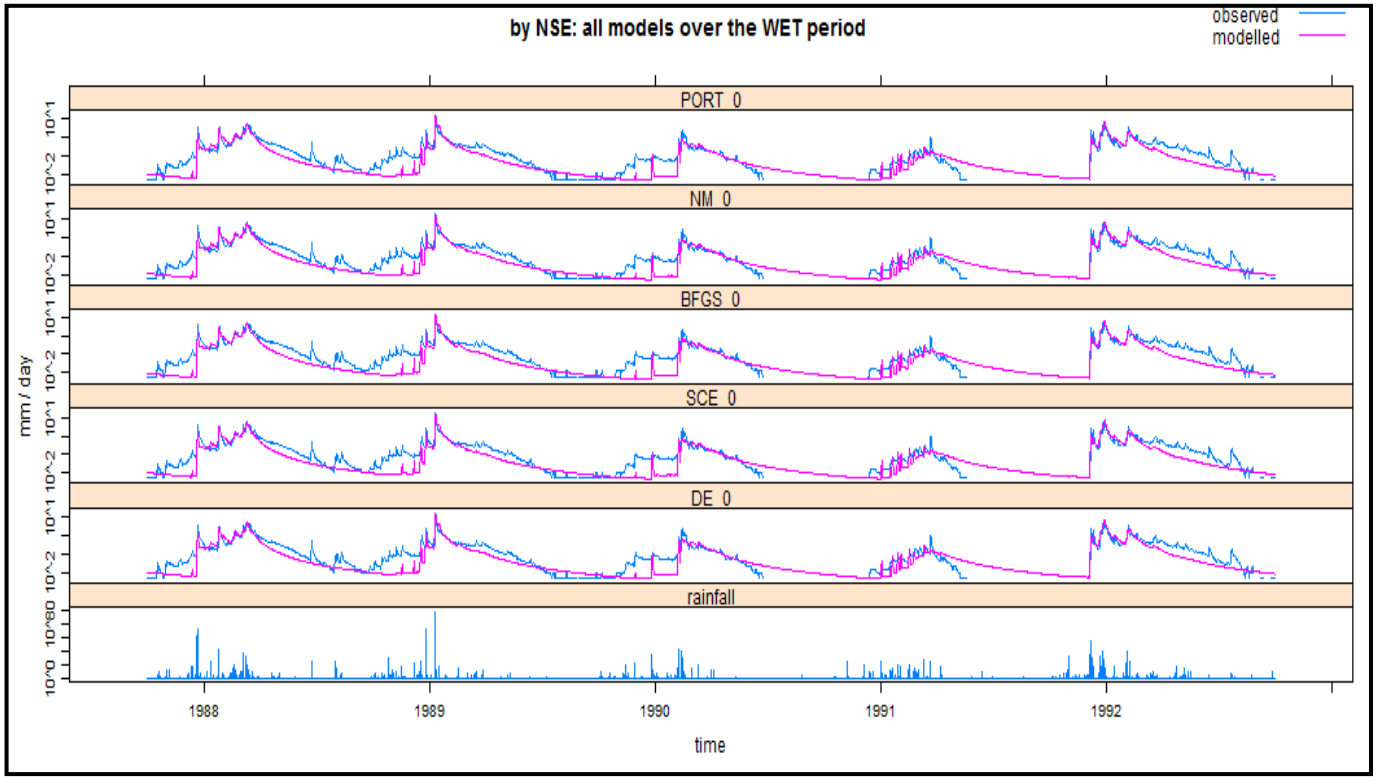
Ye, W., B.C. Bates, N.R. Viney, M. Sivapalan and A.J. Jakeman (1997), Performance of conceptual rainfall-runoff models in low-yielding ephemeral catchments, *Water Resources Research*, 33: 153-16.

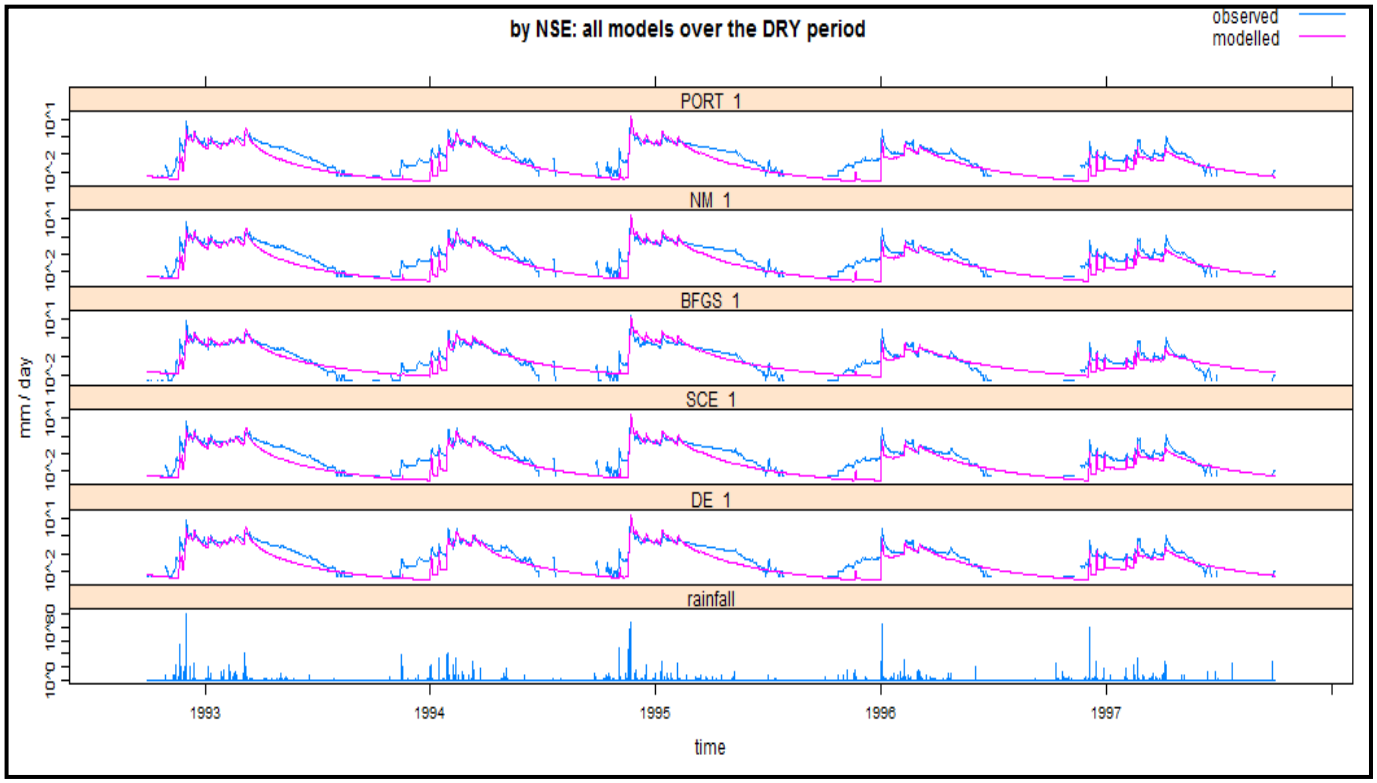
Young P.C. and Jakeman A.J. Refined instrumental variable methods of time-series analysis: Part iii, extensions. International Journal of Control, 31:741–764, 1980.

Young, P. C. (2008). The refined instrumental variable method. Journal Européen des Systèmes Automatisés 42 (2-3), 149-179

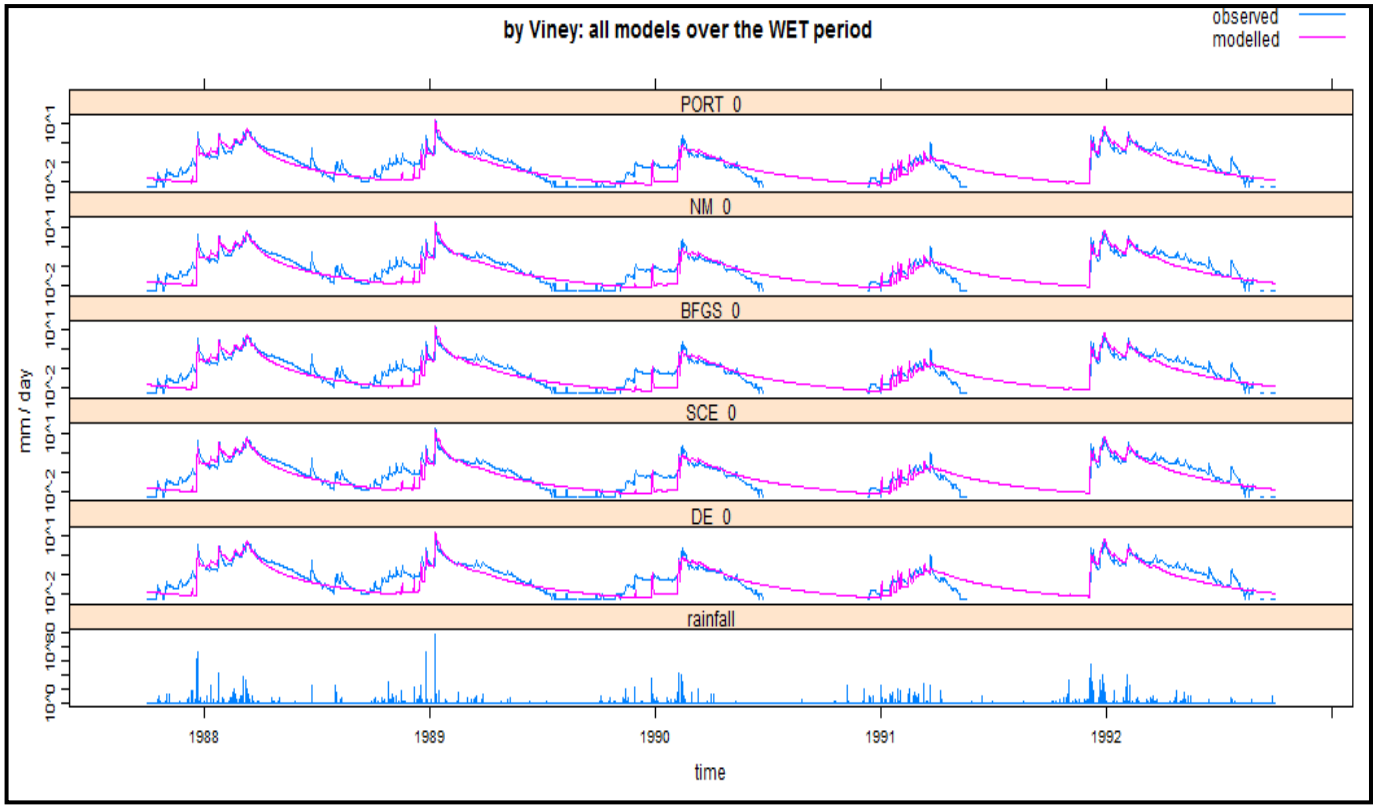
GR4J model

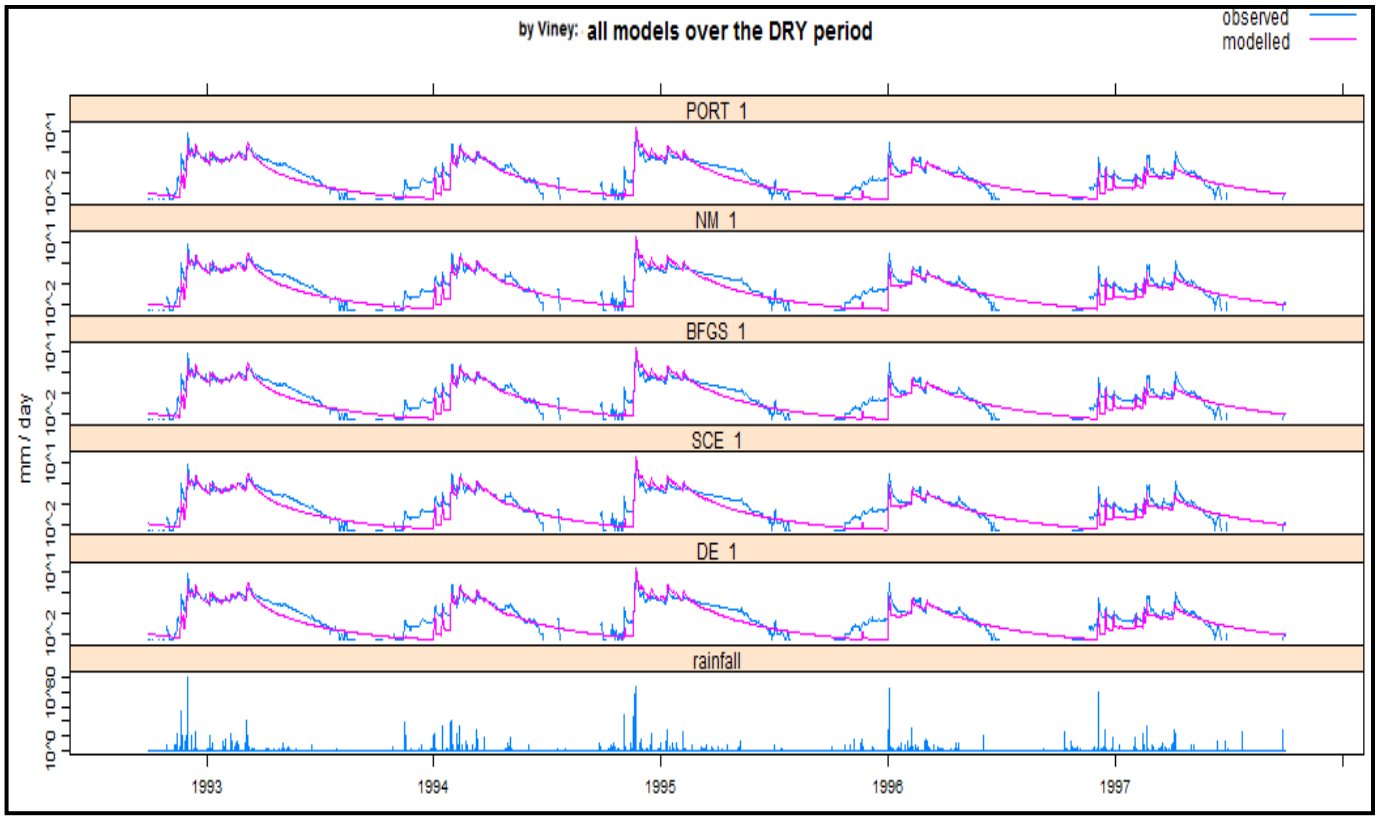
(A) Simulated streamflows (in log scale) – all models

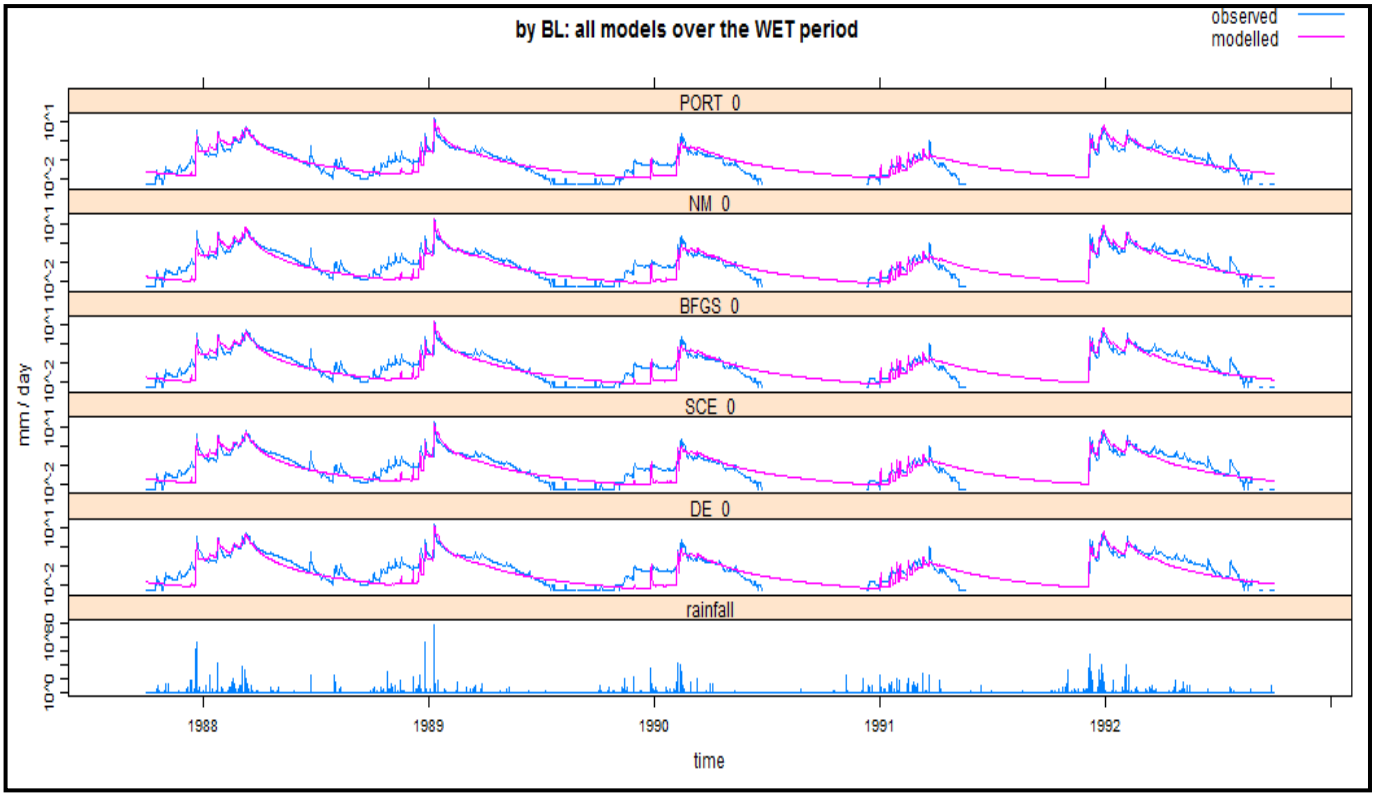


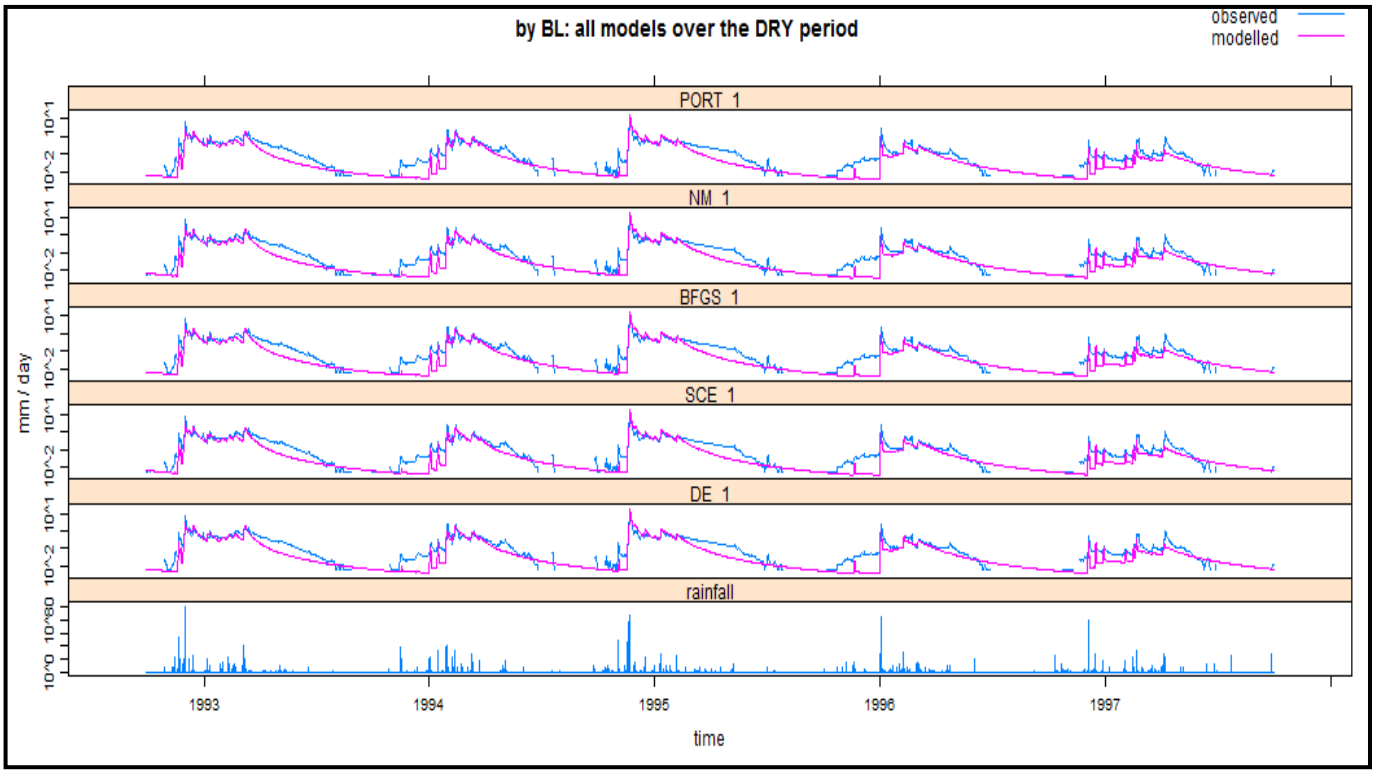


APPENDIX A: GR4J model

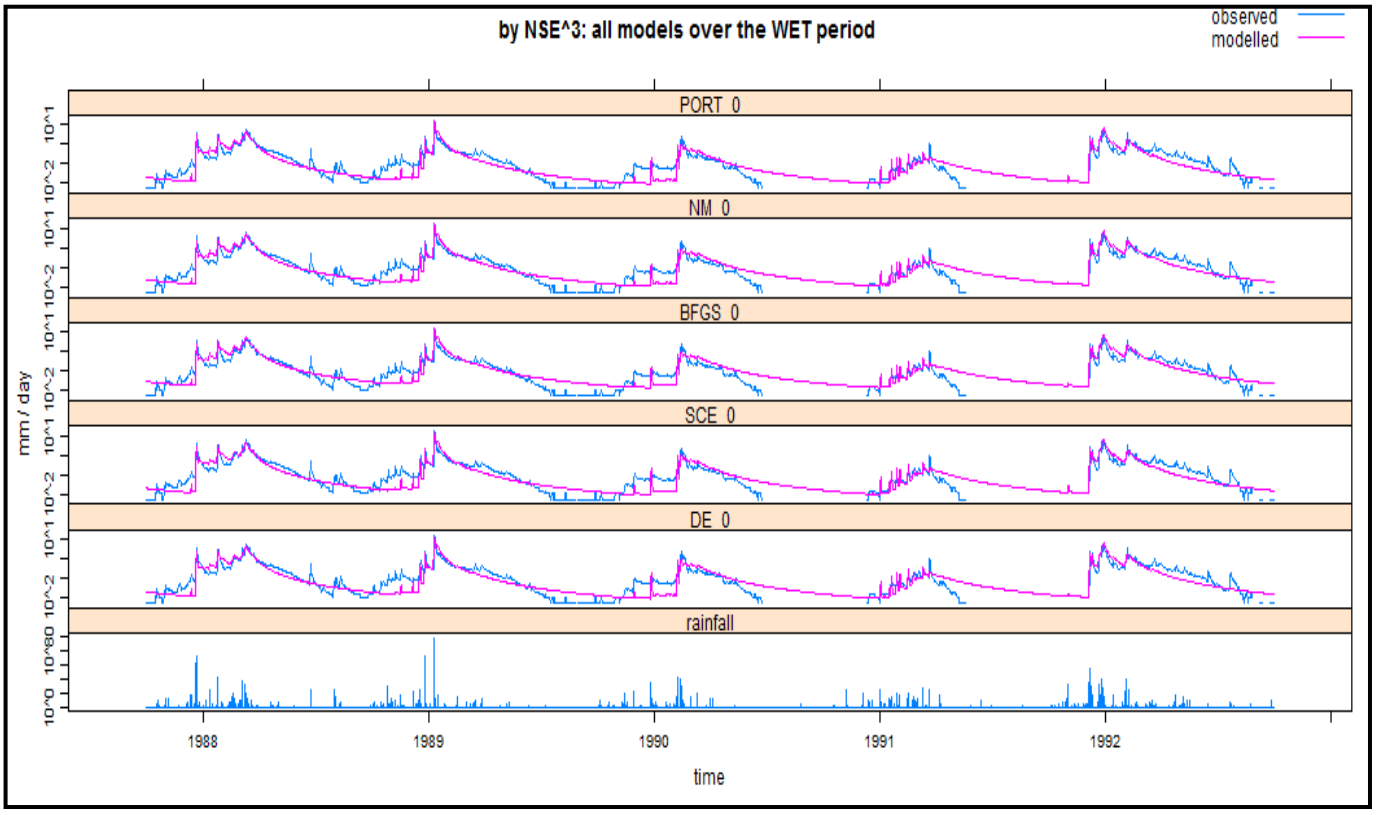


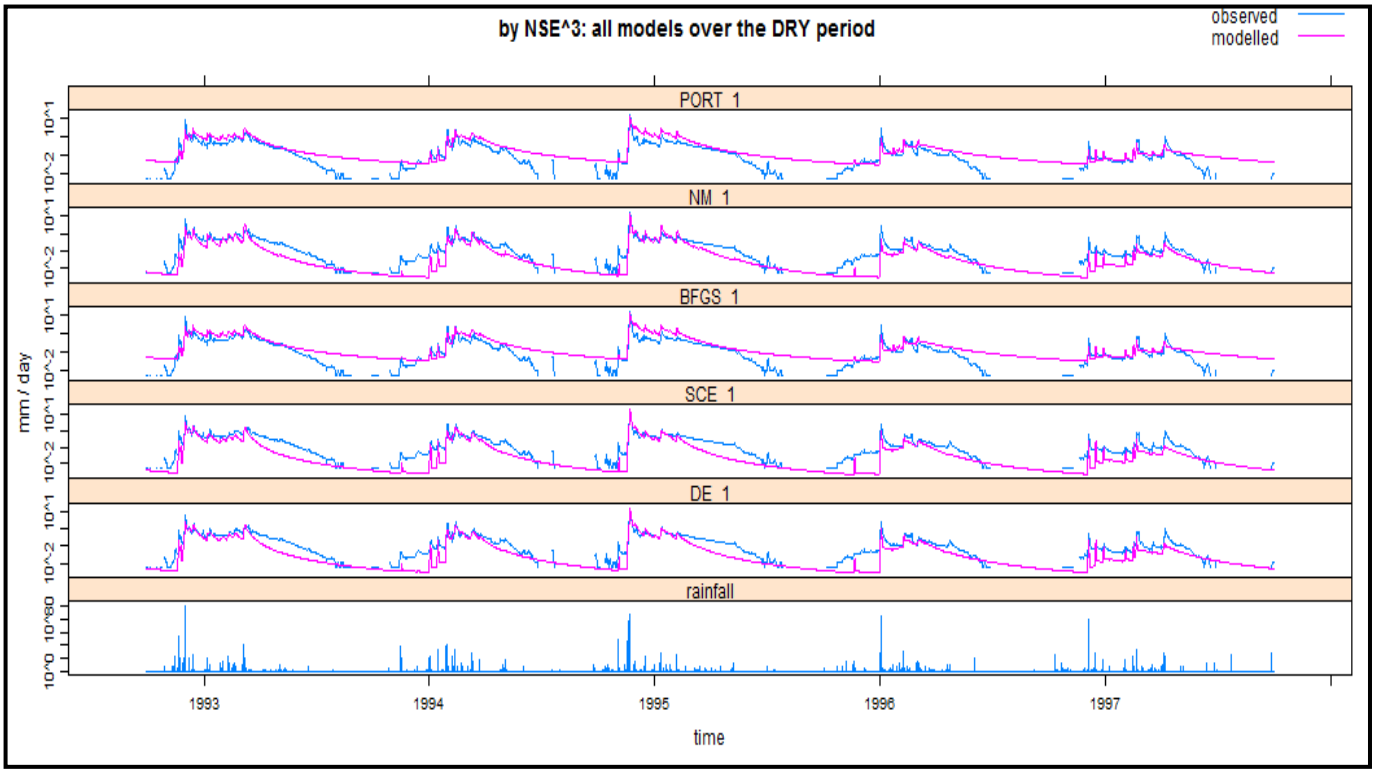






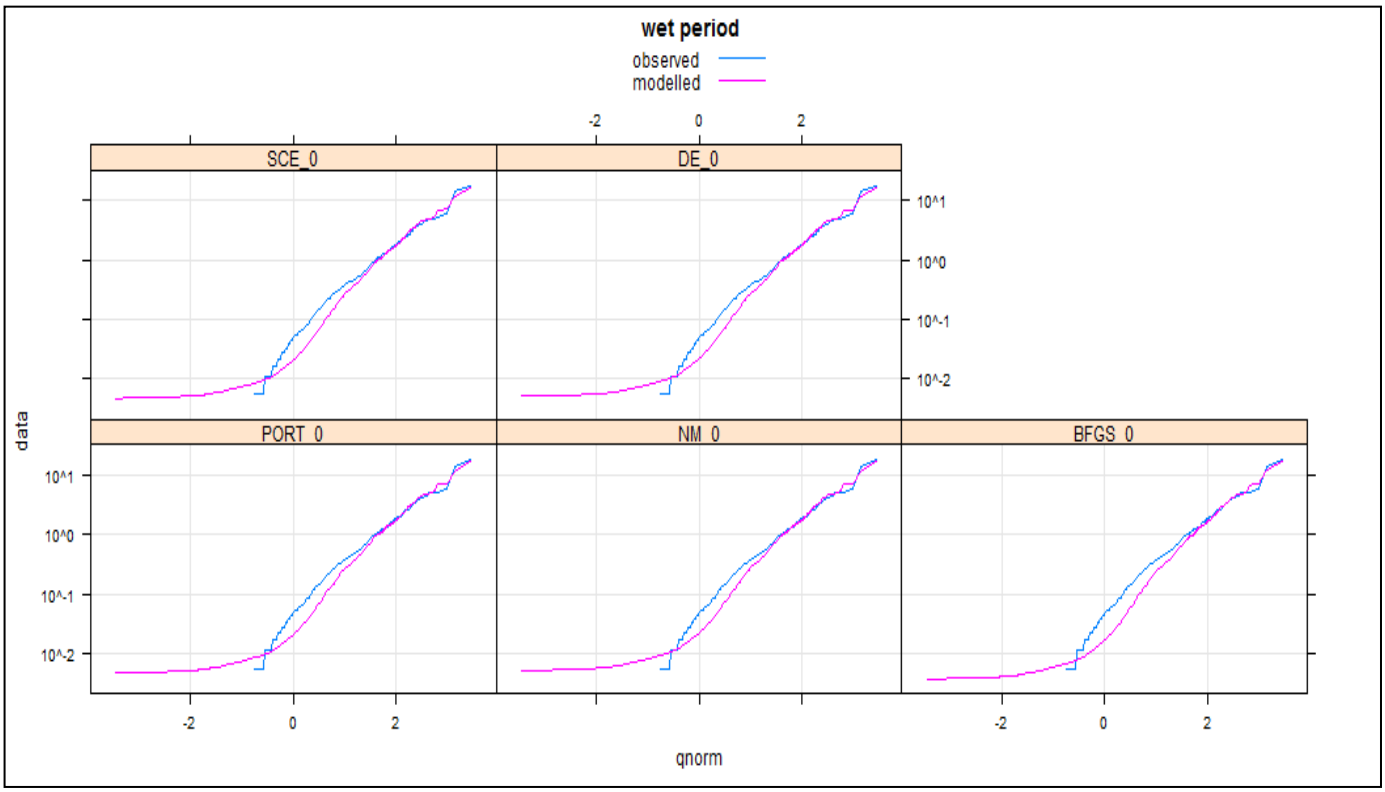
APPENDIX A: GR4J model

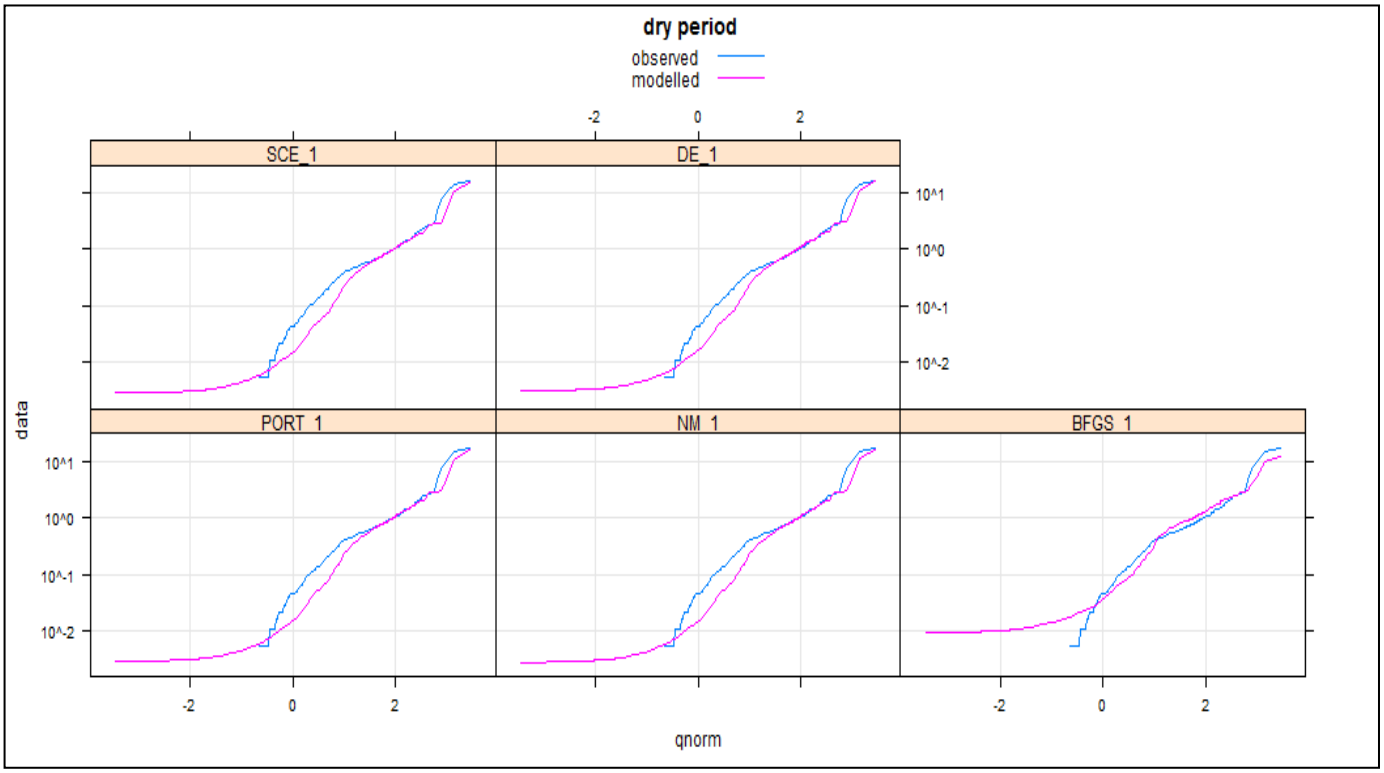




(B) Normal distribution Q-Q plot – all models

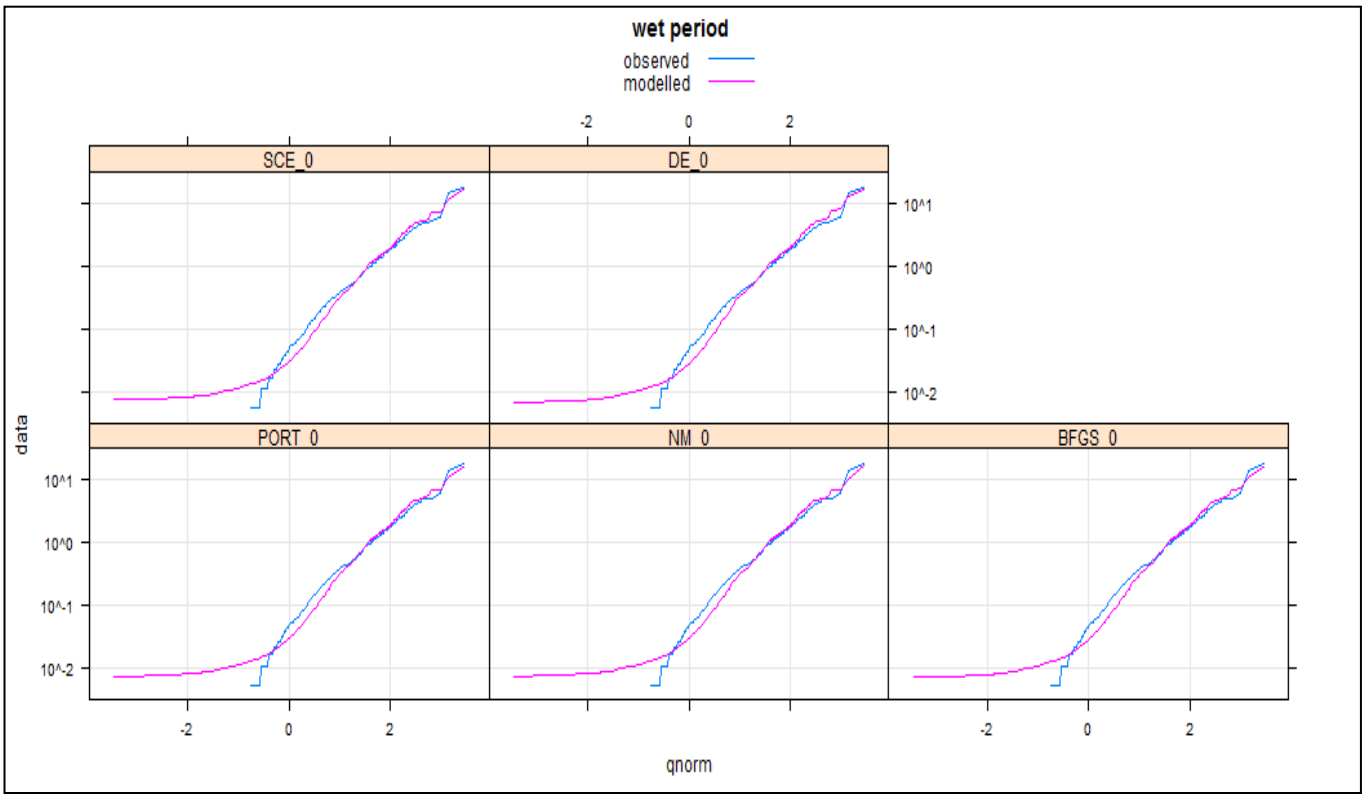
By NSE:

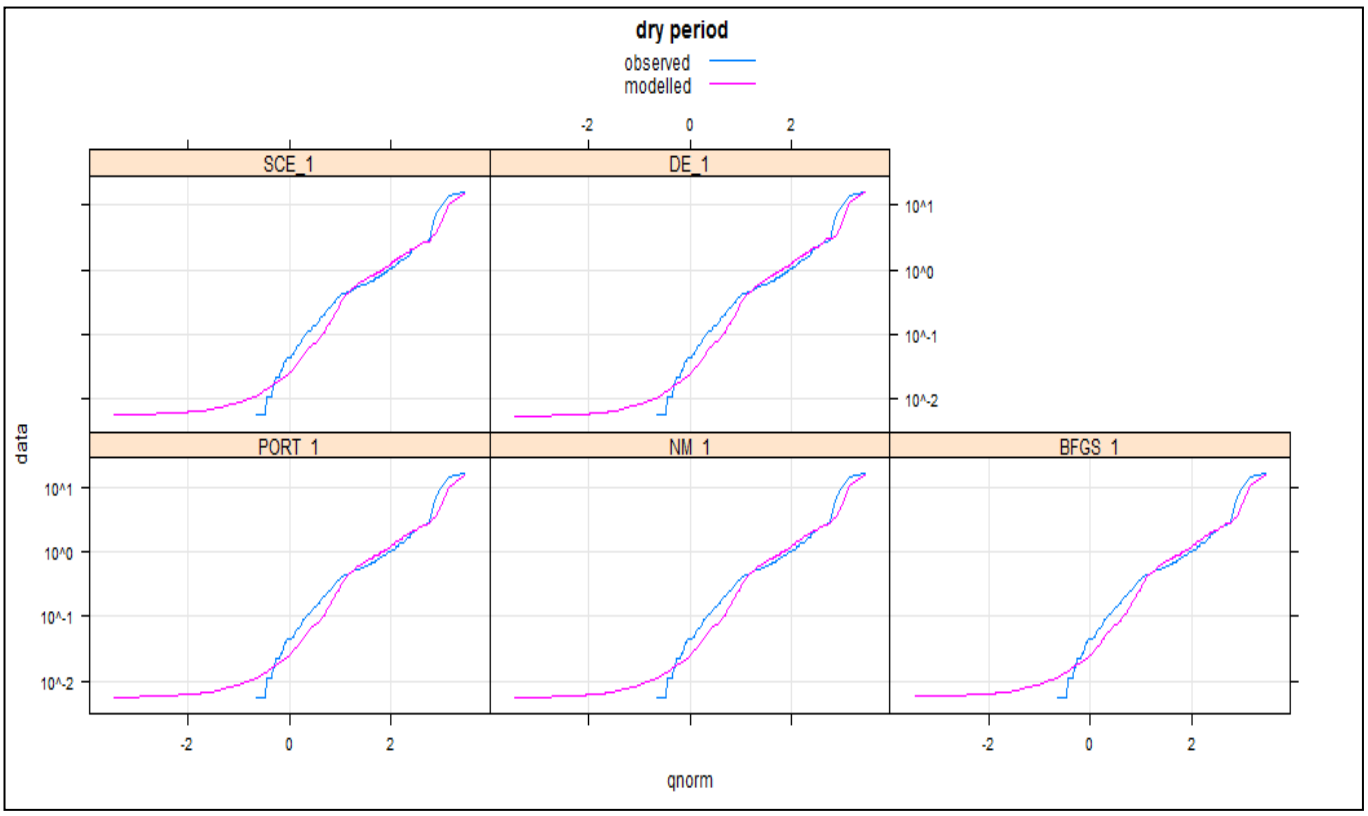




APPENDIX A: GR4J model

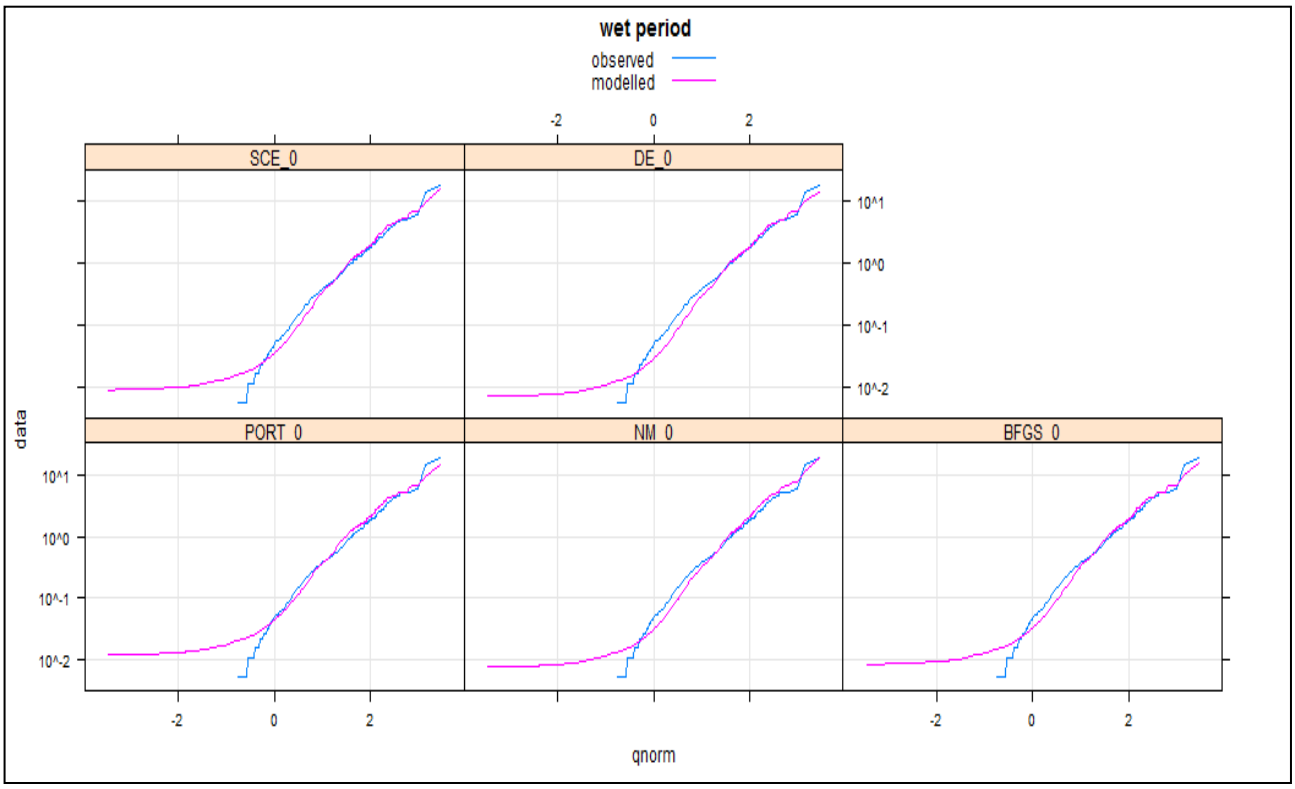
By Viney:

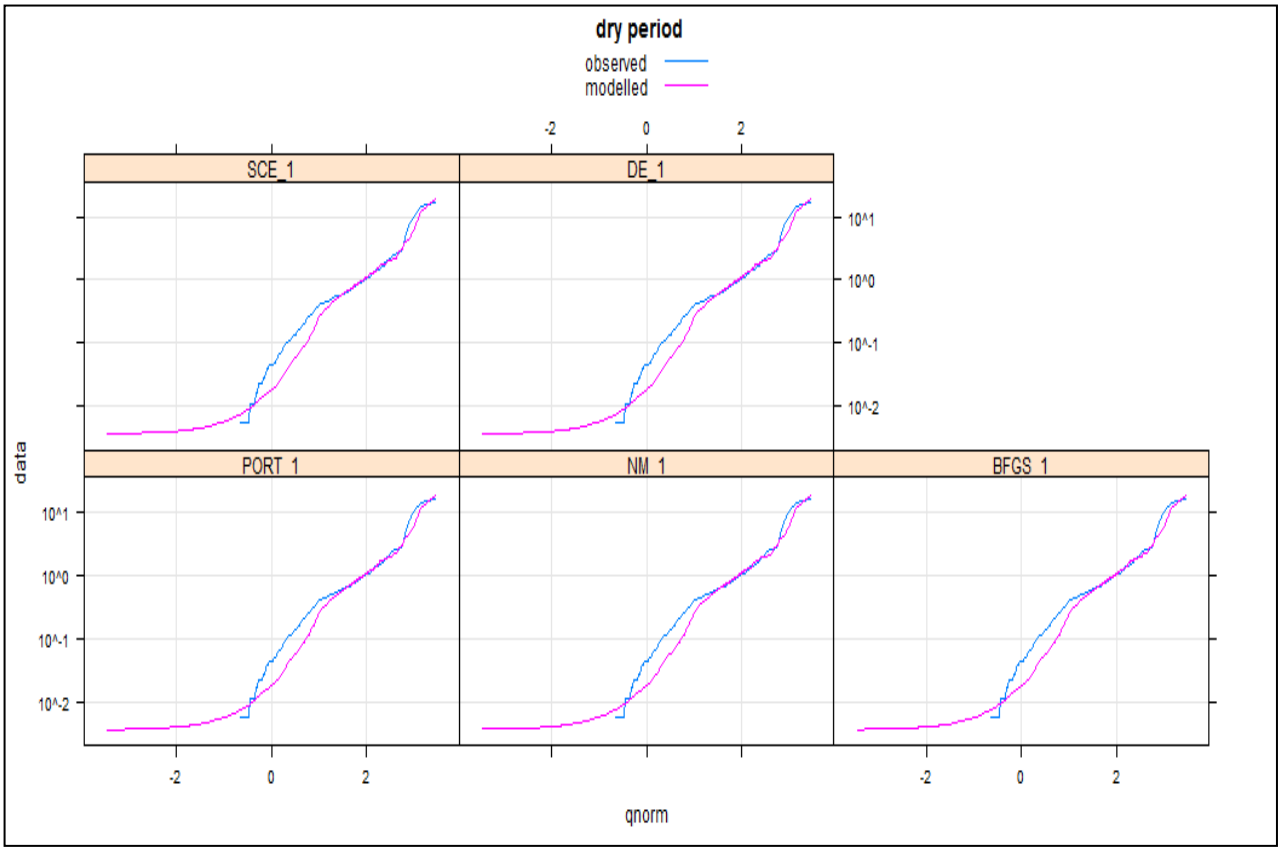




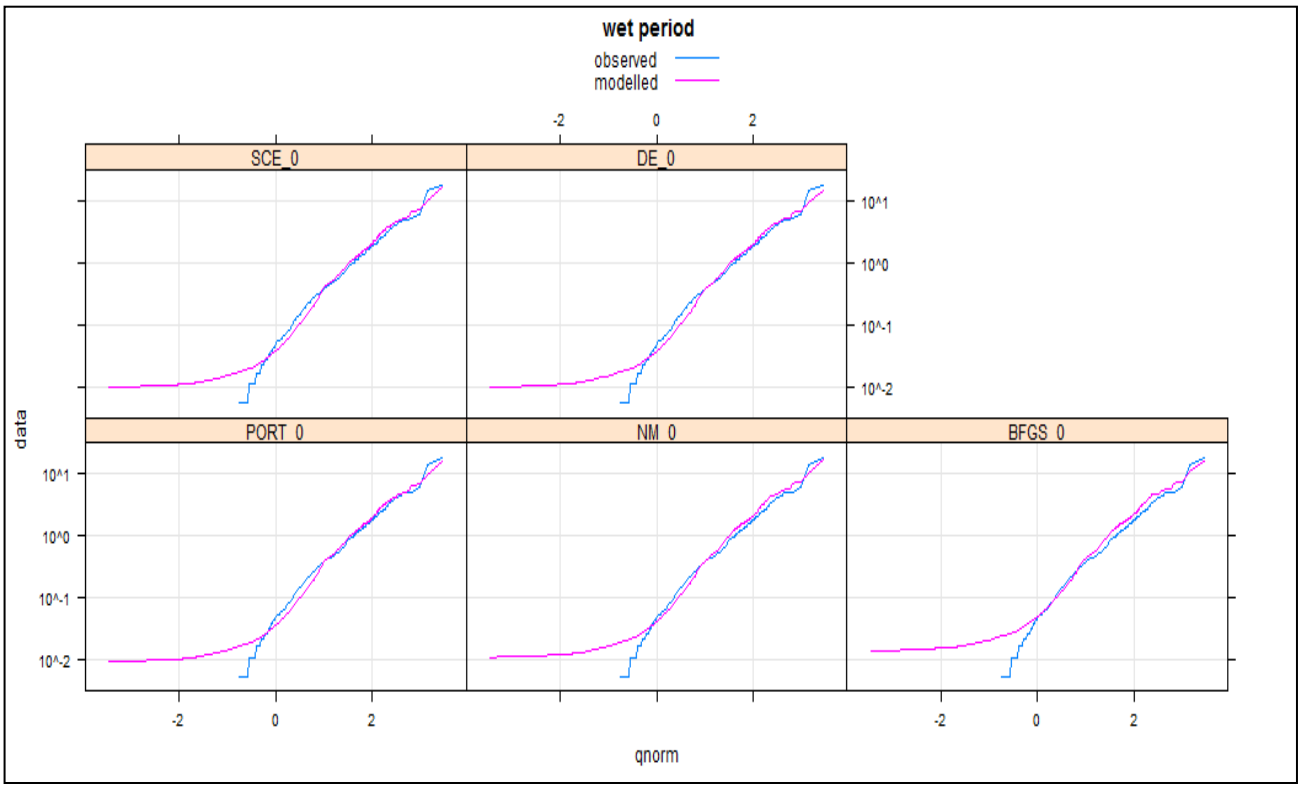
APPENDIX A: GR4J model

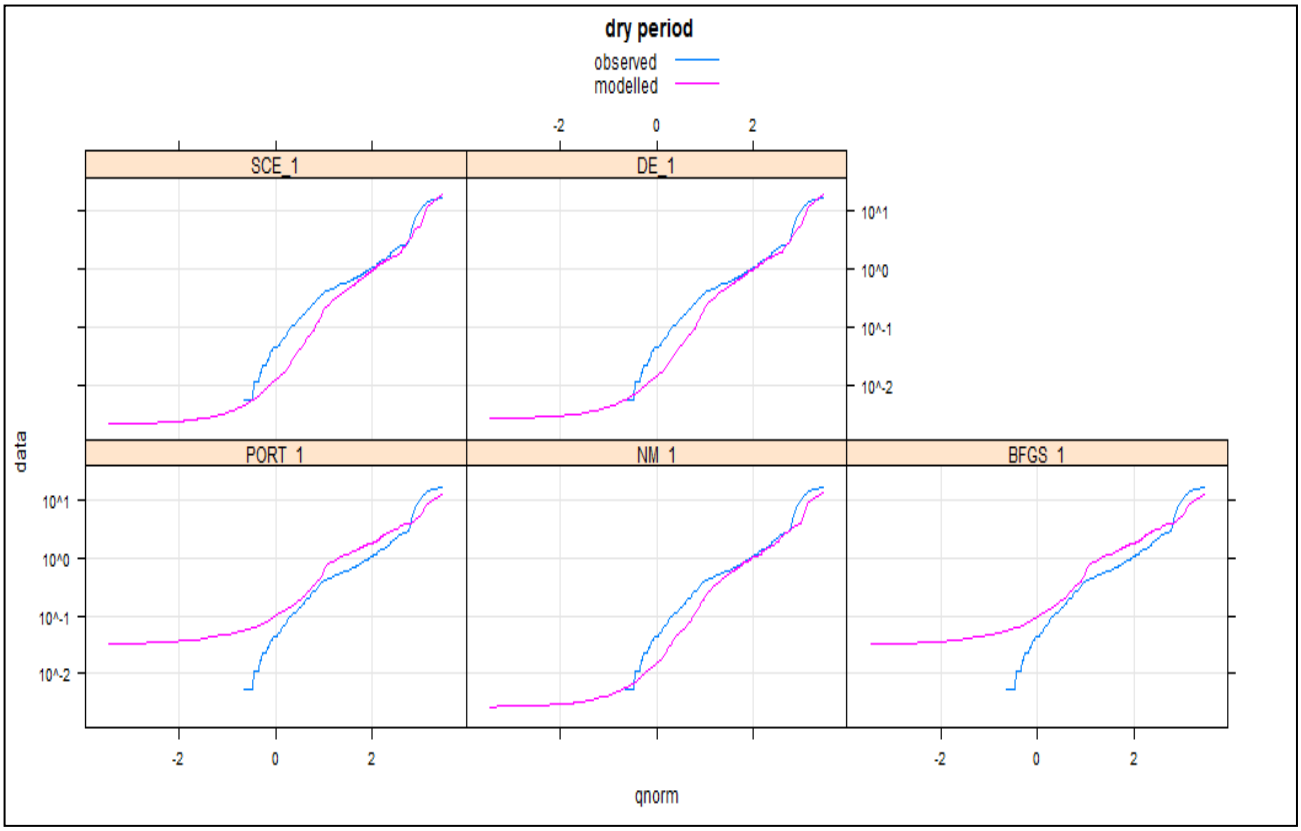
By BL:





By NSE^3:

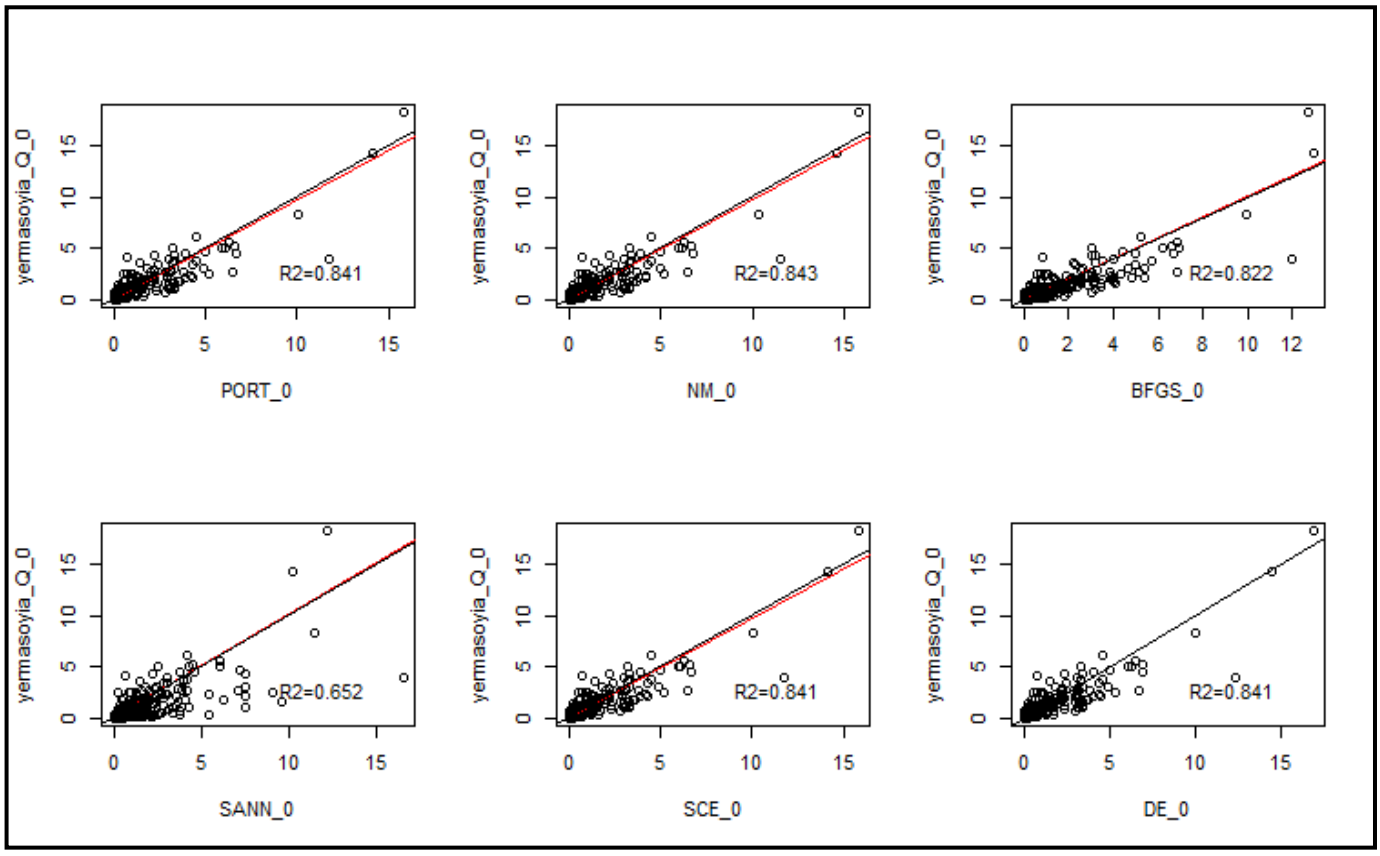




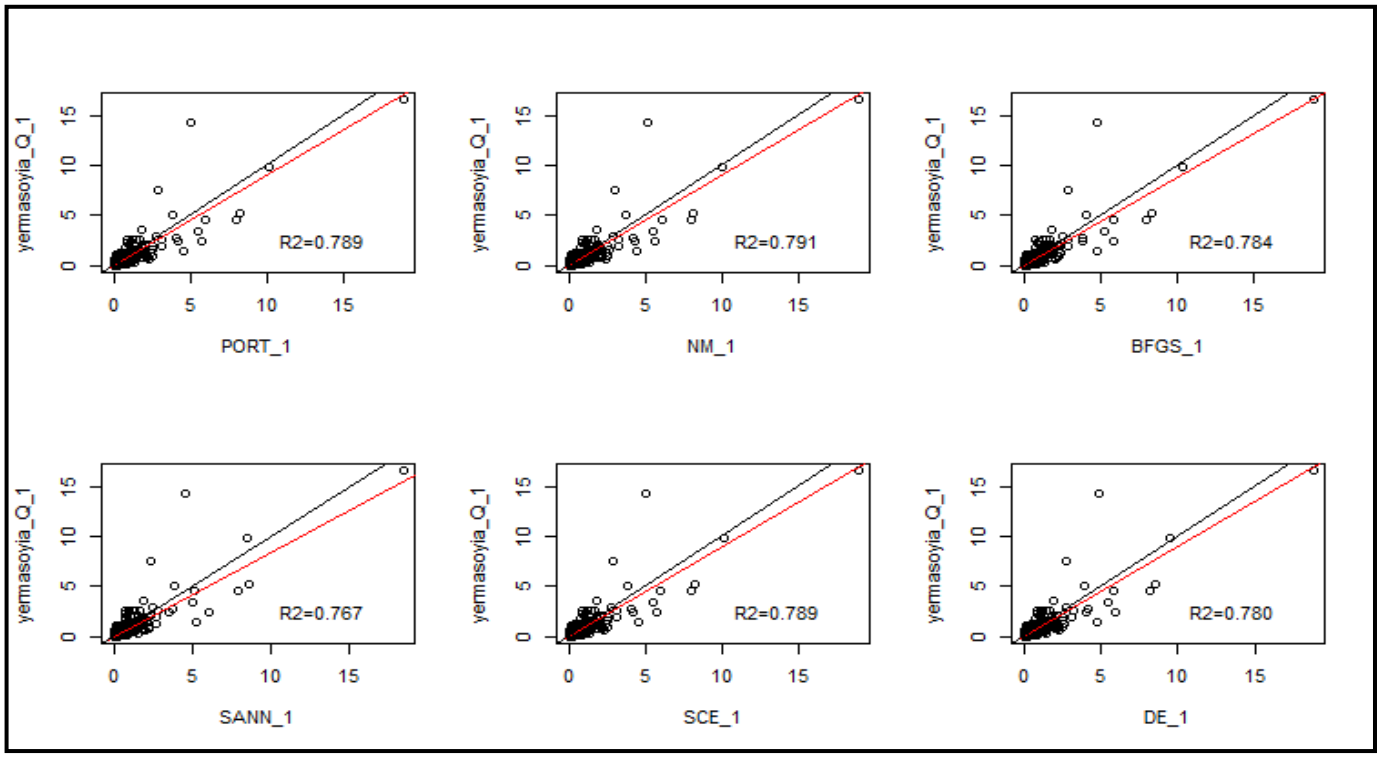
APPENDIX A: GR4J model

(C) Scatterplots

By NSE: wet period

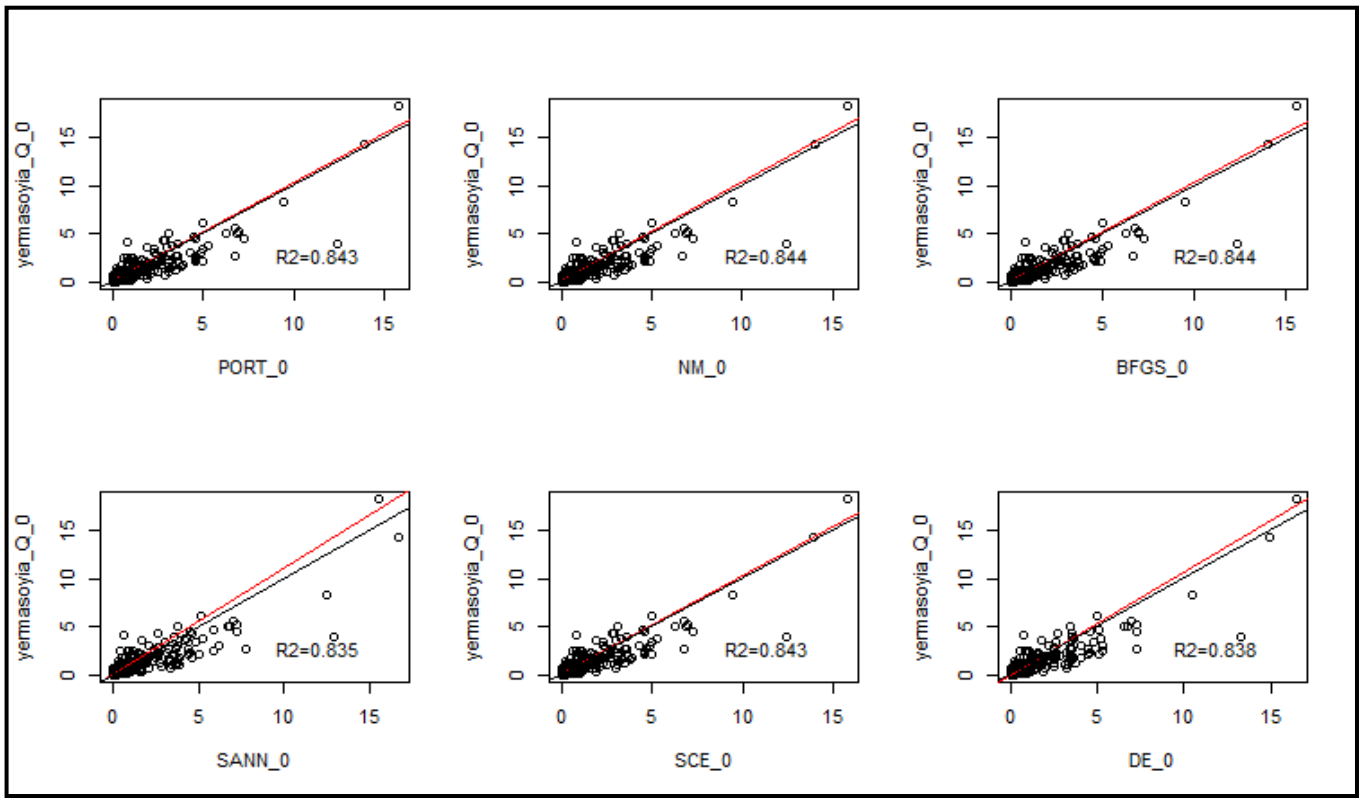


By NSE: dry period

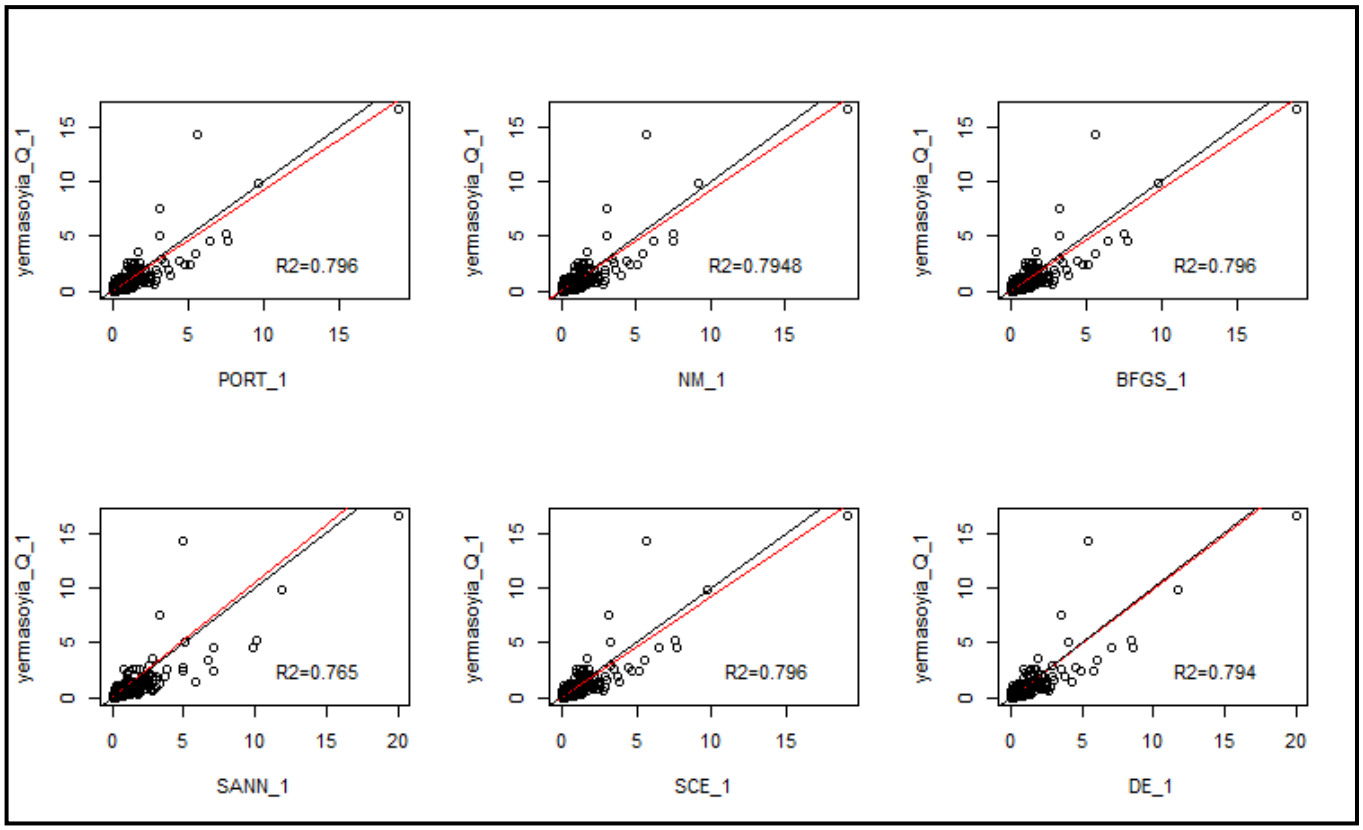


APPENDIX A: GR4J model

By Viney: wet period

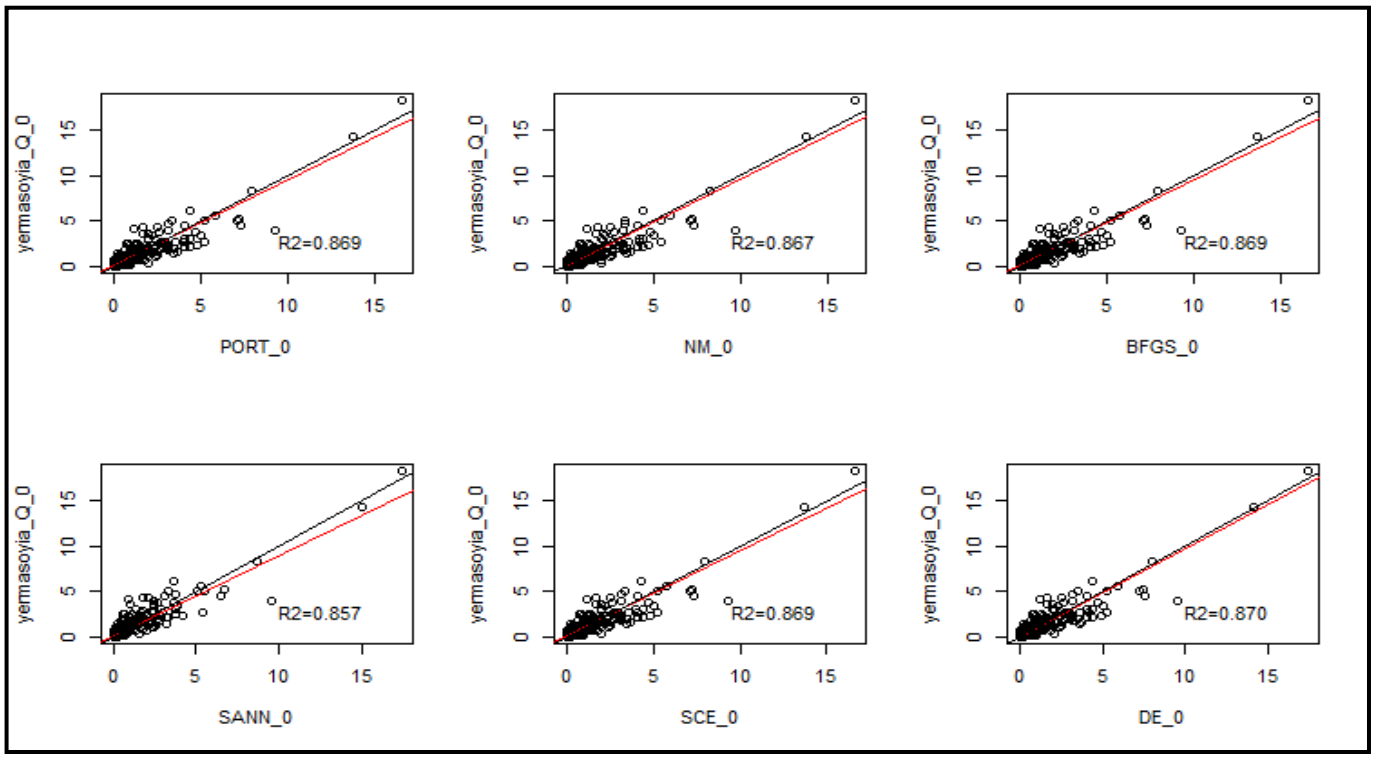


By Viney: dry period

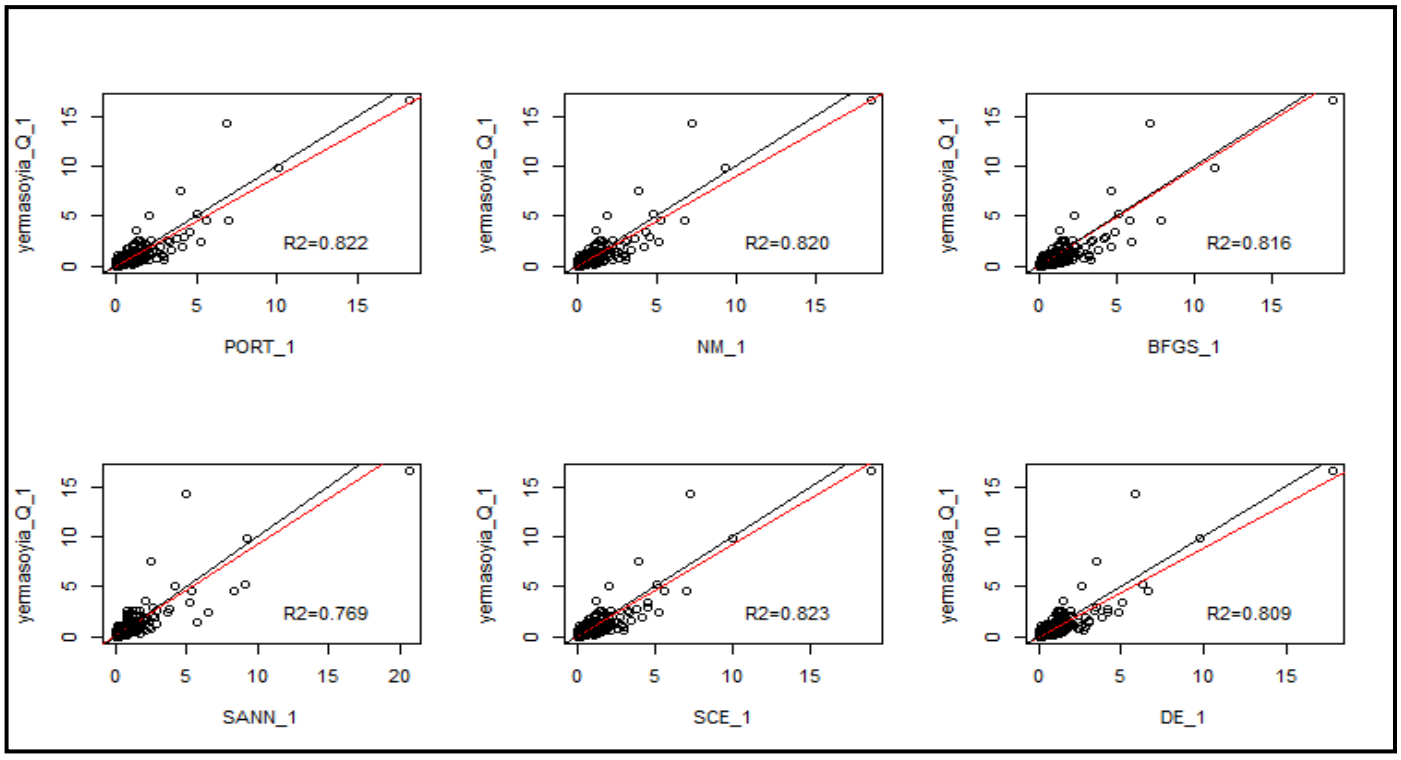


APPENDIX A: GR4J model

By BL: wet period

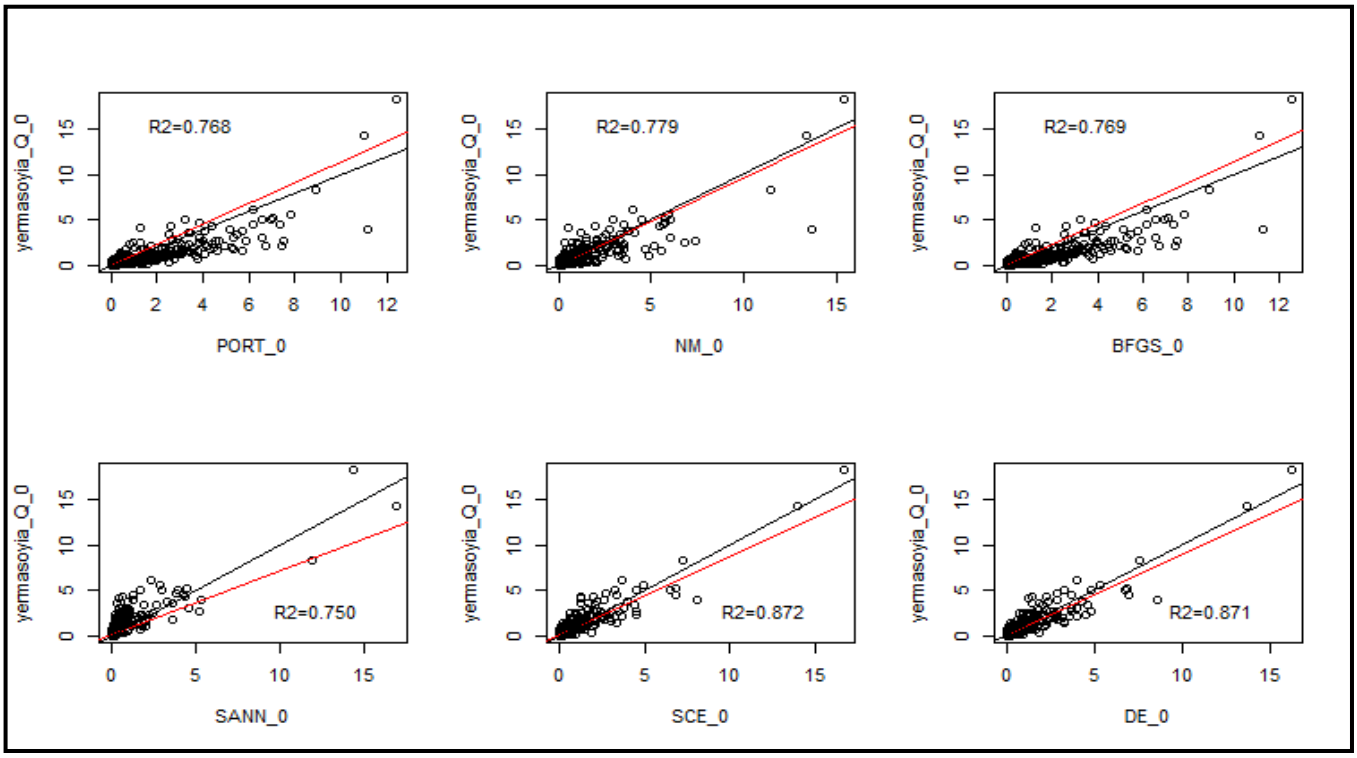


By BL: dry period

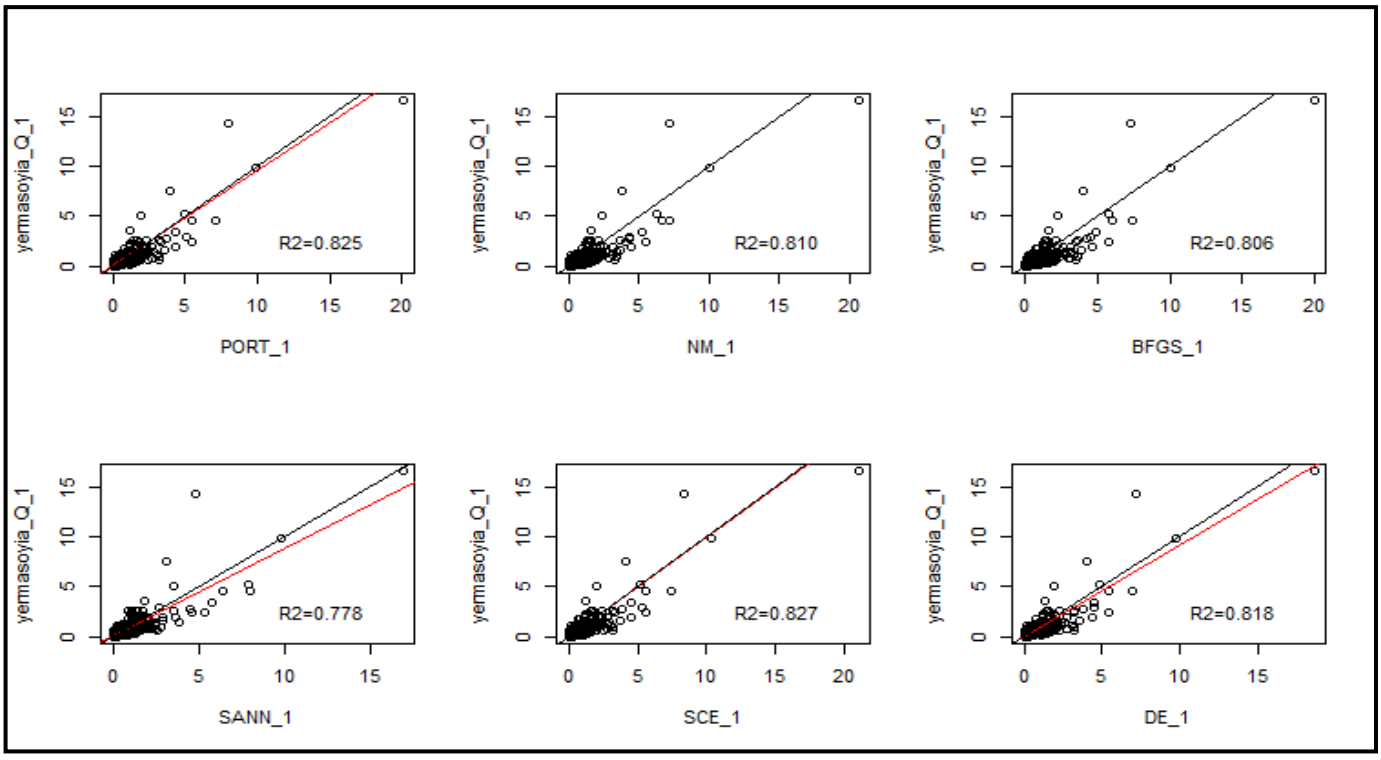


APPENDIX A: GR4J model

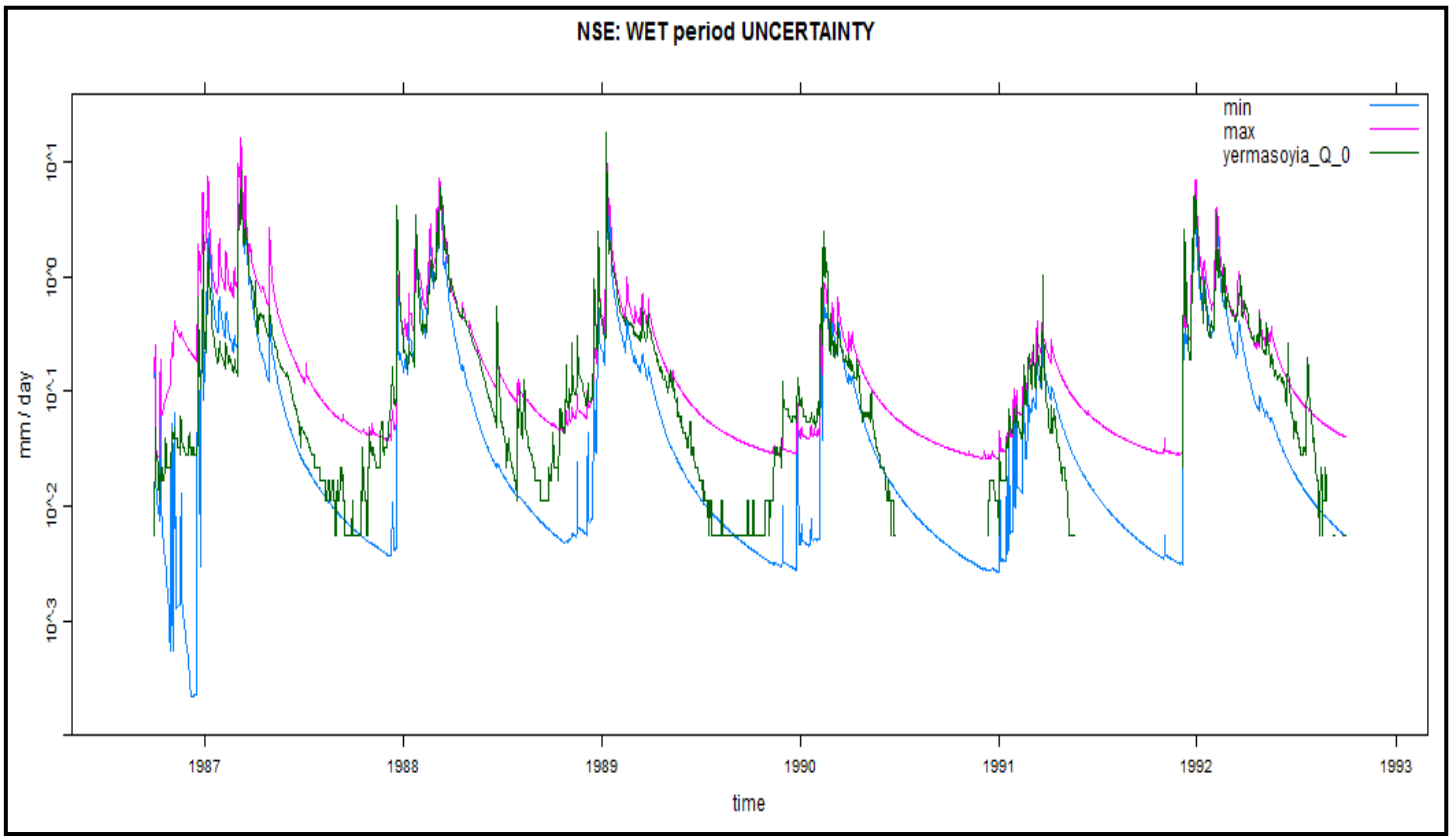
By NSE^3: wet period

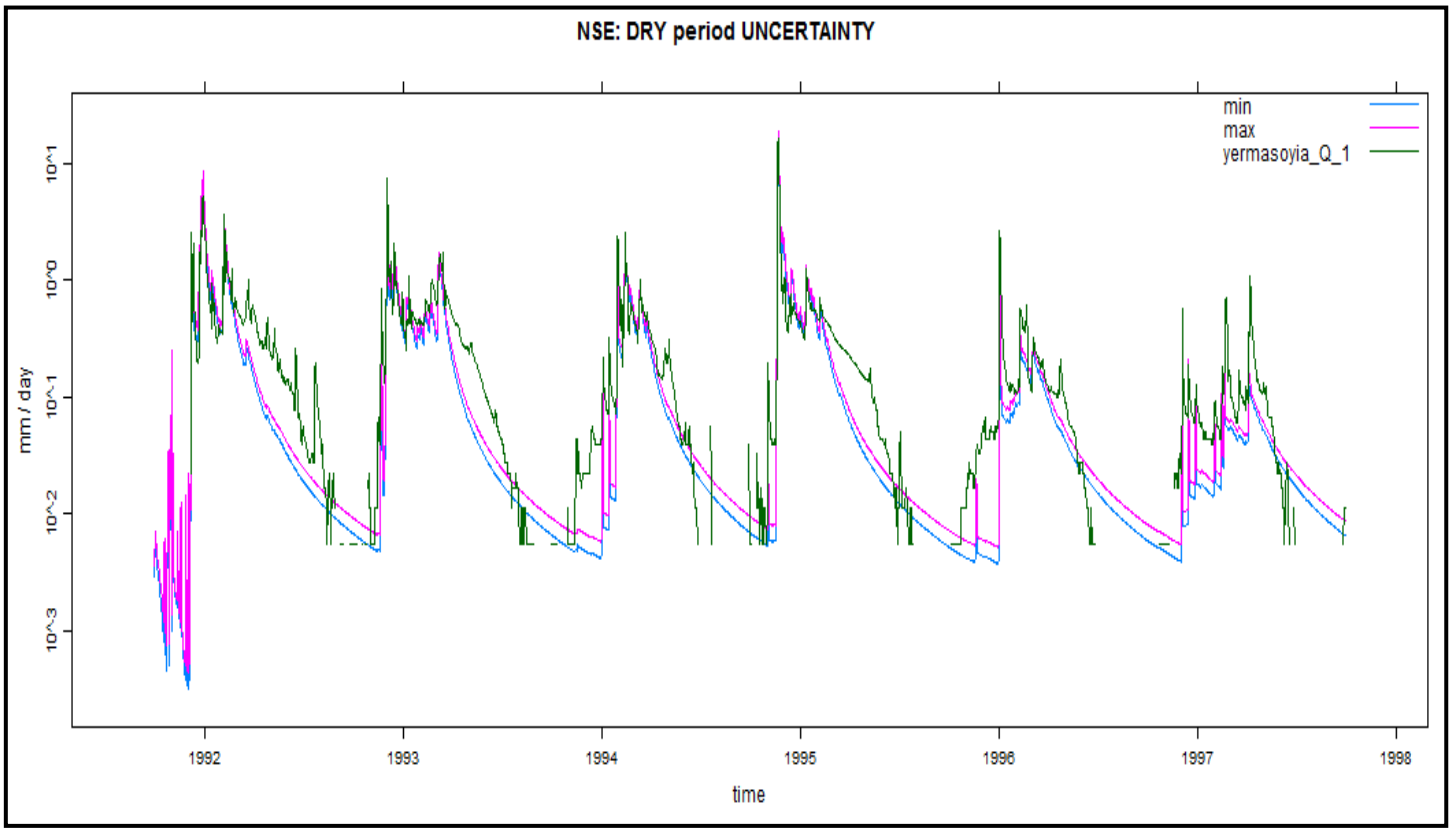


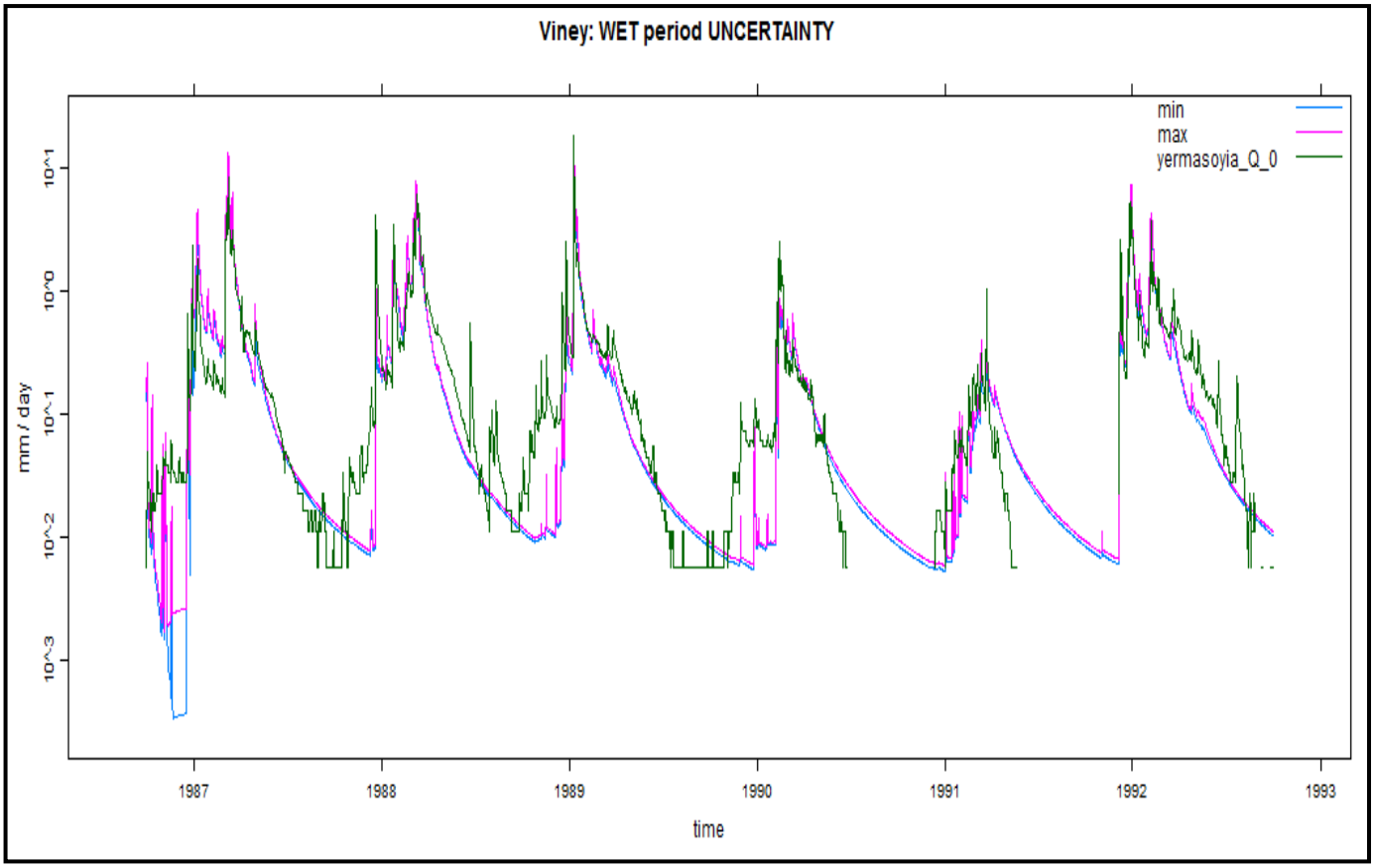
By NSE^3: dry period

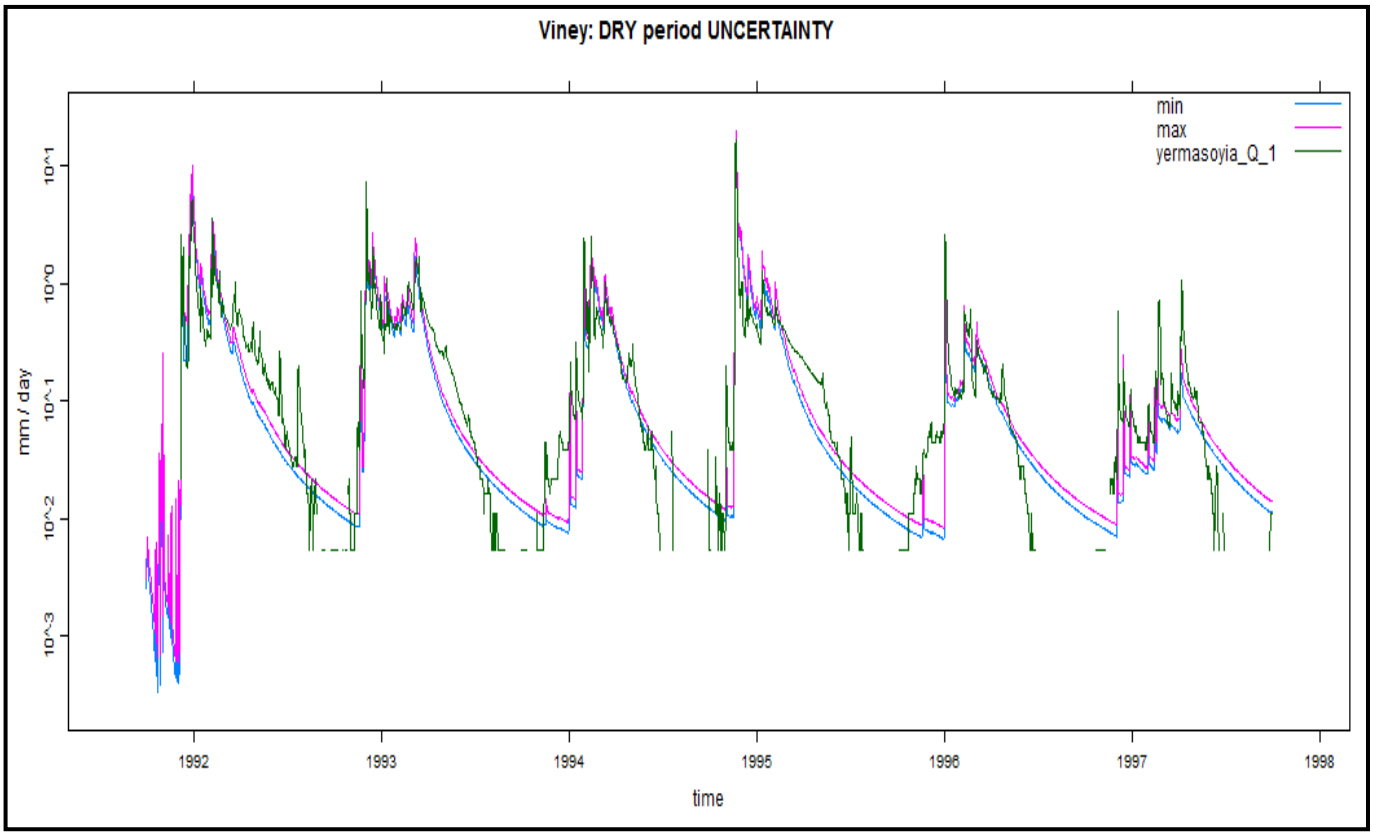


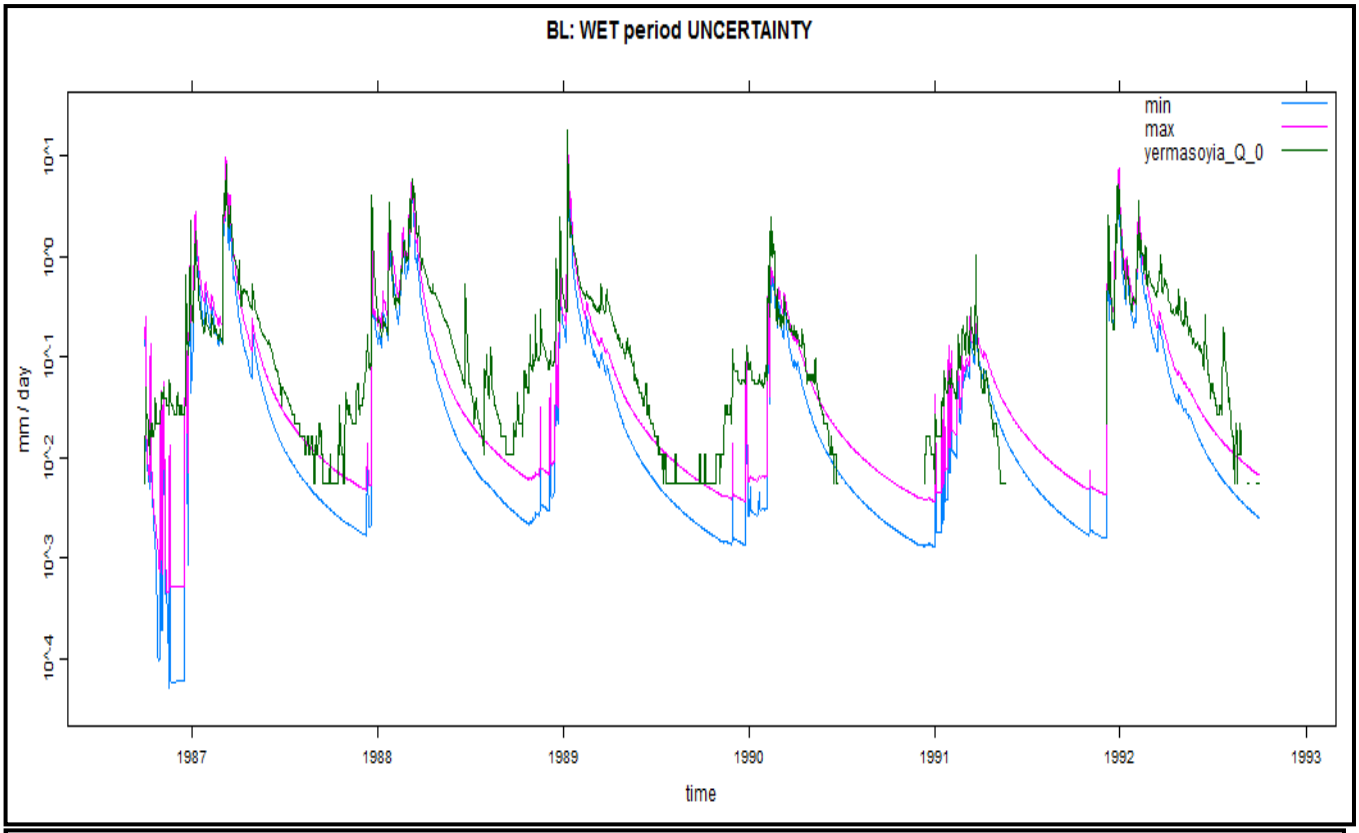
(D) Model structure uncertainty

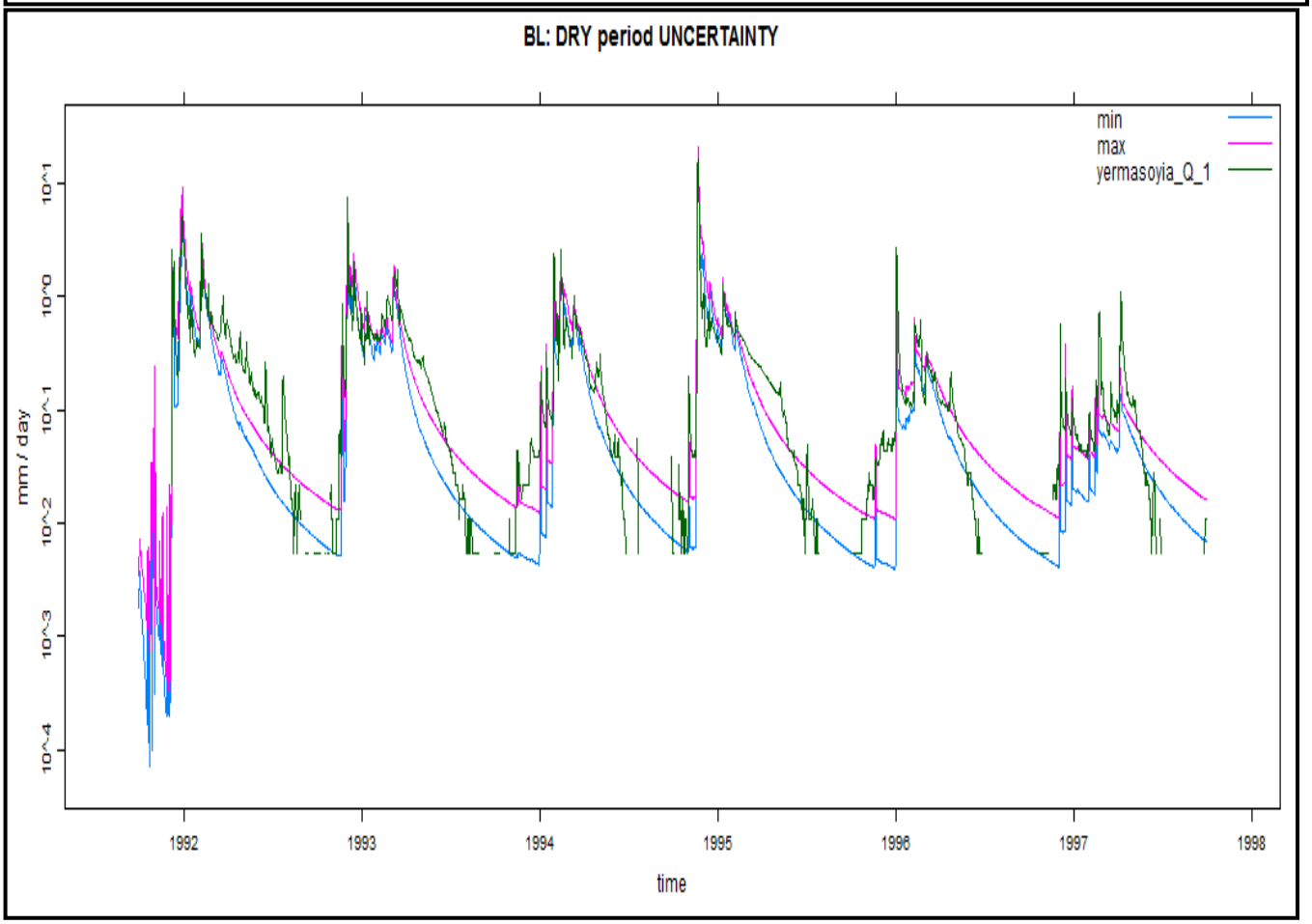


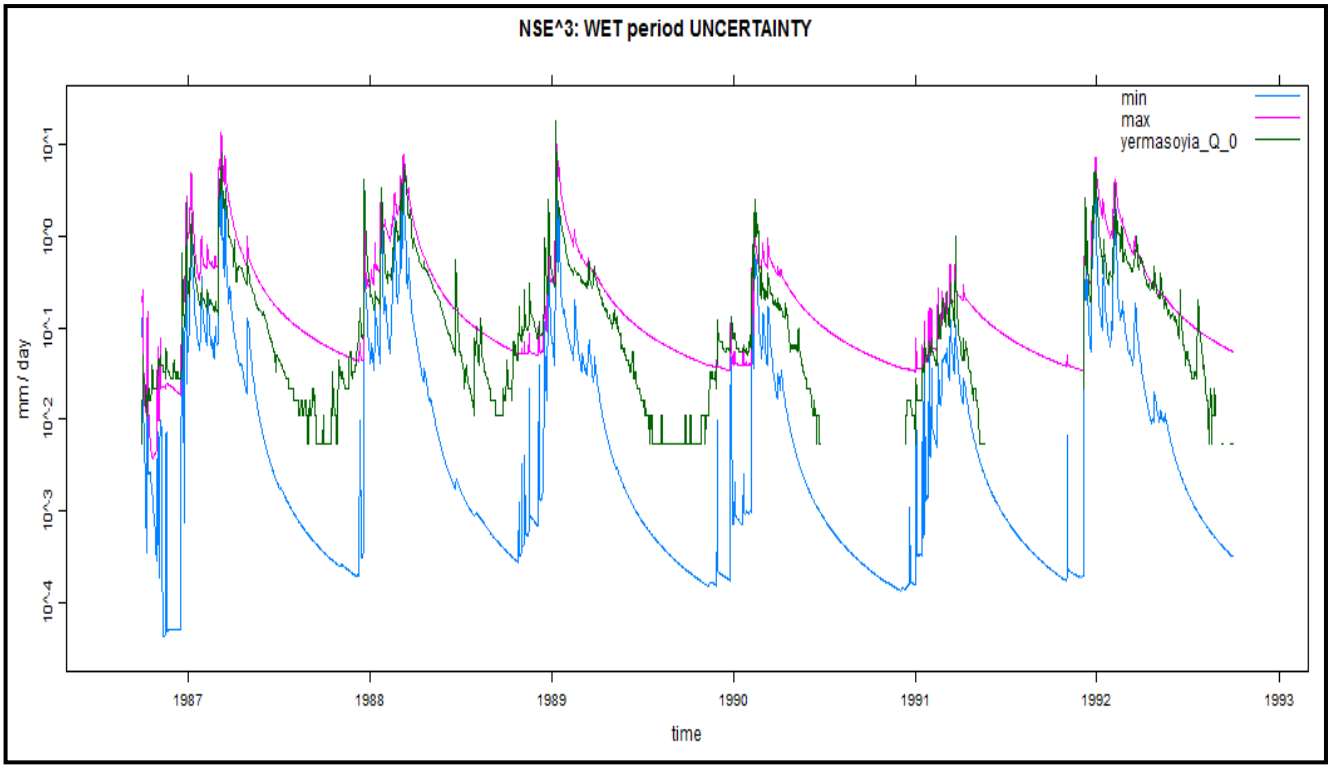


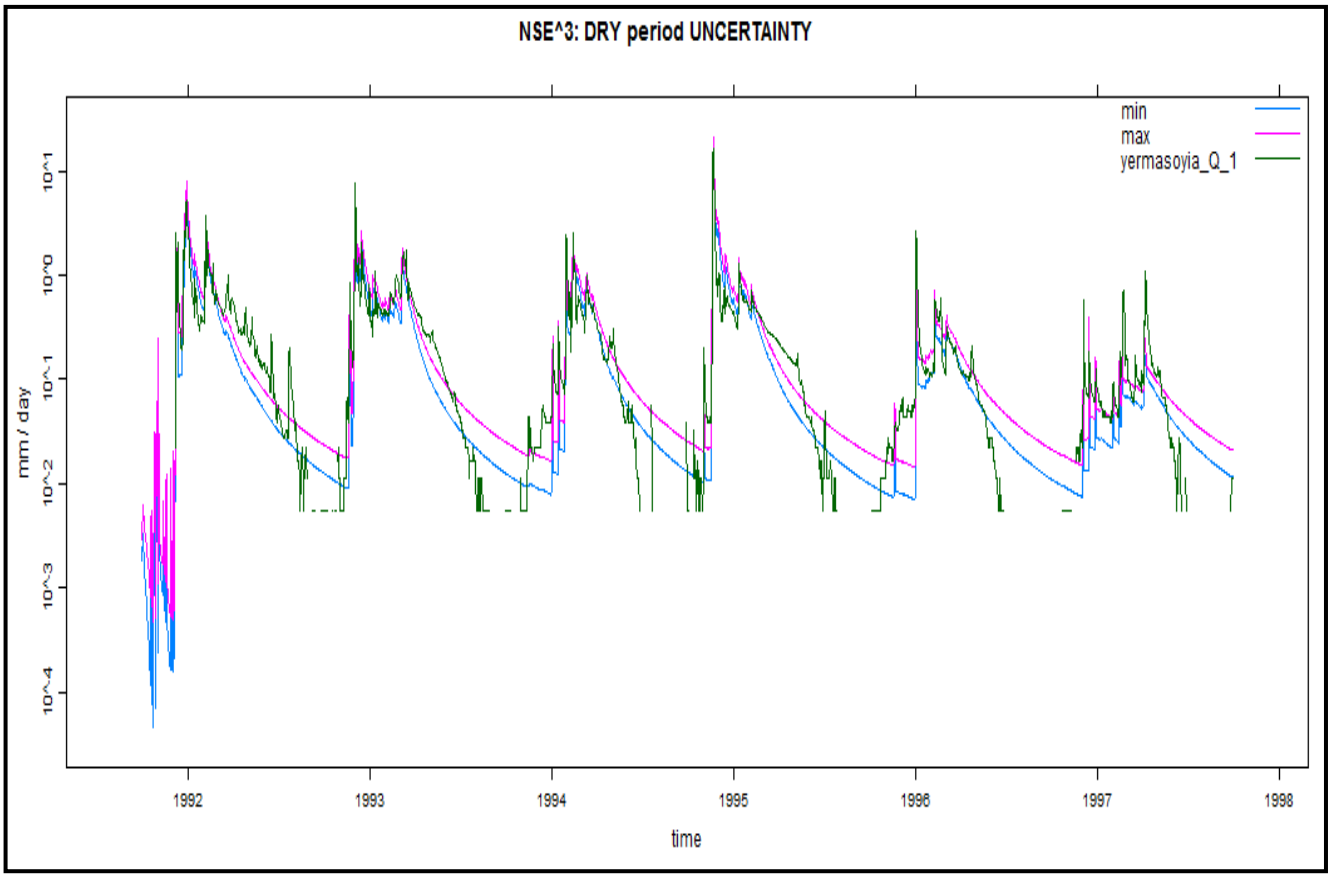








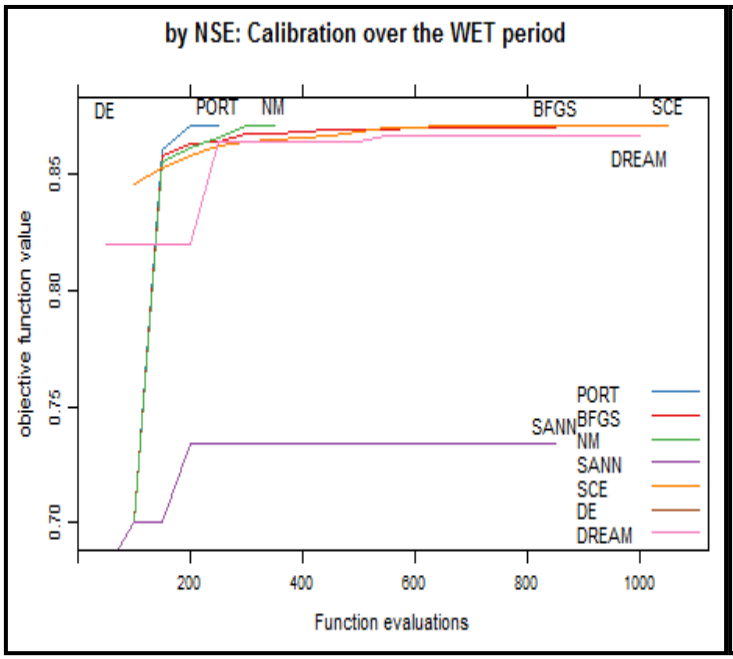


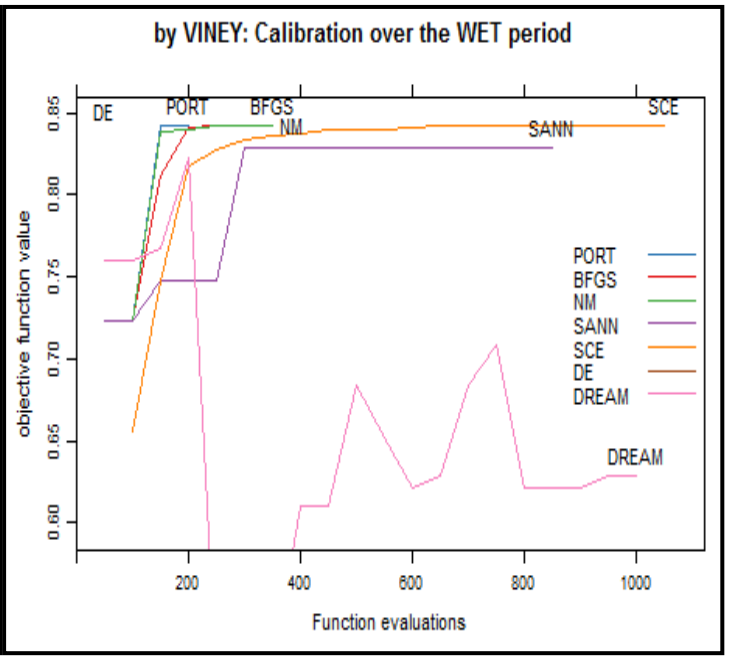


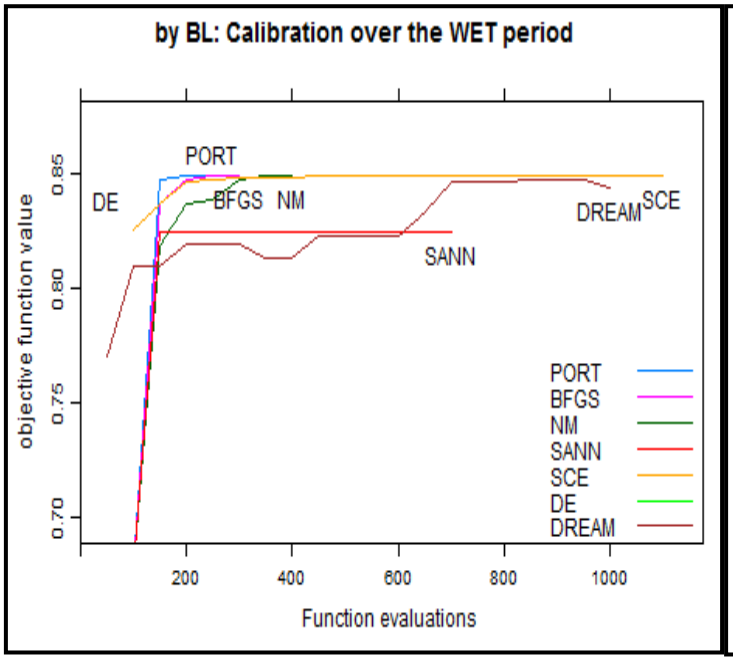
(E) Objective Function and Optimization Algorithm uncertainty

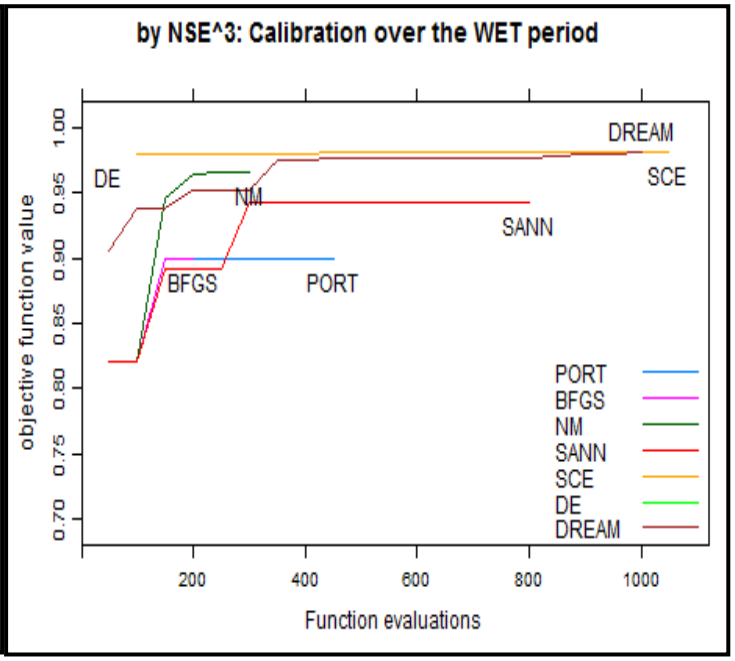
CALIBRATION: **WET** – VALIDATION: **DRY**

Optimization traces:









APPENDIX A: GR4J model

Validation FIT STATISTICS:

Objective Function = **NSE**

	rel.bias	r.sq.sqrt	**r.squared**	viney	BL	nse^3
PORT	-0.303	0.791	0.777	0.386	0.747	0.873
BFGS	-0.309	0.788	0.768	0.353	0.737	0.853
NM	-0.301	0.792	0.778	0.393	0.748	0.874
SANN	-0.622	0.568	0.542	-4.114	0.480	0.683
SCE	-0.303	0.791	0.778	0.385	0.748	0.875
DE	-0.281	0.796	**0.780**	0.468	0.752	0.887
DREAM	-0.248	0.801	0.754	0.537	0.729	0.833

Objective Function = **Viney**

	rel.bias	r.sq.sqrt	r.squared	**viney**	BL	nse^3
PORT	-0.090	0.809	0.733	0.719	0.724	0.823
BFGS	-0.084	0.809	0.730	0.719	0.722	0.822
NM	-0.093	0.810	0.733	0.718	0.724	0.823
SANN	-0.071	0.802	0.698	0.690	0.691	0.797
SCE	-0.090	0.809	0.733	0.719	0.724	0.823
DE	-0.090	0.807	0.725	0.711	0.716	0.816
DREAM	-0.301	0.779	0.725	0.340	0.695	0.802

Objective Function = **BL**

	rel.bias	r.sq.sqrt	r.squared	viney	**BL**	nse^3
PORT	-0.207	0.815	0.777	0.649	0.757	0.872
BFGS	-0.206	0.815	0.778	0.651	0.757	0.873
NM	-0.203	0.814	0.774	0.651	0.754	0.865
SANN	-0.409	0.737	0.752	-0.249	0.712	0.850
SCE	-0.205	0.815	0.778	0.651	0.757	0.872
DE	-0.192	0.816	0.775	0.670	0.756	0.877
DREAM	-0.308	0.789	0.780	0.369	0.749	0.880

Objective Function = **NSE^3**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	**nse^3**
PORT	0.608	0.529	0.551	-0.227	0.490	0.815
BFGS	0.601	0.535	0.554	-0.207	0.494	0.816
NM	-0.310	0.768	0.677	0.258	0.646	0.773
SANN	-0.681	0.426	0.655	-6.304	0.587	0.814
SCE	-0.338	0.778	0.781	0.234	0.747	**0.883**
DE	-0.293	0.796	0.781	0.426	0.752	0.874
DREAM	-0.308	0.789	0.780	0.369	0.749	0.880

APPENDIX A: GR4J model

Parameter Stability:

Objective Function = **NSE**

	x2	x3	x4	x1	etmult
PORT	-14.77	116.1	1.13	274	1
BFGS	-11.43	97.1	1.15	307	1
NM	-14.88	117.2	1.13	273	1
SANN	-2.57	20.1	1.32	642	1
SCE	-15.17	118.2	1.13	271	1
DE	-18.07	136.8	1.14	243	1
DREAM	-7.82	86.0	1.15	339	1

Objective Function = **Viney**

	x2	x3	x4	x1	etmult
PORT	-5.00	92.3	1.11	357	1
BFGS	-5.00	93.5	1.11	356	1
NM	-5.00	91.5	1.11	357	1
SANN	-2.47	62.0	1.19	413	1
SCE	-5.00	92.3	1.11	357	1
DE	-3.71	75.8	1.13	383	1
DREAM	-4.92	59.0	1.18	406	1

Objective Function = **BL**

	x2	x3	x4	x1	etmult
PORT	-12.312	124.090	1.112	277.803	1
BFGS	-12.440	125.214	1.110	276.384	1
NM	-11.197	117.924	1.114	288.558	1
SANN	-11.810	76.829	1.182	320.338	1
SCE	-12.399	124.971	1.110	276.781	1
DE	-12.234	123.484	1.109	273.257	1
DREAM	-16.403	122.988	1.129	261.423	1

Objective Function = **NSE^3**

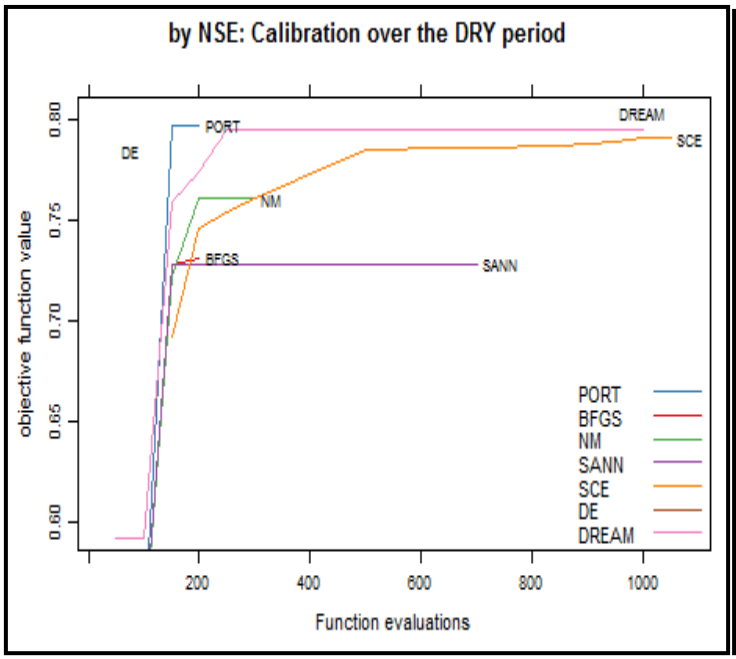
	x2	x3	x4	x1	etmult
PORT	-1.418	210.008	1.000	308.409	1
BFGS	-1.425	206.931	1.000	309.410	1
NM	-1.935	35.986	1.192	514.059	1
SANN	-11.789	31.774	1.442	396.009	1
SCE	-18.017	124.644	1.129	256.532	1
DE	-15.238	123.484	1.123	270.180	1
DREAM	-16.403	122.988	1.129	261.423	1

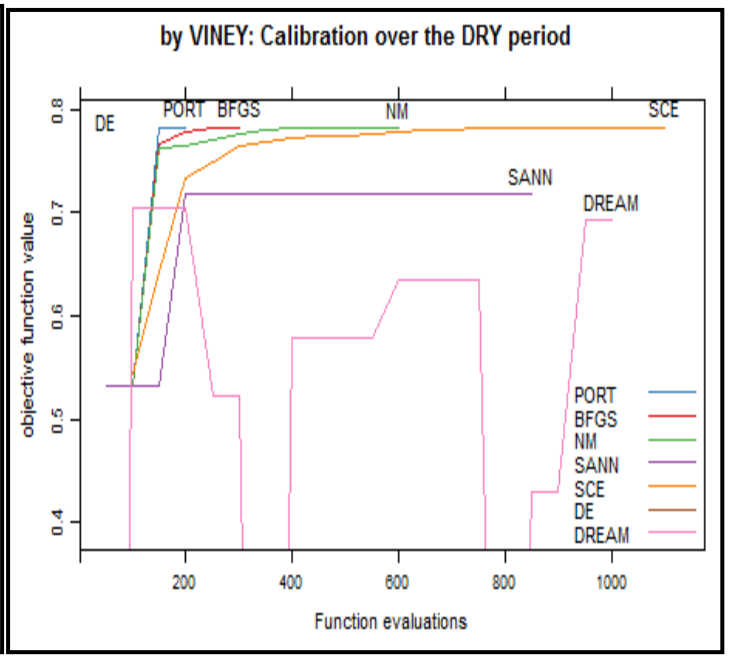
APPENDIX A: GR4J model

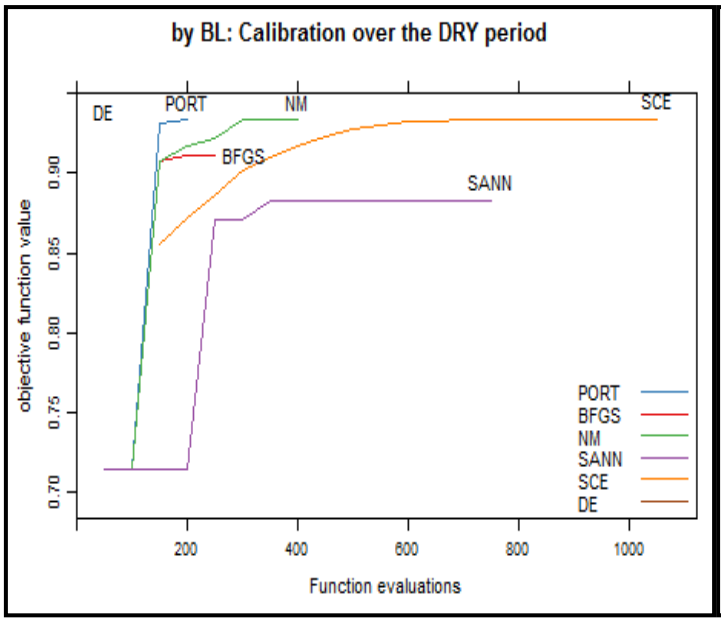
(F) Objective Function and Optimization Algorithm uncertainty

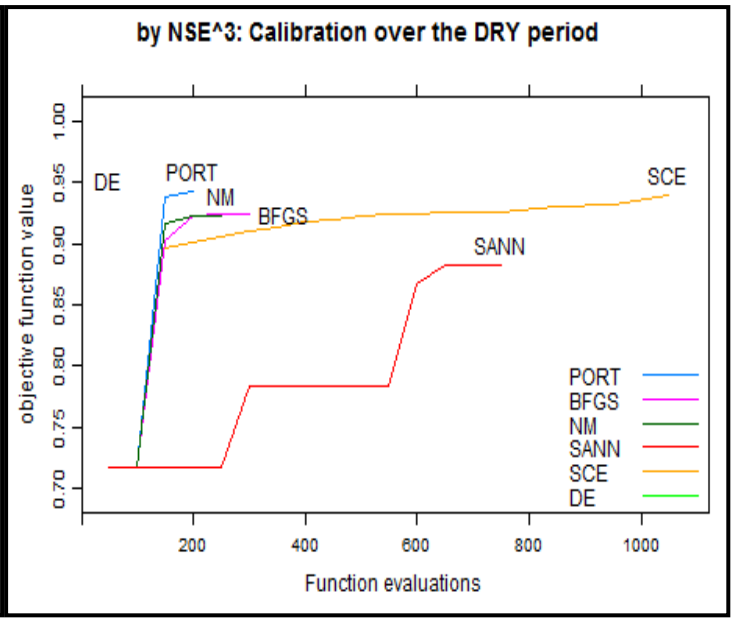
CALIBRATION: **DRY** – VALIDATION: **WET**

Optimization traces:









APPENDIX A: GR4J model

Validation FIT STATISTICS:

Objective Function = **NSE**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.206	0.814	**0.862**	0.734	0.842	0.977
BFGS	0.079	0.817	0.791	0.787	0.790	0.946
SANN	-0.367	0.756	0.849	0.138	0.813	0.973
SCE	-0.127	0.828	0.843	0.809	0.830	0.966
DE	-0.159	0.826	0.845	0.782	0.829	0.965
DREAM	-0.177	0.820	0.859	0.775	0.841	0.975

Objective Function = **Viney**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.019	0.824	0.837	0.837	0.835	0.969
BFGS	-0.026	0.825	0.839	0.838	0.836	0.970
NM	-0.021	0.823	0.835	0.835	0.833	0.968
SANN	0.120	0.813	0.795	0.773	0.783	0.952
SCE	-0.020	0.824	0.837	0.837	0.835	0.969
DE	-0.001	0.82	0.836	0.836	0.836	0.968
DREAM	-0.133	0.88	0.843	0.804	0.830	0.971

Objective Function = **BL**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.089	0.809	0.804	0.794	0.796	0.954
BFGS	-0.010	0.822	0.829	0.829	0.828	0.964
NM	0.025	0.821	0.818	0.818	0.816	0.962
SANN	-0.130	0.821	0.860	0.825	0.847	0.976
SCE	0.000	0.821	0.827	0.827	0.827	0.963
DE	-0.082	0.829	0.837	0.826	0.829	0.965

Objective Function = **NSE^3**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.017	0.816	0.823	0.823	0.821	0.961
BFGS	0.216	0.784	0.780	0.695	0.758	0.952
NM	0.125	0.804	0.802	0.778	0.789	0.957
SANN	-0.040	0.829	0.840	0.838	0.836	0.970
SCE	0.047	0.812	0.820	0.817	0.815	0.961
DE	0.022	0.818	0.820	0.819	0.817	0.959

APPENDIX A: GR4J model

Parameter Stability:

Objective Function = **NSE**

	x2	x3	x4	x1	etmult
PORT	-15.000	158.635	1.056	252.297	1
BFGS	-14.998	142.246	1.074	262.663	1
NM	-14.999	165.698	1.049	248.320	1
SANN	-13.712	137.058	1.030	282.453	1
SCE	-14.999	158.633	1.056	252.439	1
DE	-13.529	154.495	1.041	262.426	1
DREAM	-9.724	118.458	1.026	301.581	1

Objective Function = **Viney**

	x2	x3	x4	x1	etmult
PORT	-15.00	199	1.02	231	1
BFGS	-15.00	195	1.03	231	1
NM	-15.00	200	1.01	232	1
SANN	-7.27	140	1.08	274	1
SCE	-15.00	198	1.03	230	1
DE	-13.00	173	1.07	236	1
DREAM	-9.72	118	1.03	302	1

Objective Function = **BL**

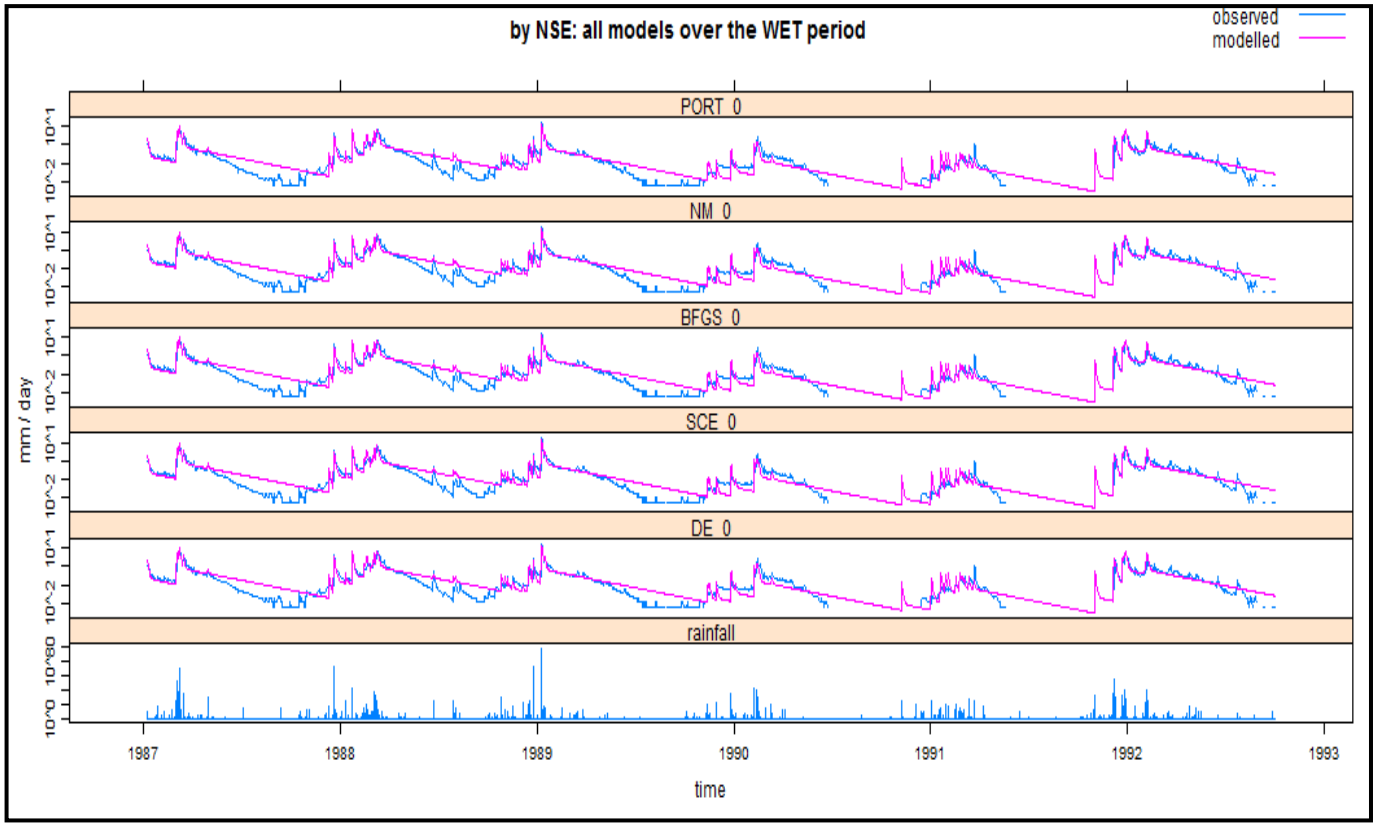
	x2	x3	x4	x1	etmult
PORT	-20.0	216.3	1.10	203	1
BFGS	-16.1	279.3	1.10	211	1
NM	-20.0	236.0	1.10	195	1
SANN	-11.8	76.8	1.28	320	1
SCE	-20.0	231.8	1.10	194	1
DE	-16.0	188.4	1.11	224	1
DREAM	-19.6	179.9	1.10	224	1

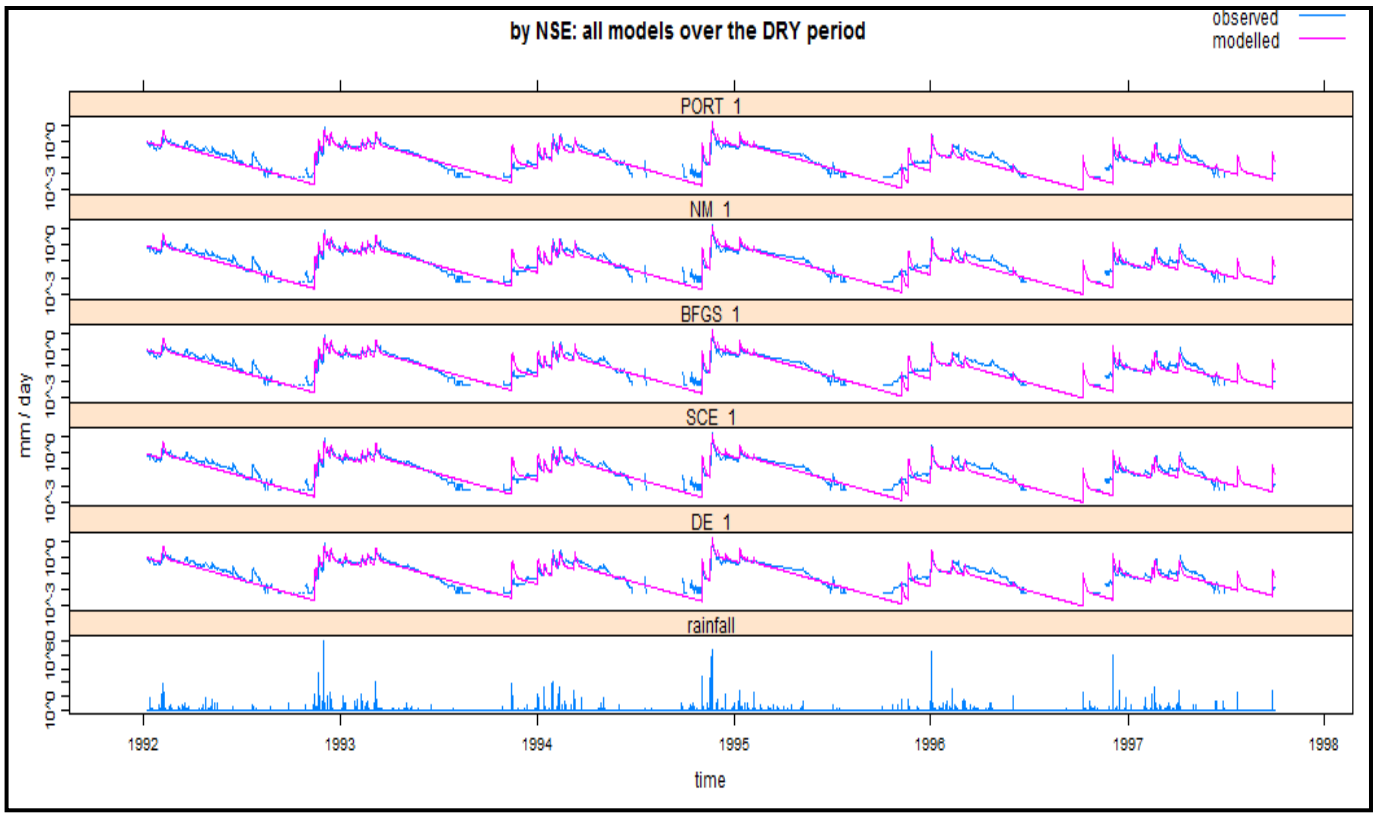
Objective Function = **NSE^3**

	x2	x3	x4	x1	etmult
PORT	-25.000	300.000	1.000	159.501	1
BFGS	-15.716	282.162	1.002	182.173	1
NM	-16.336	254.201	1.000	192.225	1
SANN	-13.197	178.954	1.058	251.173	1
SCE	-24.385	299.928	1.002	154.493	1
DE	-23.314	293.179	1.016	170.050	1

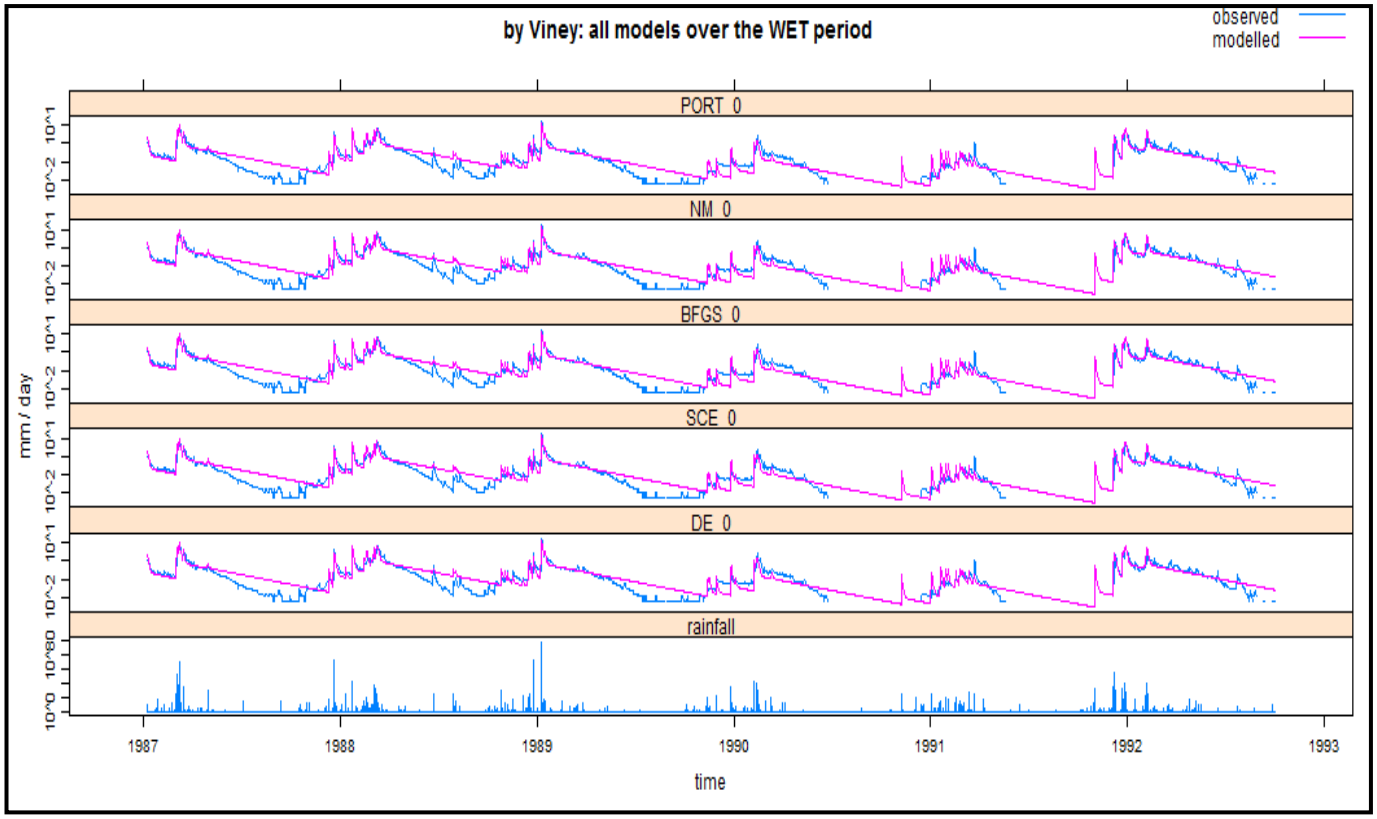
AWBM model

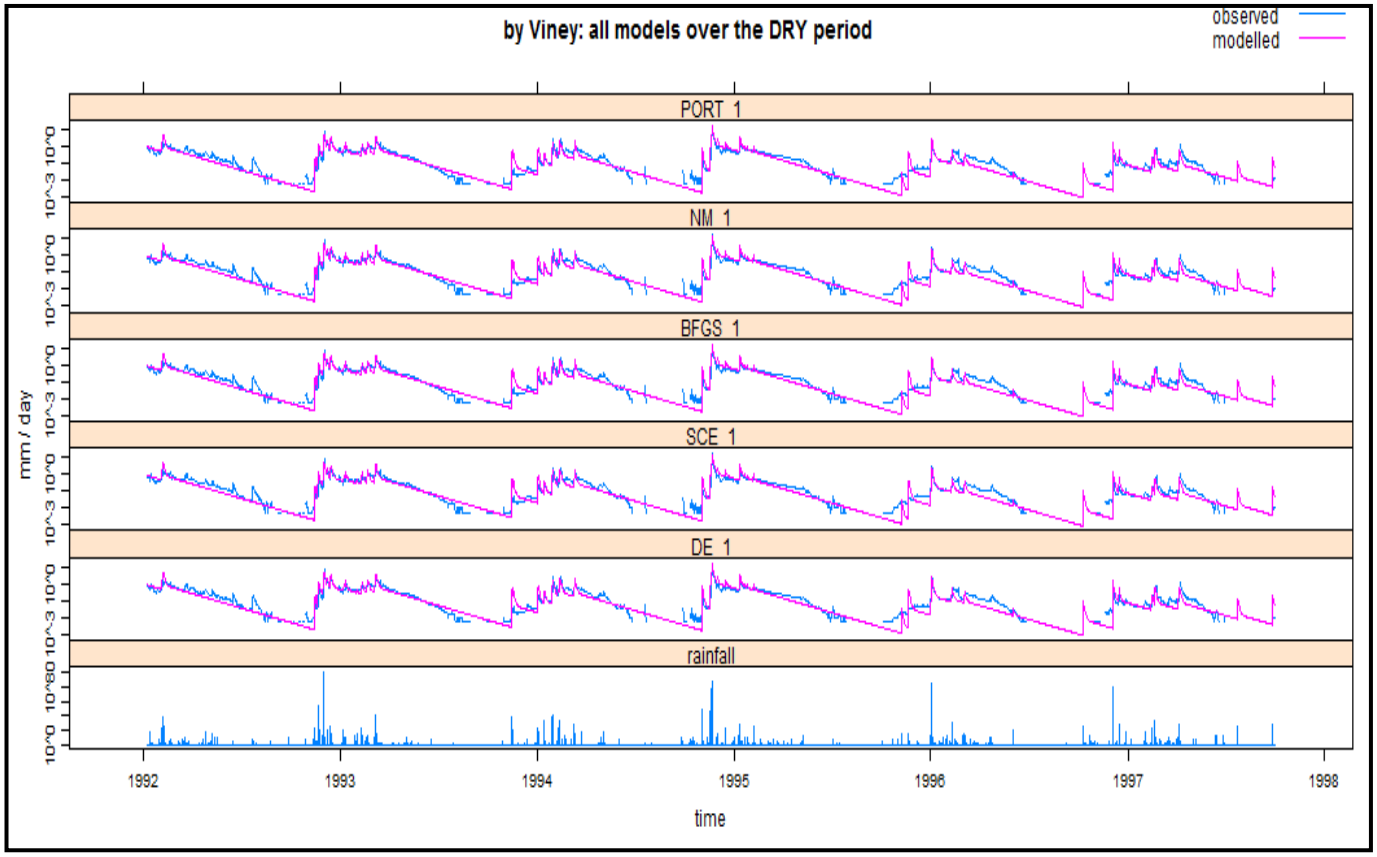
(A) Simulated streamflows (in log scale) – all models



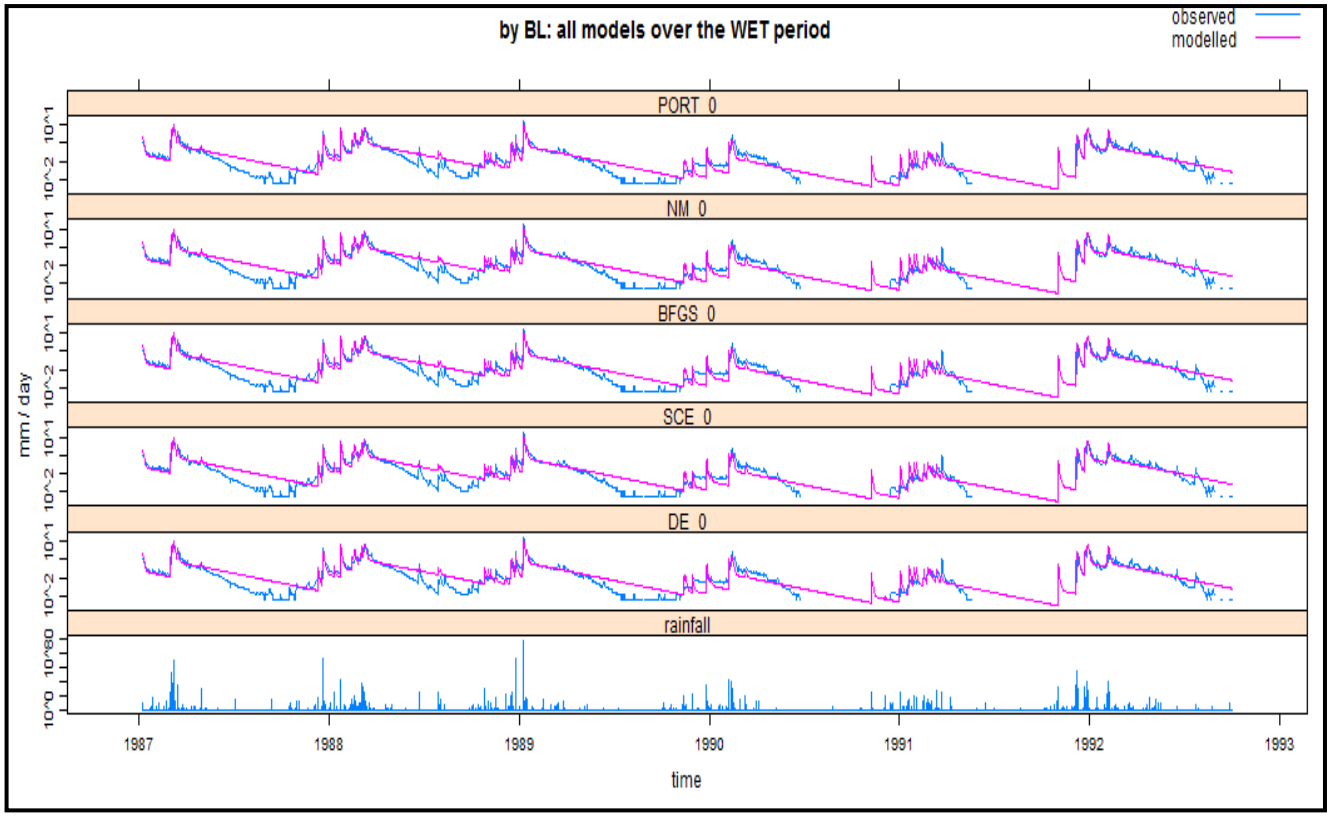


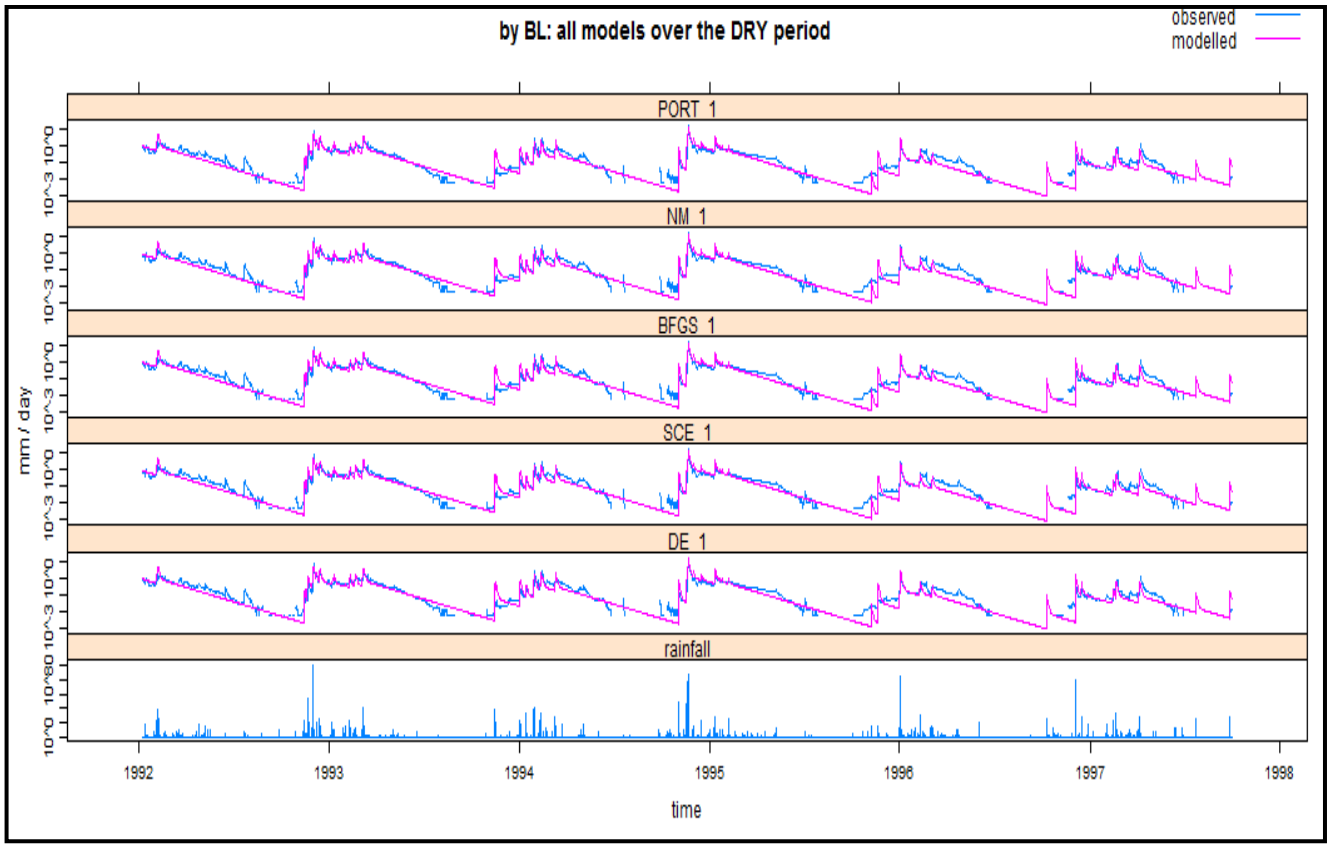
APPENDIX B: AWBM model



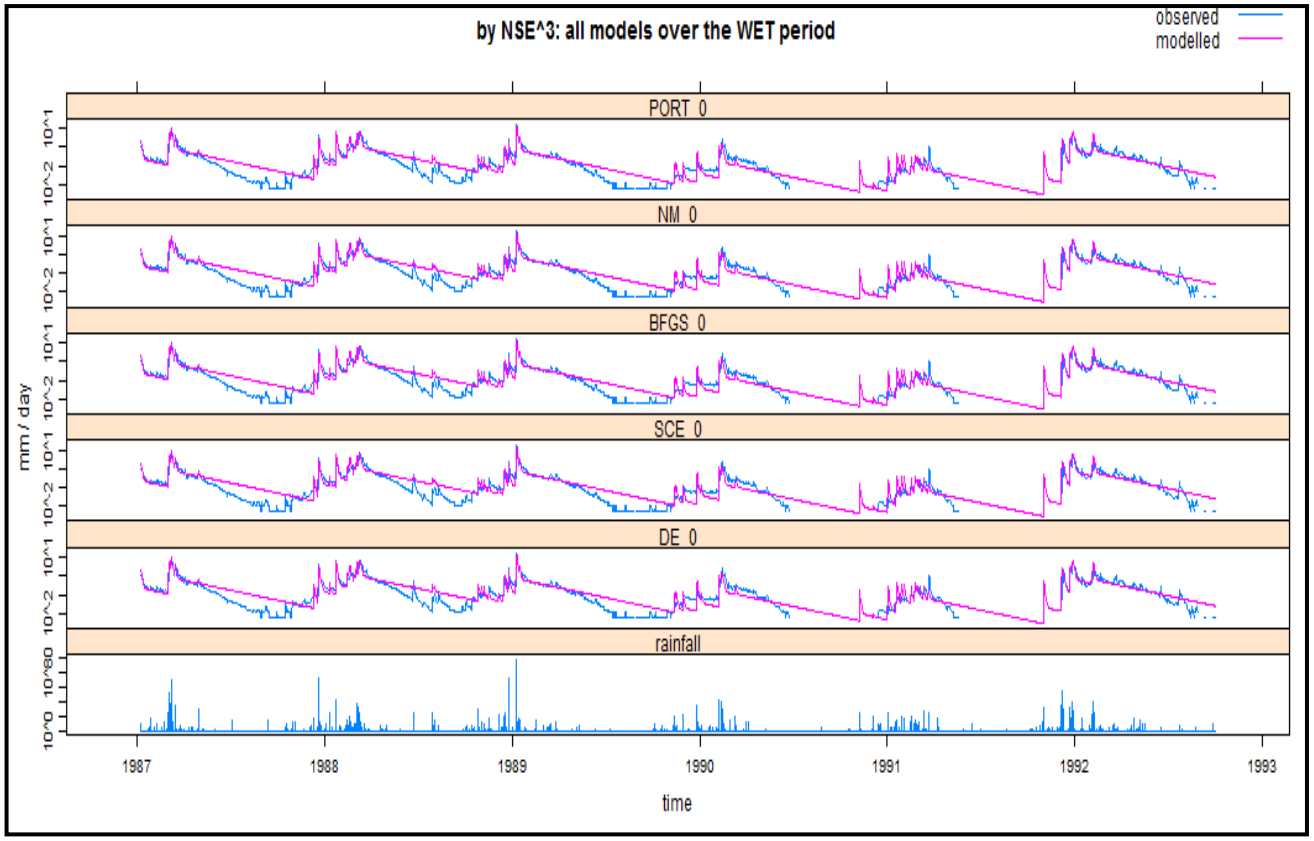


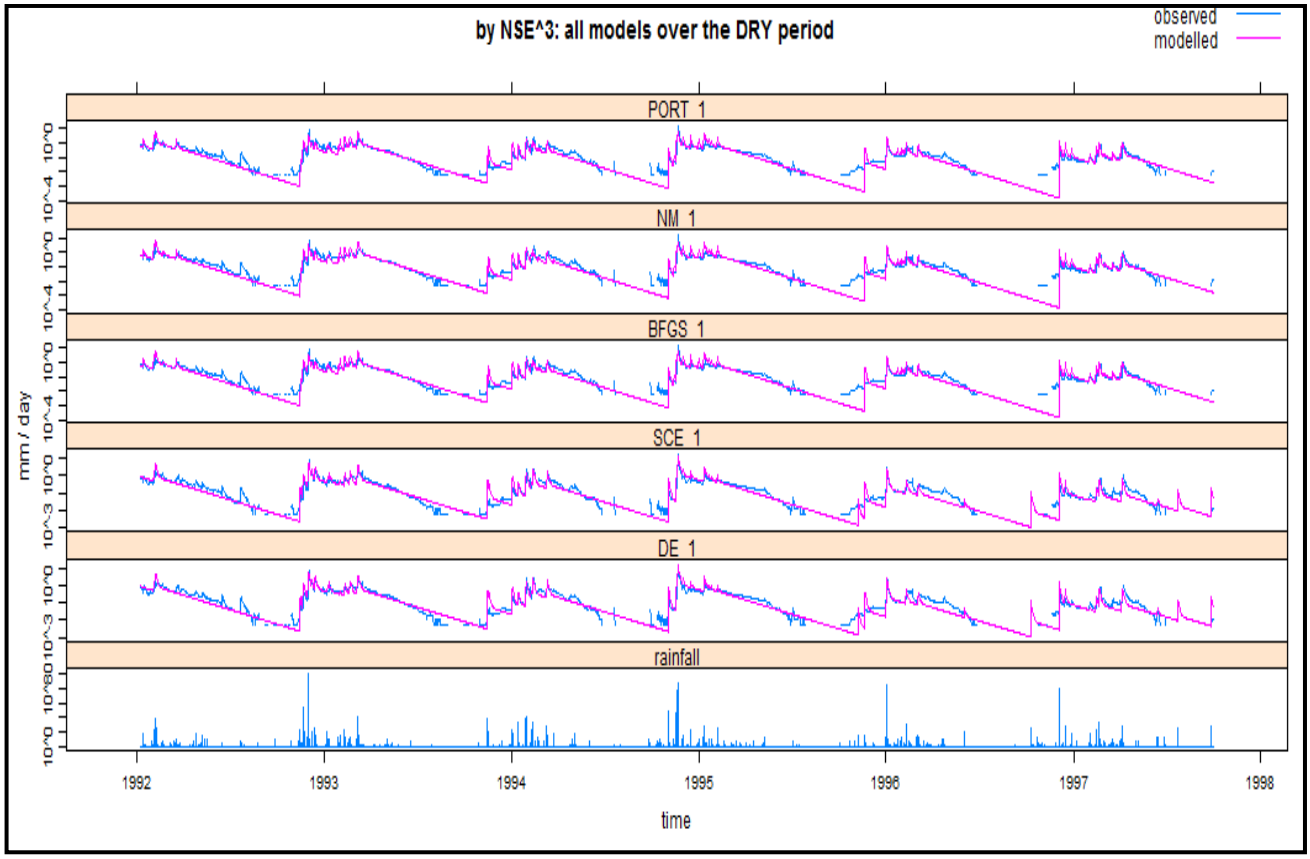
APPENDIX B: AWBM model





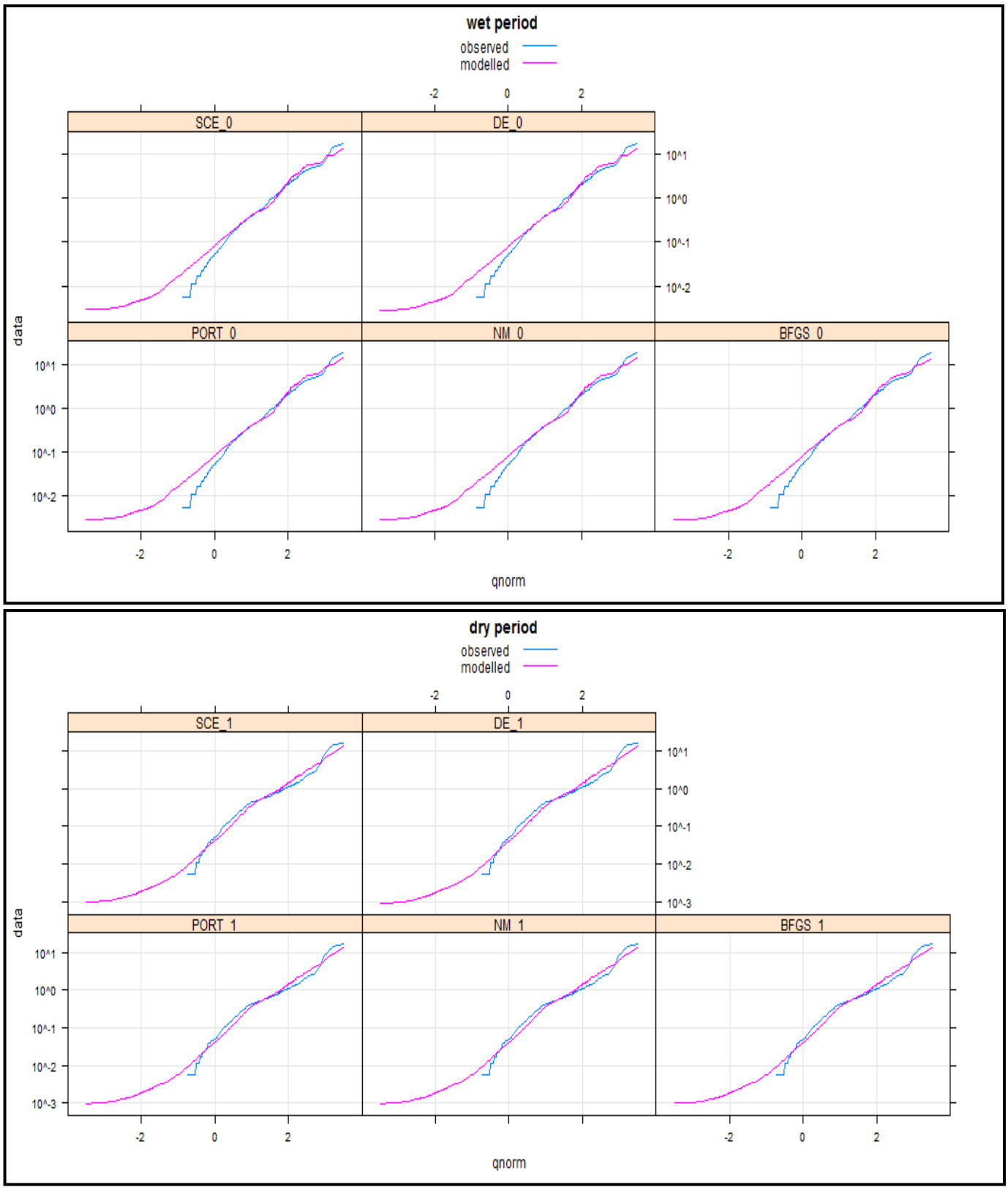
APPENDIX B: AWBM model





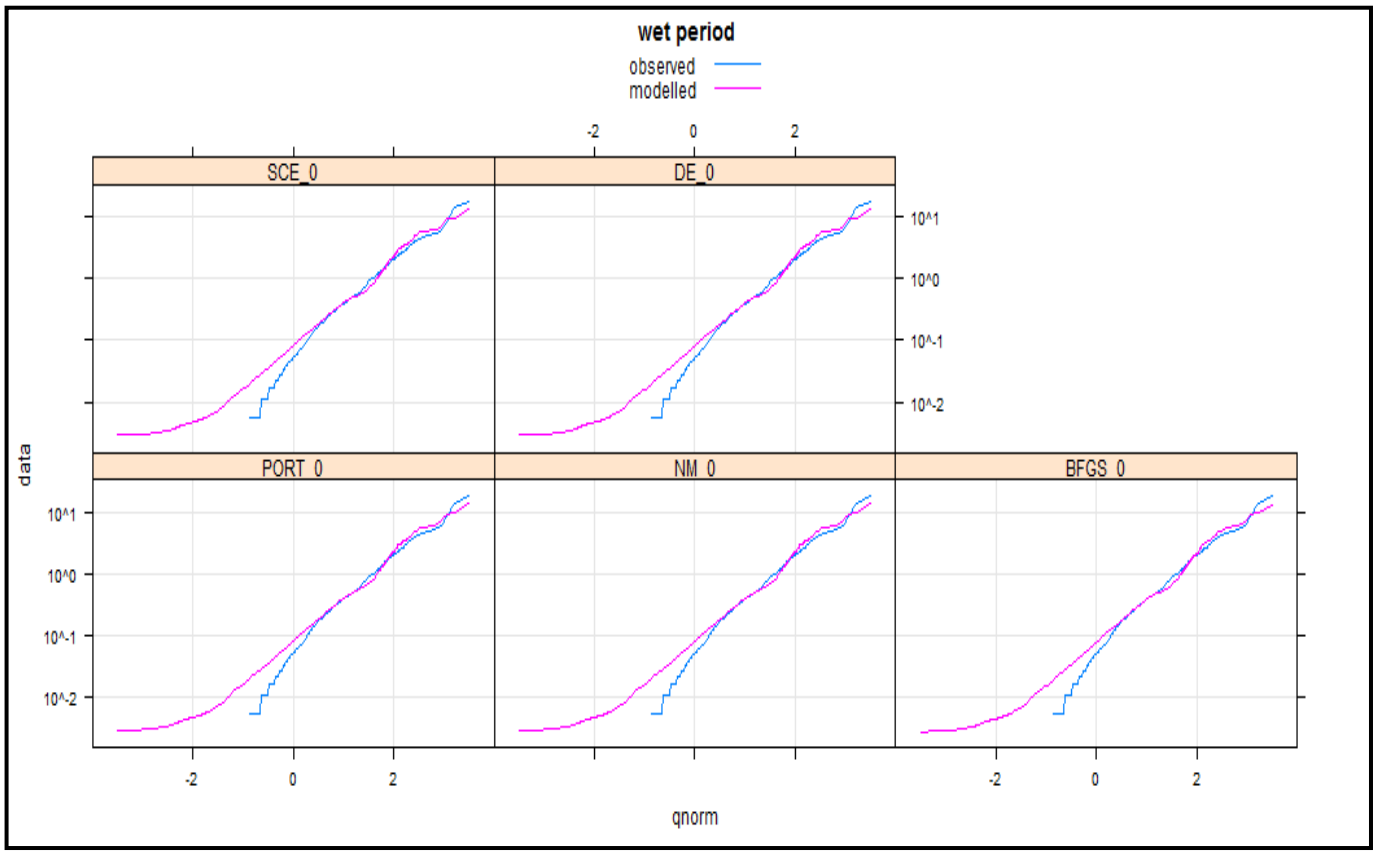
(B) Normal distribution Q-Q plot – all models

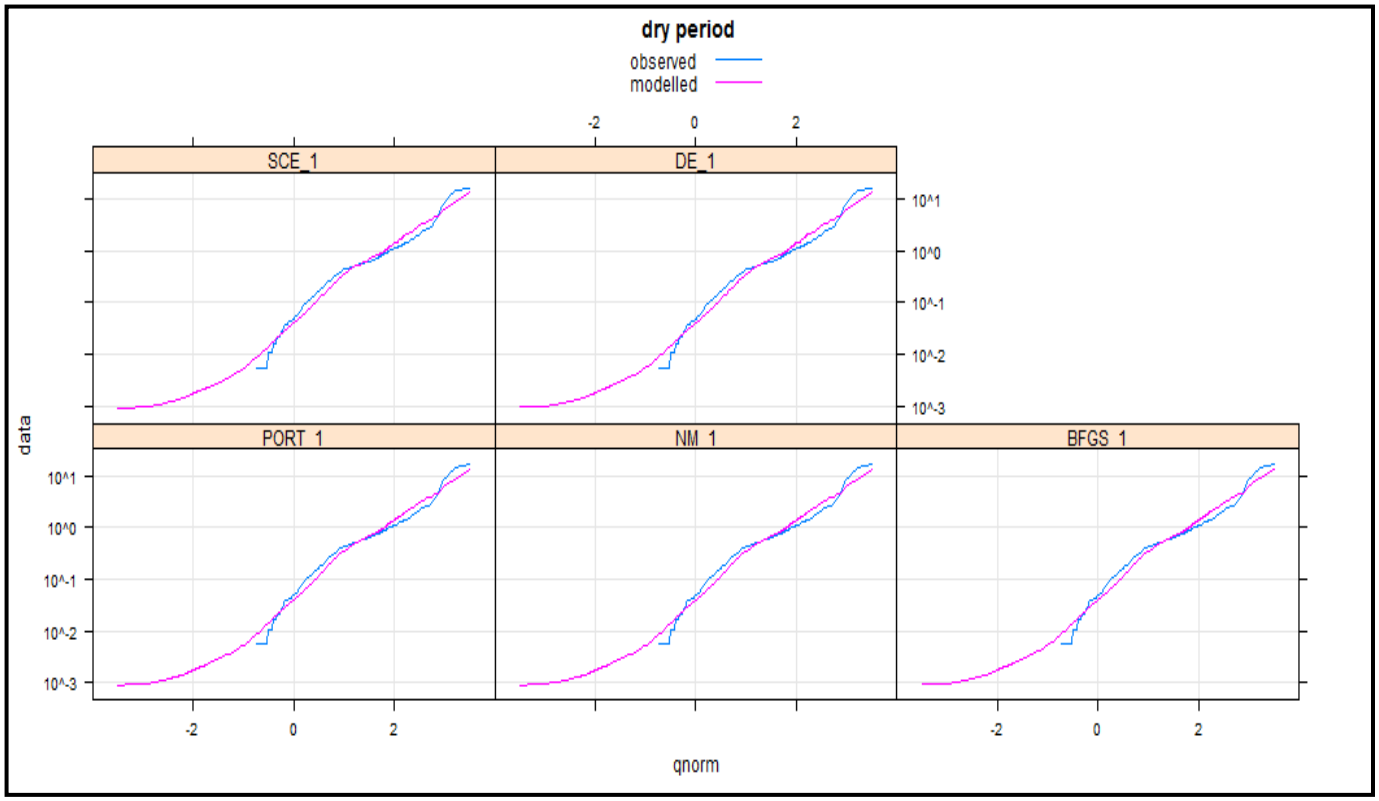
By NSE:



APPENDIX B: AWBM model

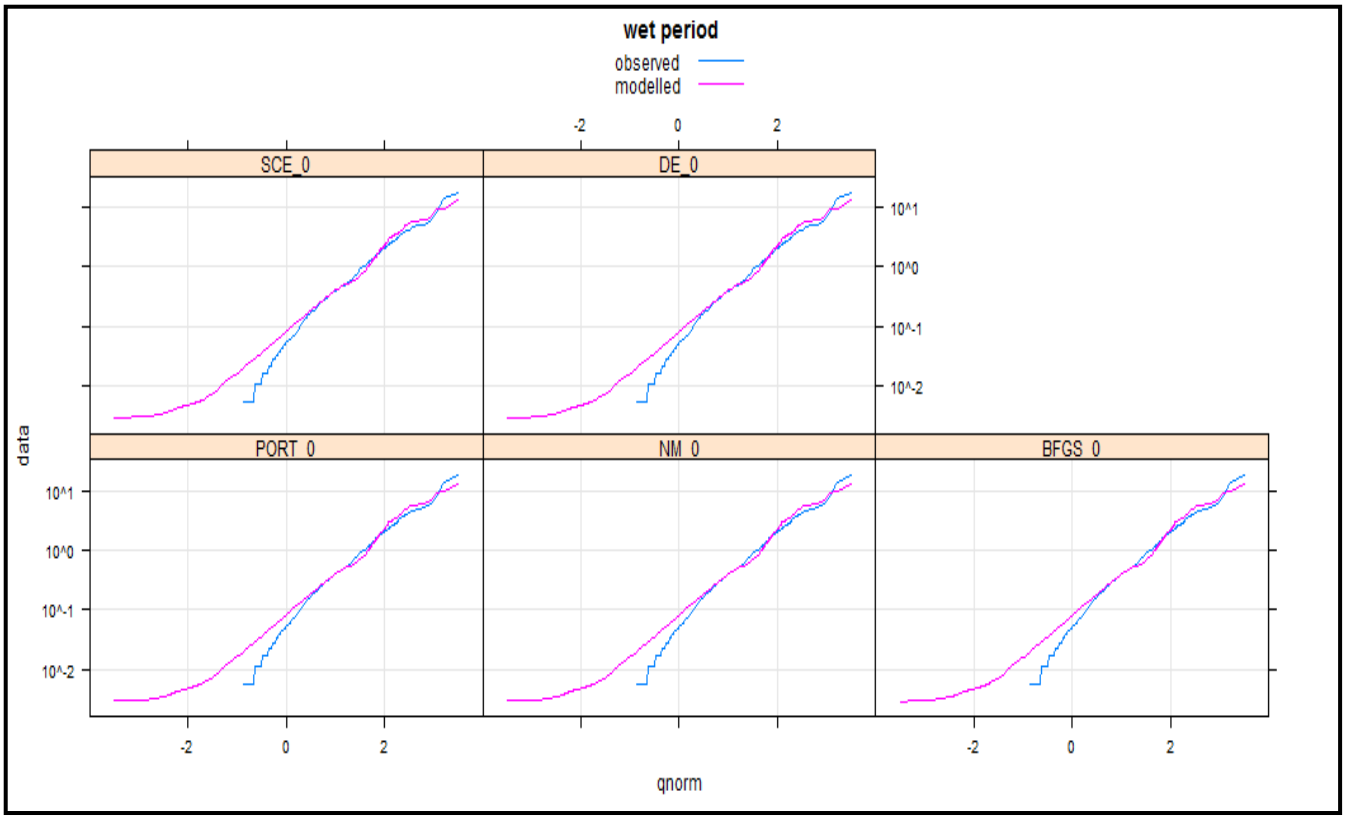
By Viney:

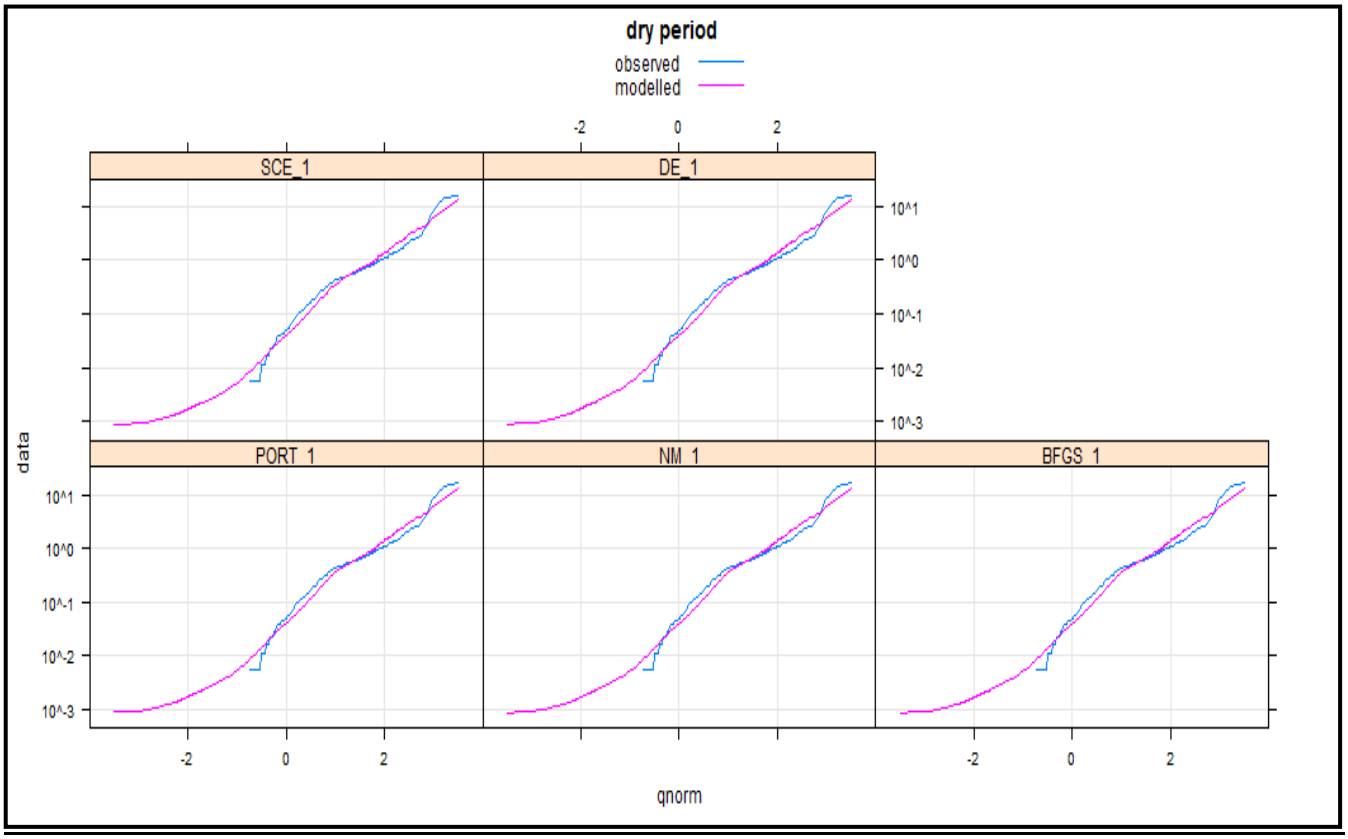




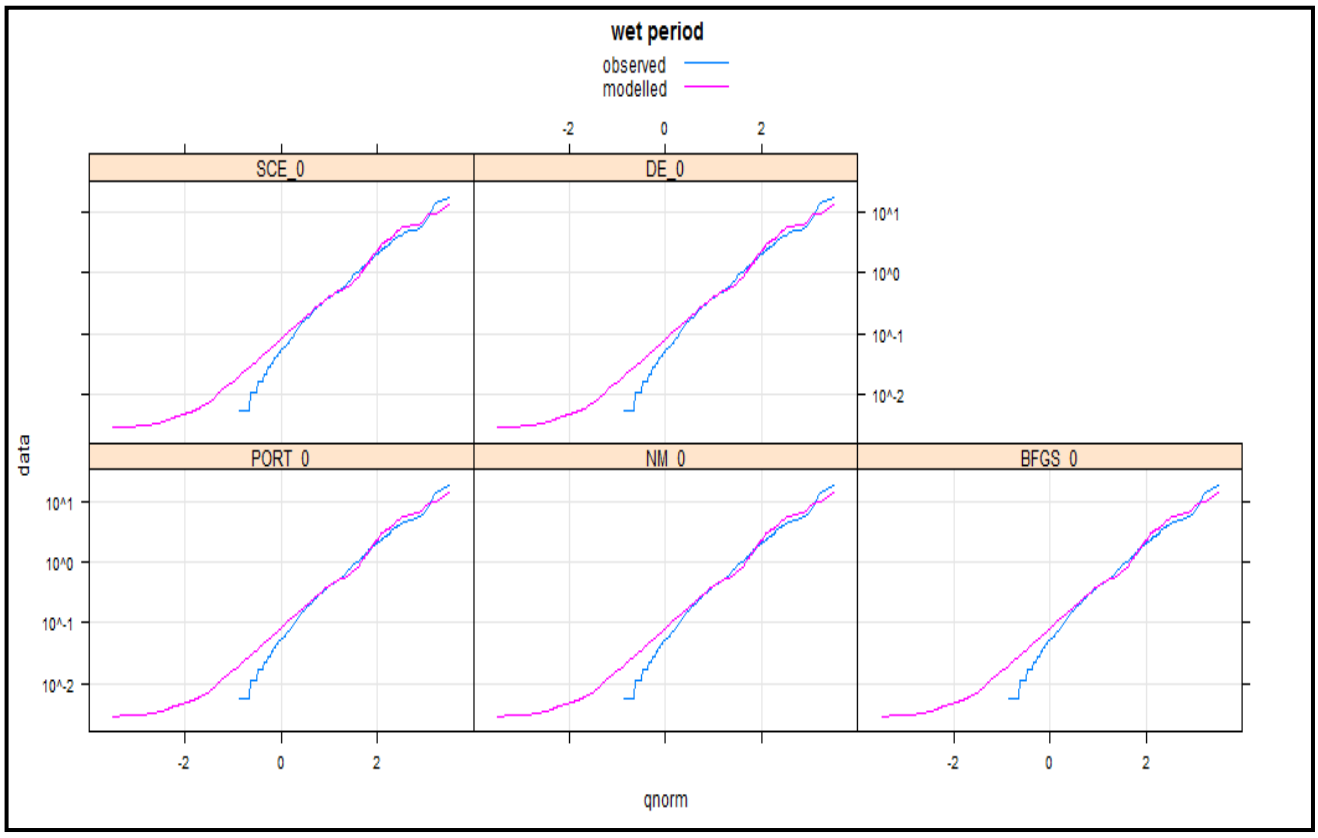
APPENDIX B: AWBM model

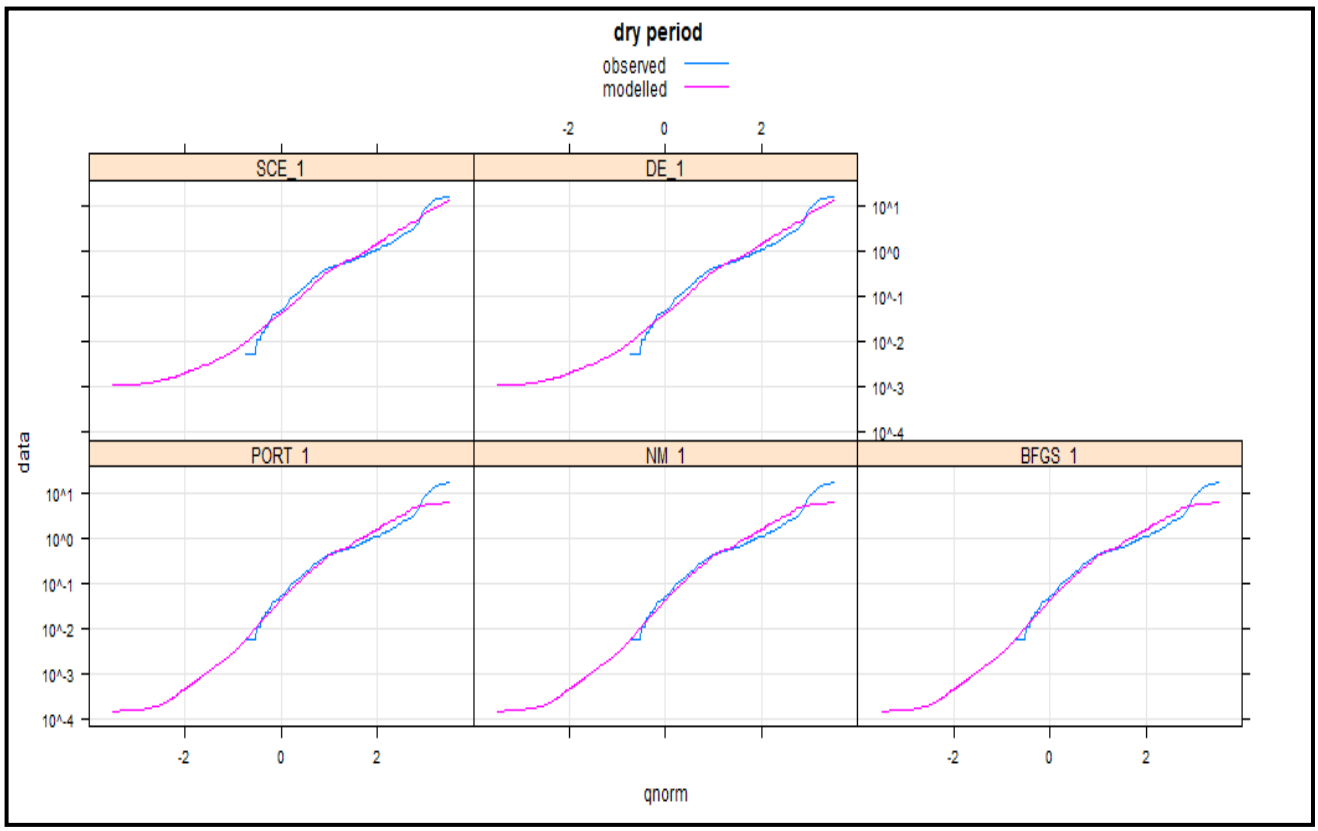
By BL:





By NSE^3:

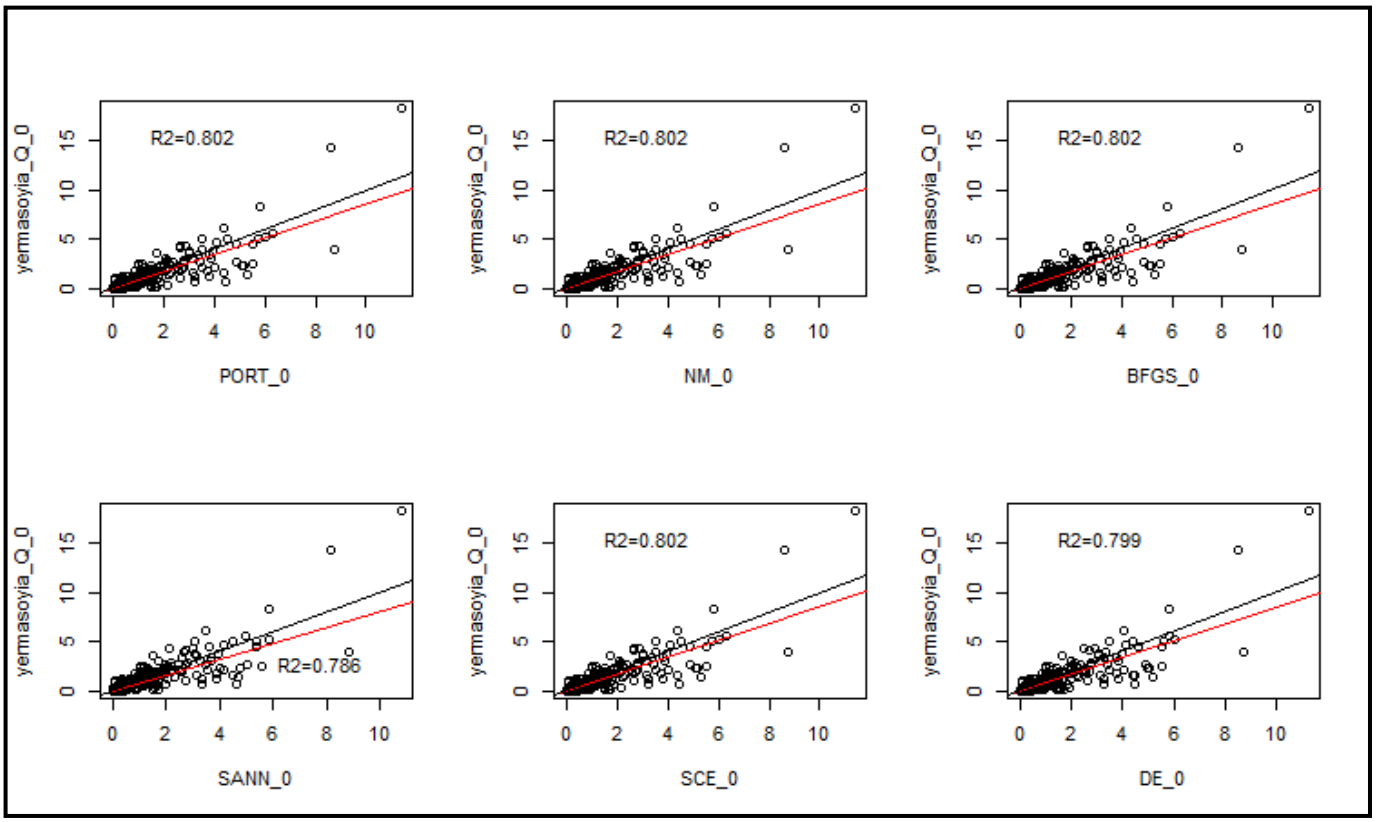




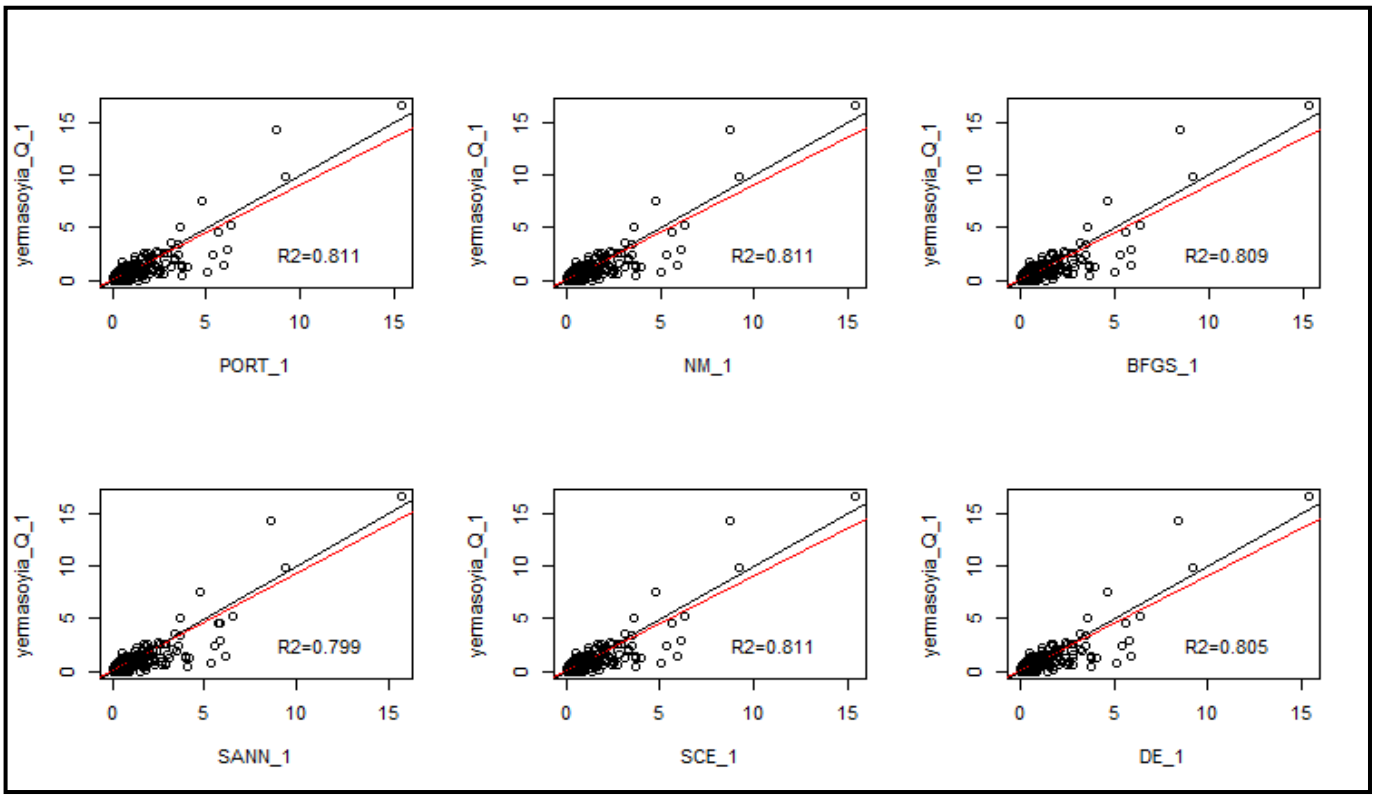
APPENDIX B: AWBM model

(C) Scatterplots

By NSE: wet period

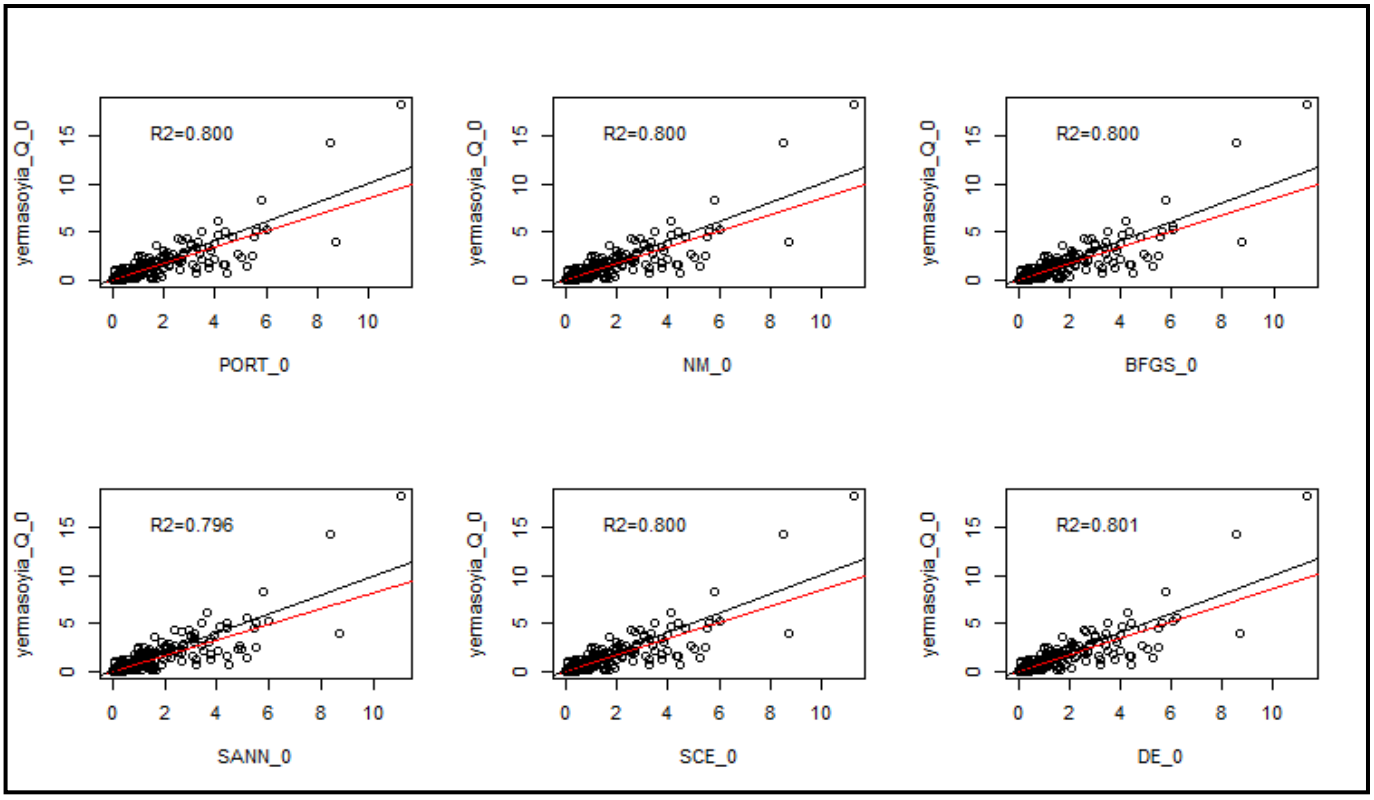


By NSE: dry period

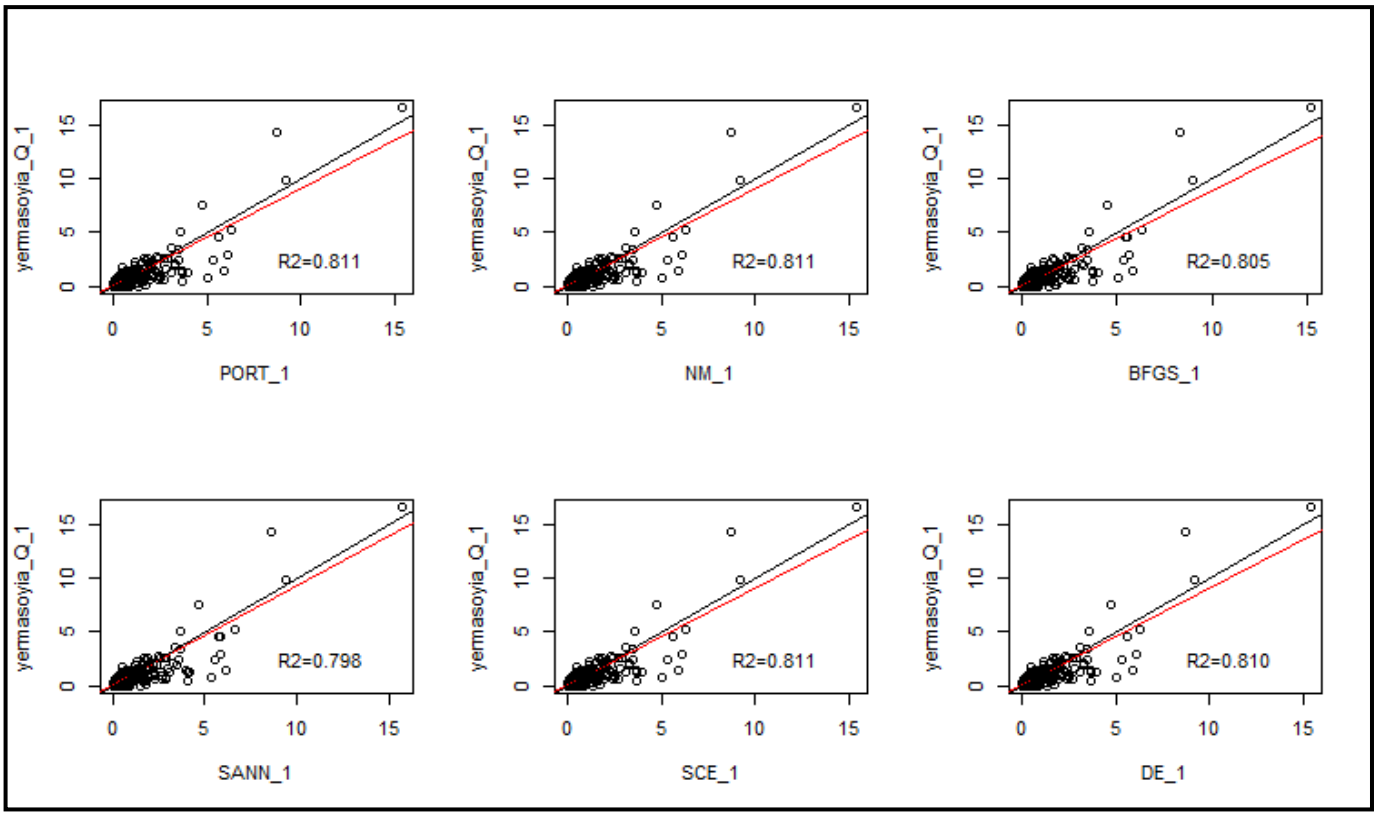


APPENDIX B: AWBM model

By Viney: wet period

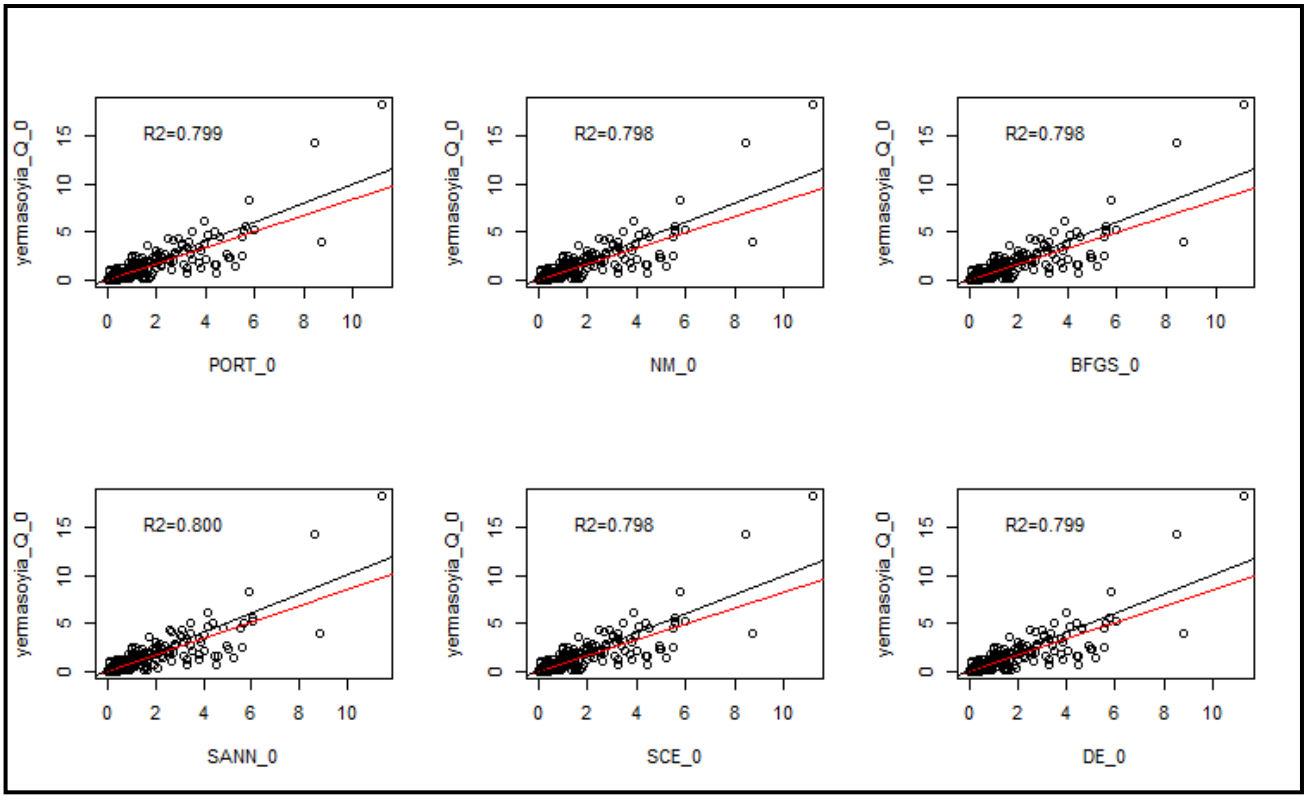


By Viney: dry period

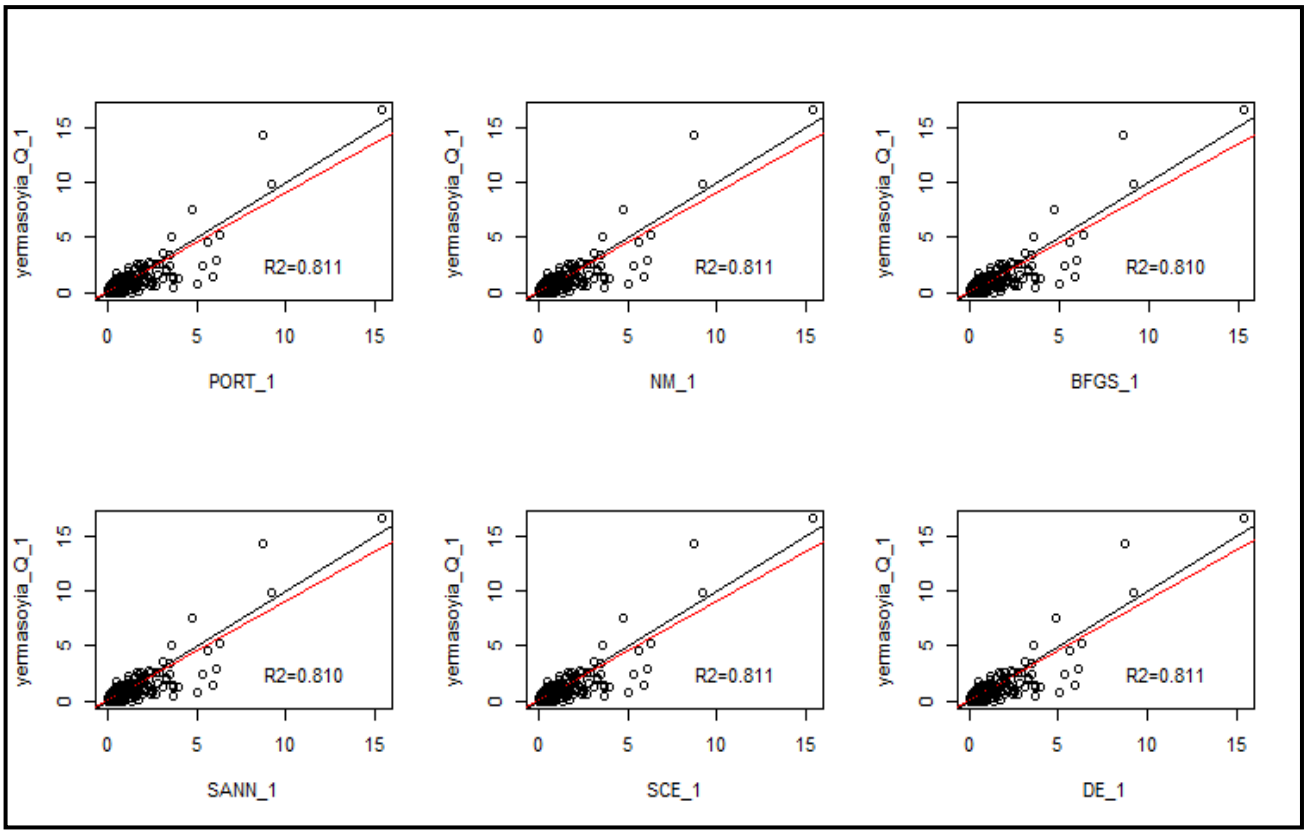


APPENDIX B: AWBM model

By BL: wet period

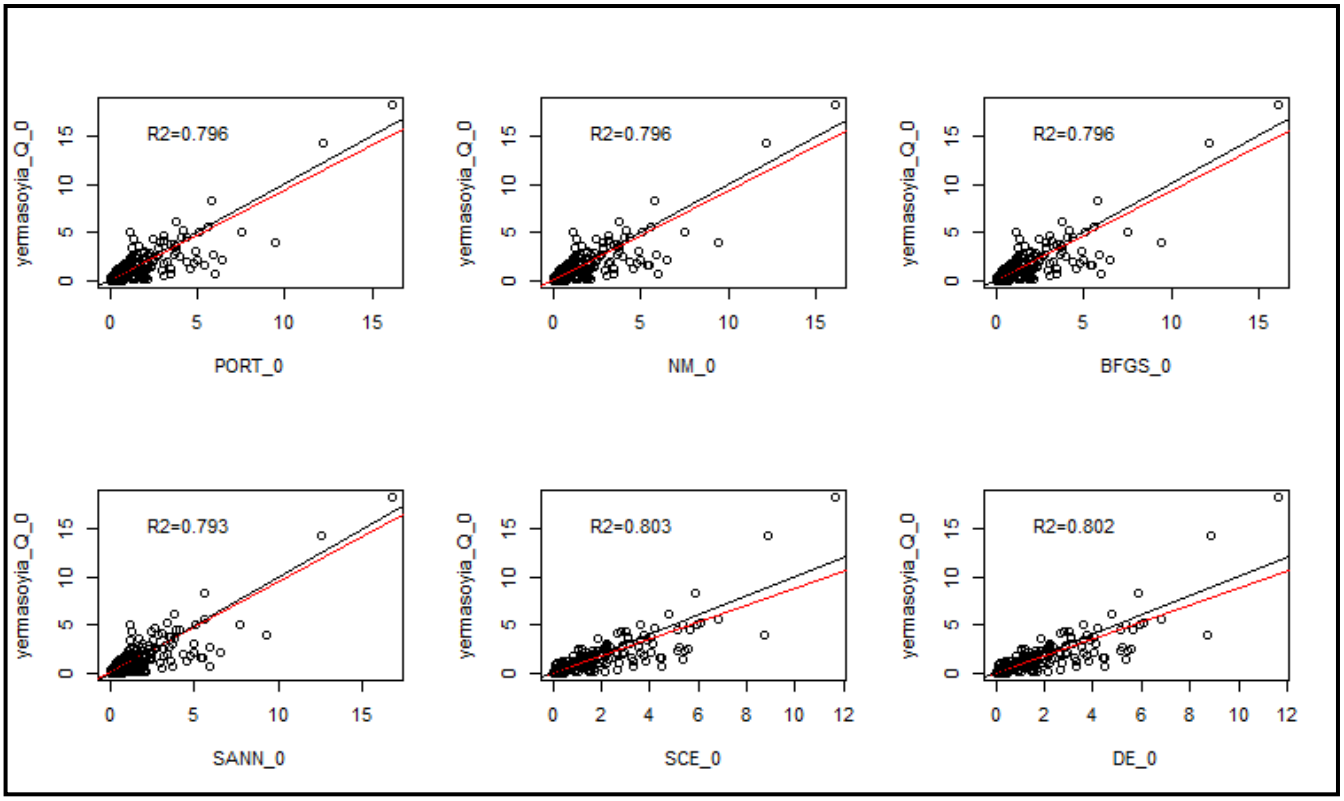


By BL: dry period

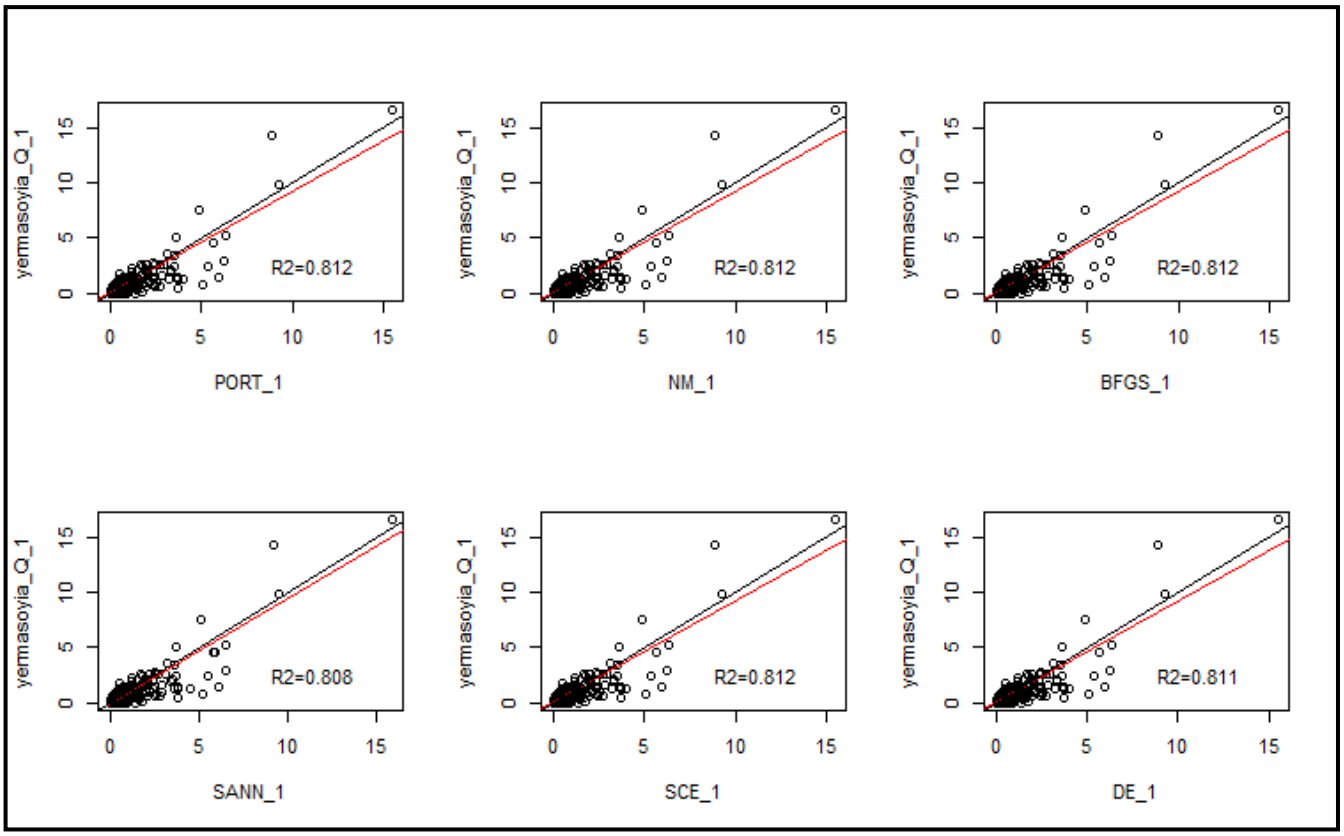


APPENDIX B: AWBM model

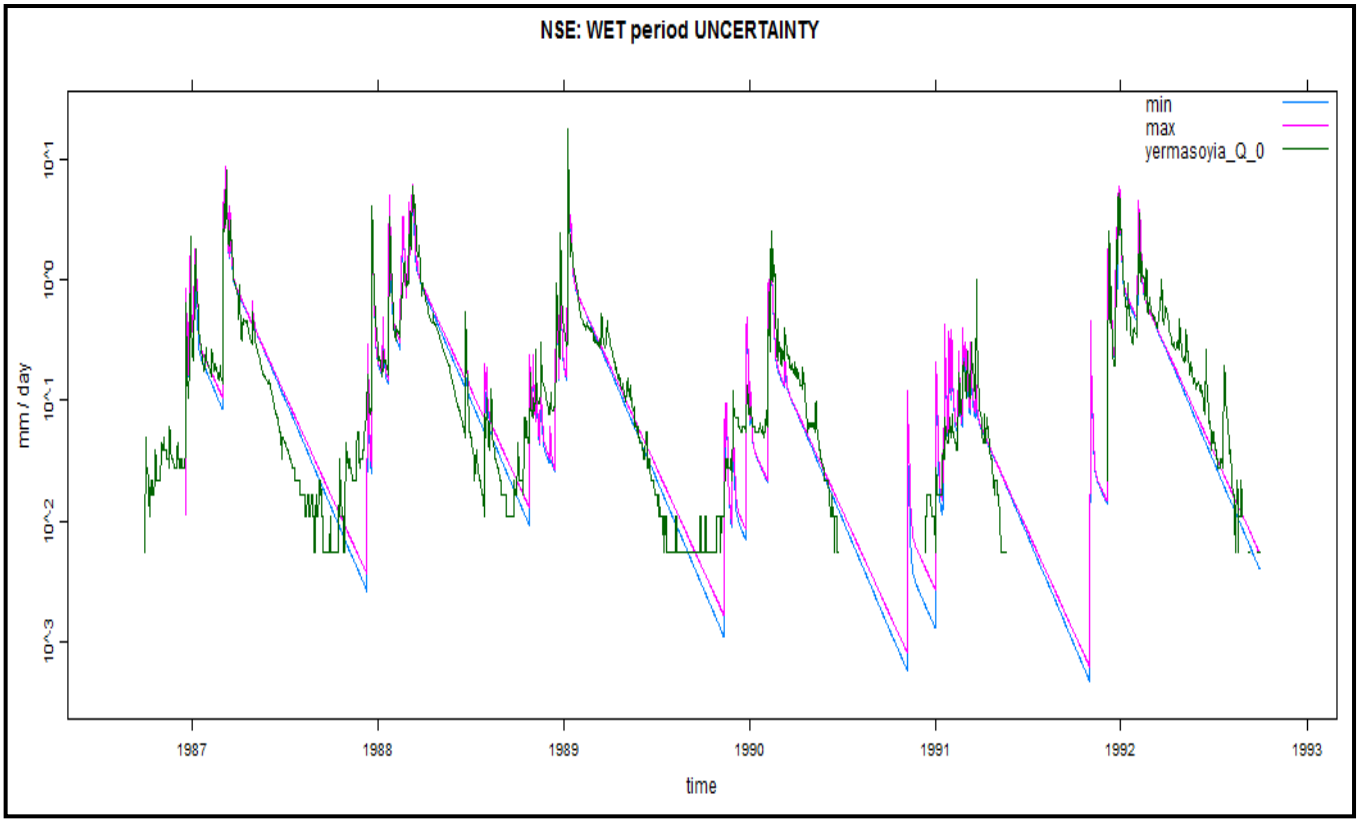
By NSE^3: wet period

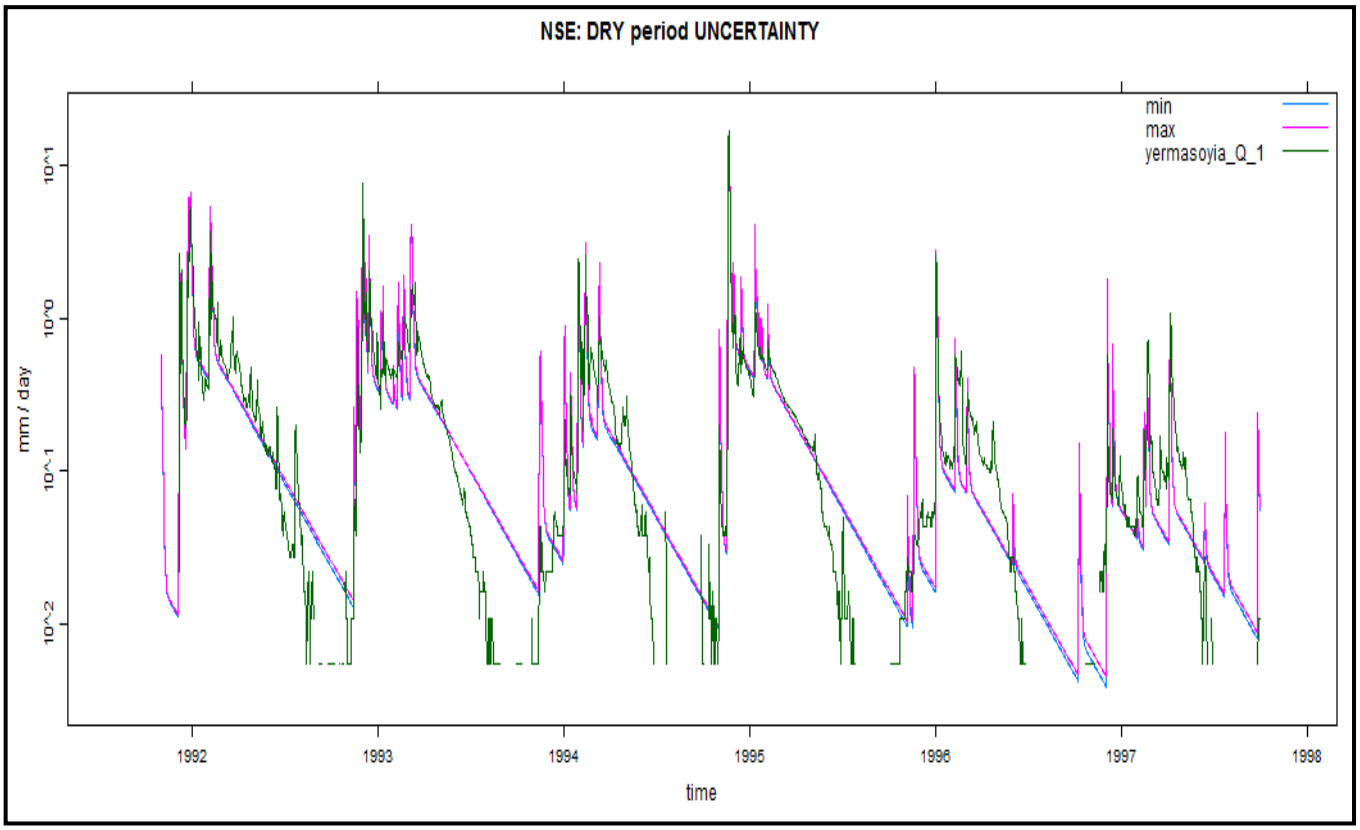


By NSE^3: dry period

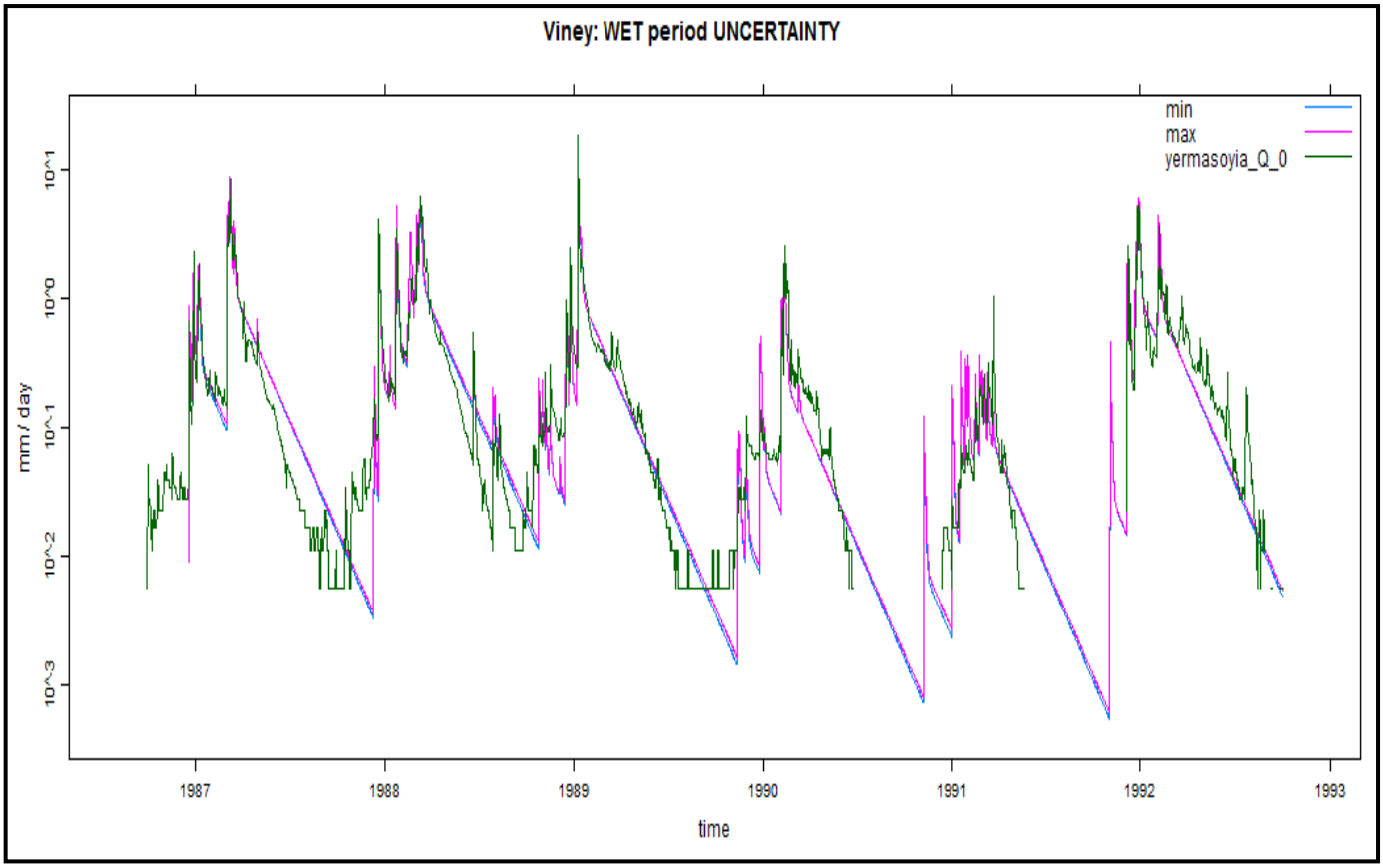


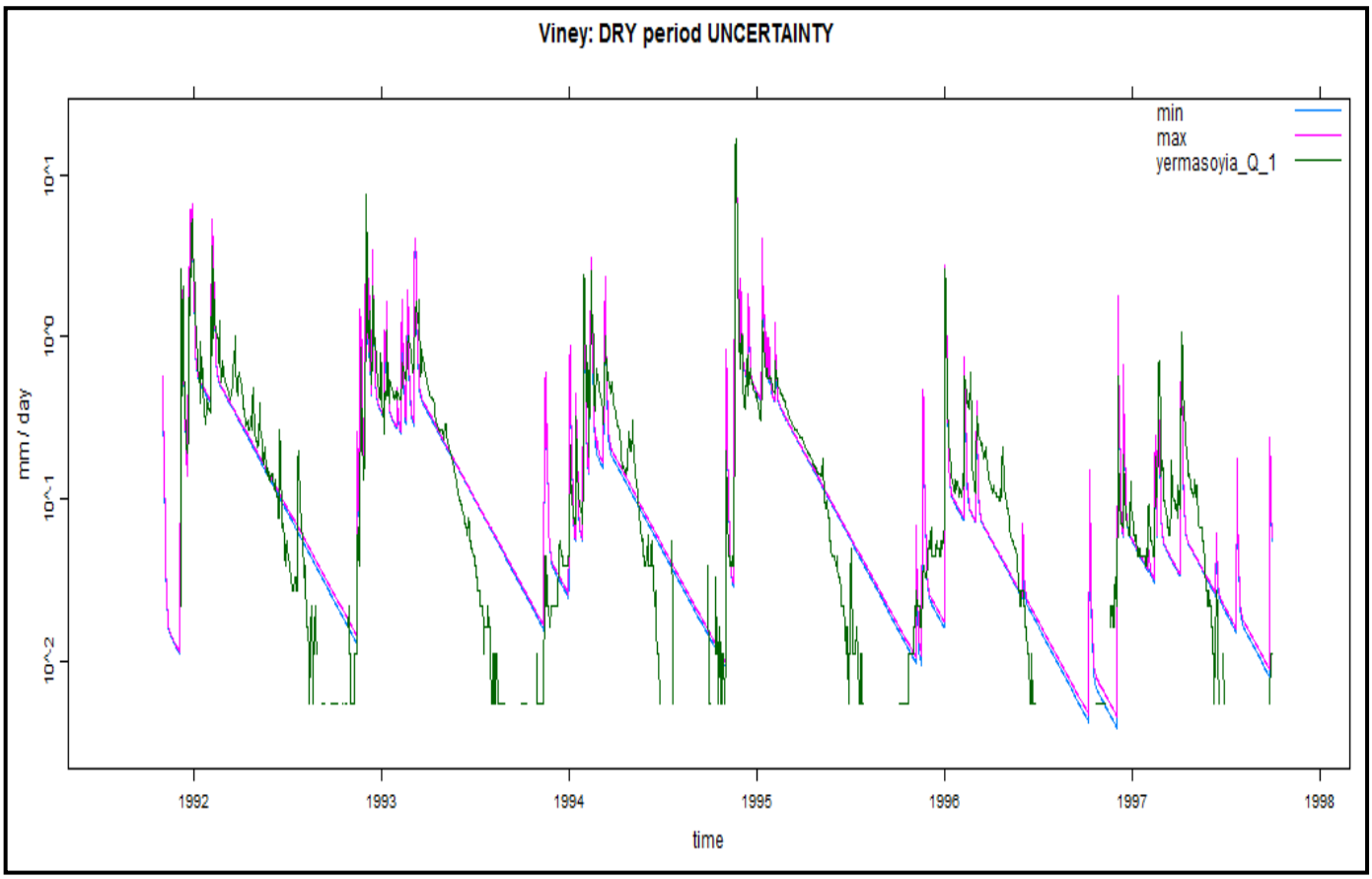
(D) Model structure uncertainty



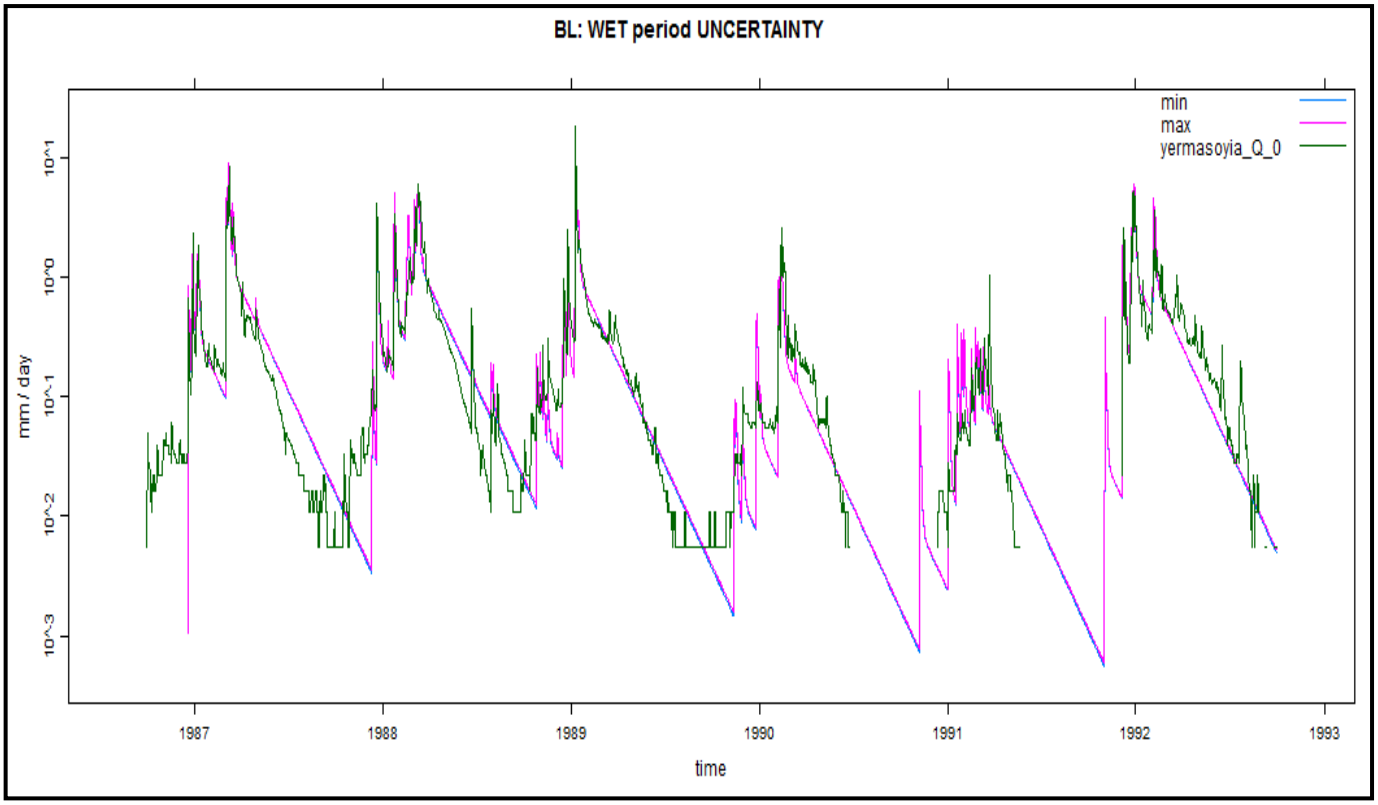


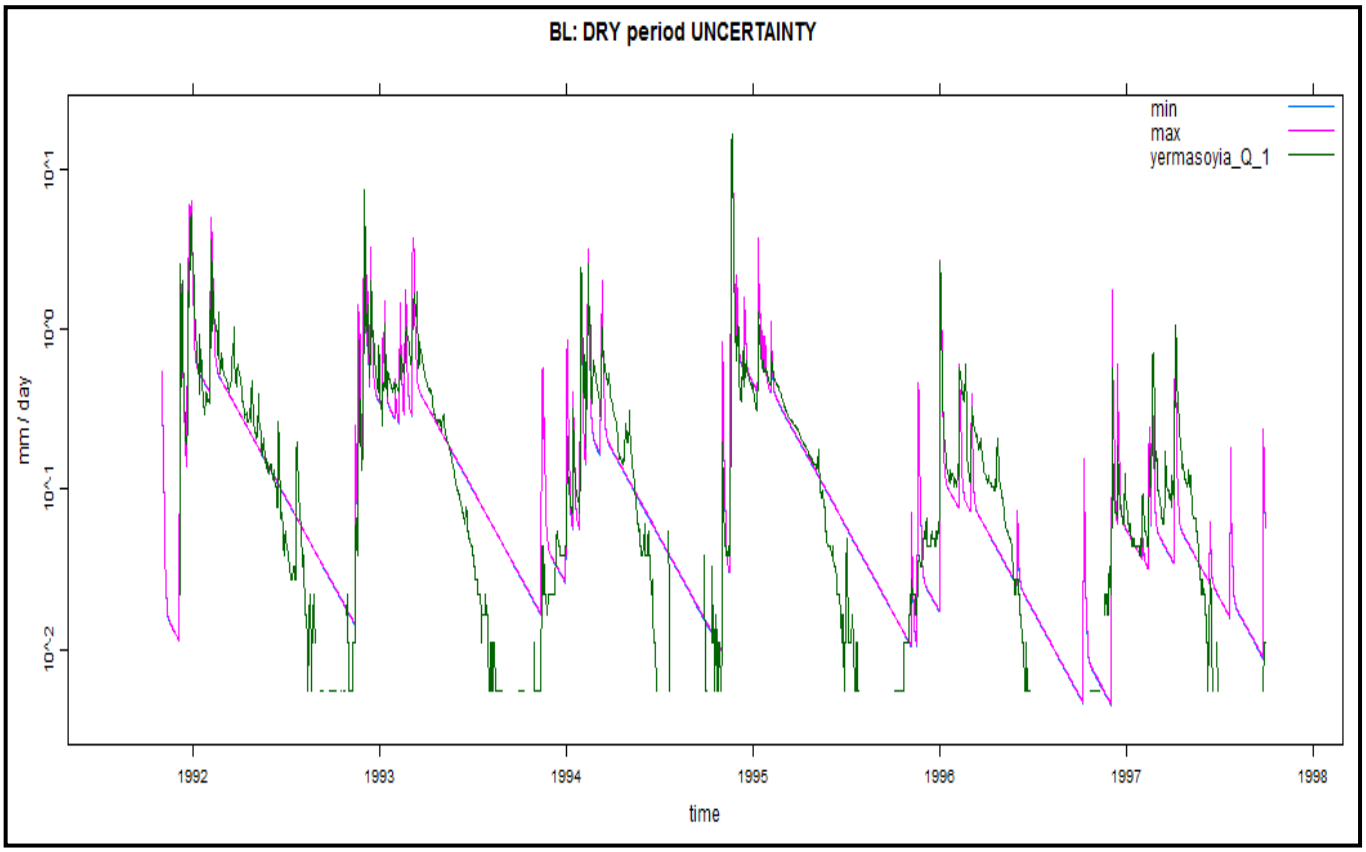
APPENDIX B: AWBM model

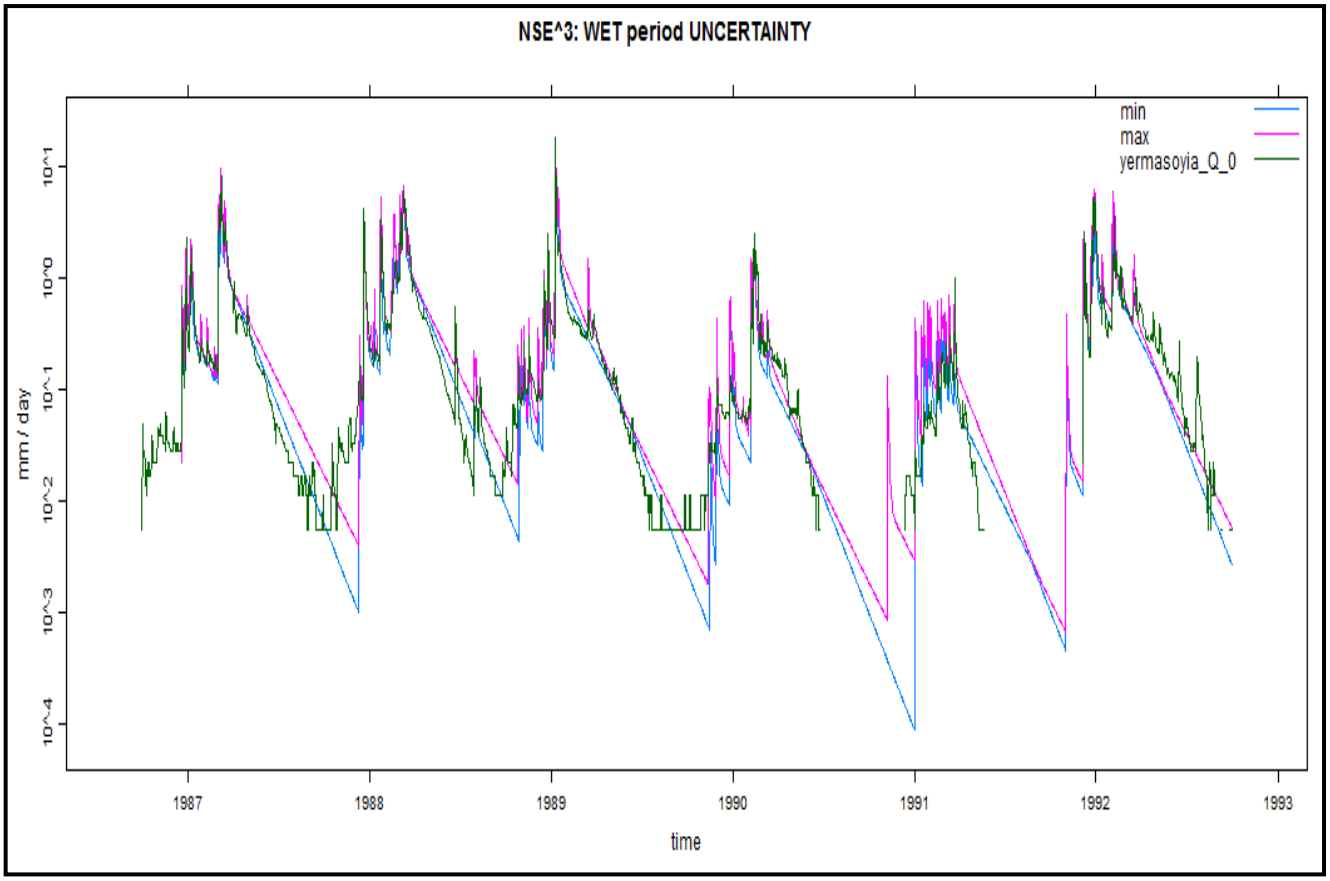


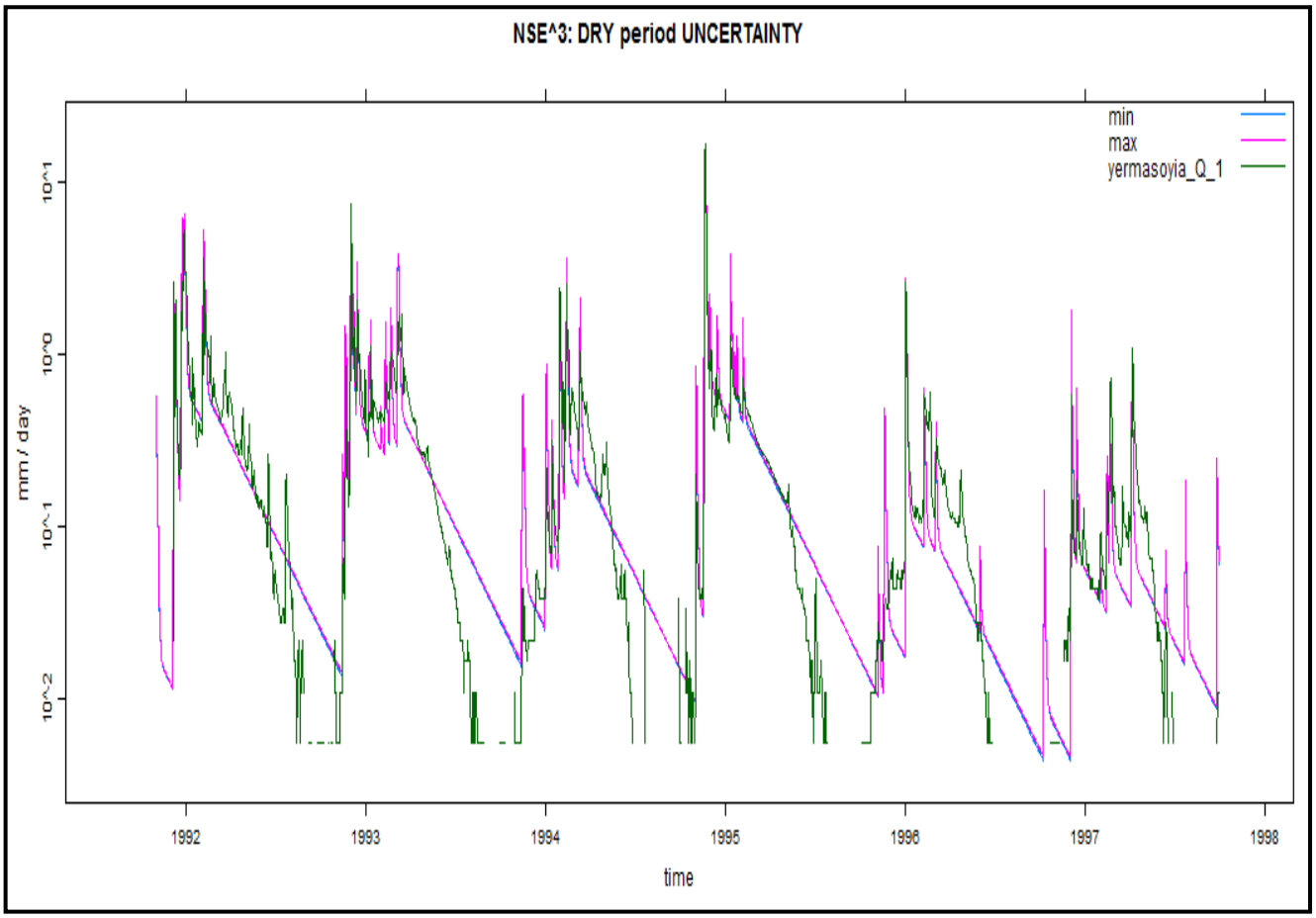


APPENDIX B: AWBM model







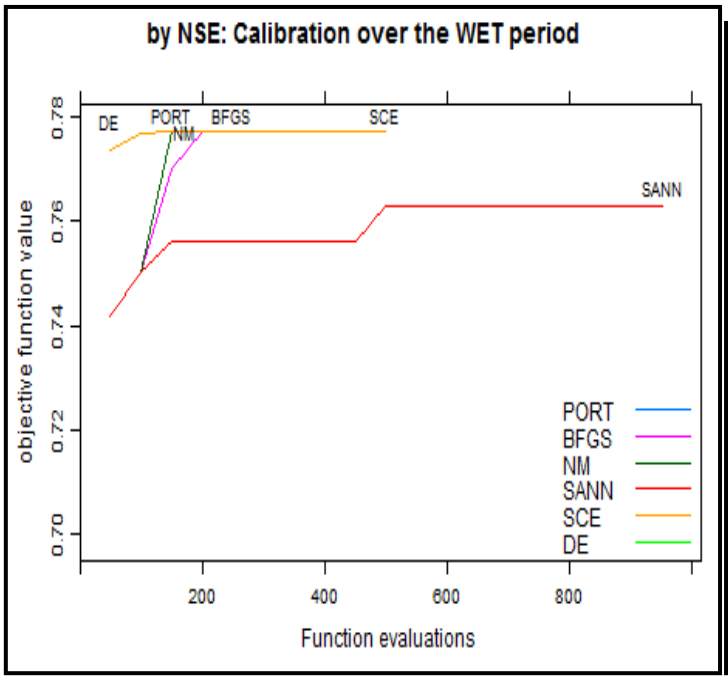


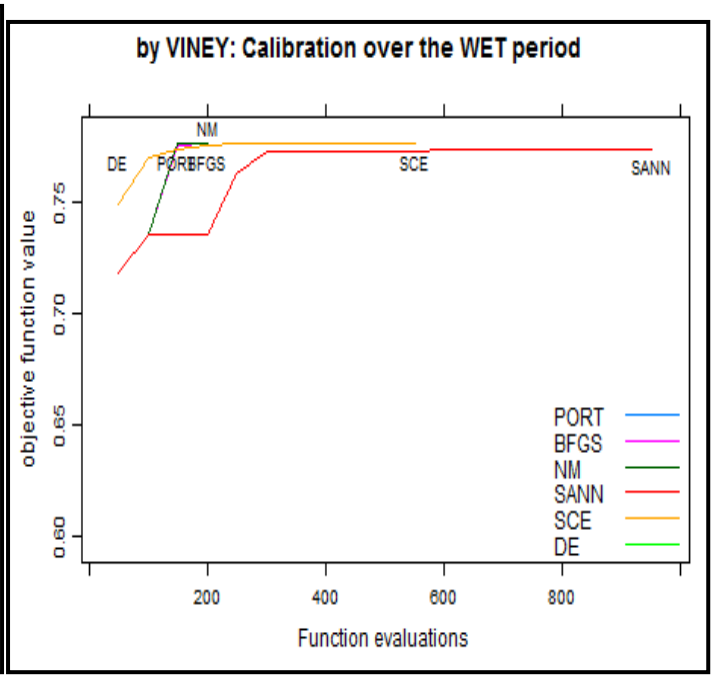
APPENDIX B: AWBM model

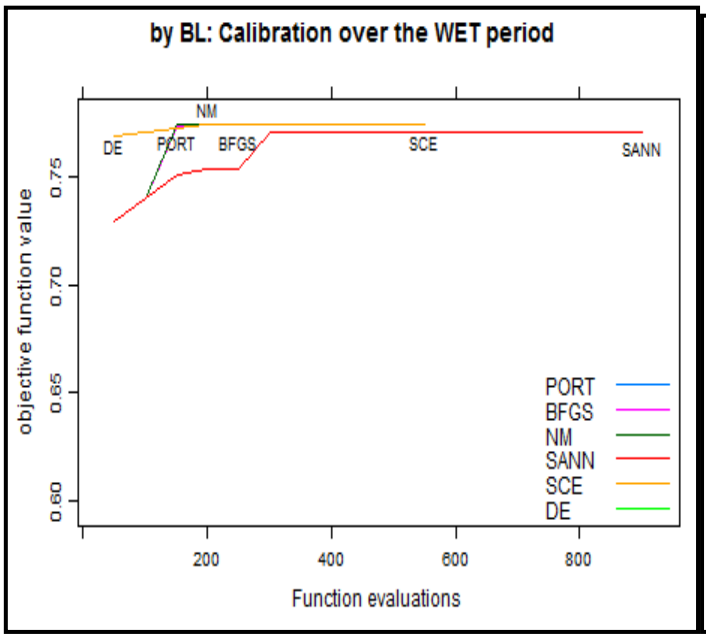
(E) Objective Function and Optimization Algorithm uncertainty

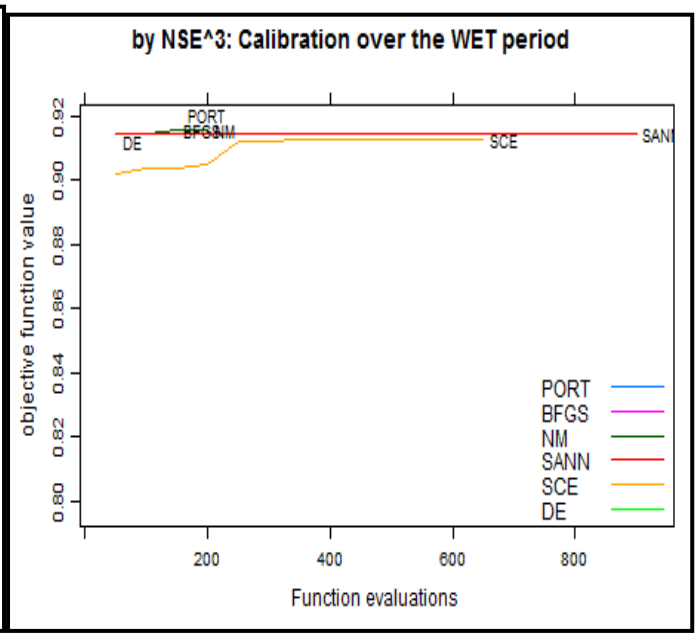
CALIBRATION: **WET** – VALIDATION: **DRY**

Optimization traces:









APPENDIX B: AWBM model

Validation FIT STATISTICS:

Objective Function = **NSE**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.014	0.826	0.782	0.782	0.780	0.930
BFGS	-0.013	0.826	0.782	0.782	0.781	0.930
NM	-0.014	0.826	0.782	0.782	0.780	0.930
SANN	-0.036	0.813	0.715	0.713	0.711	0.875
SCE	-0.014	0.826	0.782	0.782	0.780	0.930
DE	-0.025	0.825	0.774	0.773	0.772	0.923

Objective Function = **Viney**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.026	0.825	0.777	0.777	0.775	0.925
BFGS	-0.022	0.825	0.779	0.778	0.777	0.927
NM	-0.026	0.825	0.777	0.777	0.775	0.925
SANN	-0.042	0.824	0.767	0.765	0.763	0.915
SCE	-0.026	0.825	0.777	0.777	0.775	0.925
DE	-0.018	0.825	0.781	0.780	0.779	0.929

Objective Function = **BL**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.032	0.825	0.774	0.773	0.771	0.922
BFGS	-0.036	0.825	0.773	0.772	0.769	0.920
NM	-0.034	0.825	0.773	0.772	0.770	0.921
SANN	-0.012	0.824	0.773	0.773	0.772	0.924
SCE	-0.035	0.825	0.773	0.772	0.770	0.921
DE	-0.024	0.825	0.773	0.772	0.770	0.922

Objective Function = **NSE^3**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.039	0.734	0.371	0.370	0.367	0.604
BFGS	0.039	0.734	0.371	0.370	0.367	0.604
NM	0.040	0.734	0.372	0.370	0.368	0.605
SANN	0.017	0.732	0.361	0.361	0.359	0.584
SCE	0.015	0.825	0.785	0.784	0.783	0.936
DE	0.015	0.825	0.785	0.785	0.783	0.936

Parameter Stability:

Objective Function = **NSE**

	cap.ave	etmult	tau_s	tau_q	v_s	v_q	delay
PORT	272.790	1.000	46.845	2.126	0.548	0.452	0
BFGS	272.621	1.000	46.861	2.128	0.547	0.453	0
NM	272.829	1.000	46.843	2.126	0.548	0.452	0
SANN	309.238	0.892	43.831	2.065	0.558	0.442	0
SCE	272.791	1.000	46.845	2.126	0.548	0.452	0
DE	278.833	0.990	46.408	2.103	0.550	0.450	0

Objective Function = **Viney**

	cap.ave	etmult	tau_s	tau_q	v_s	v_q	delay
PORT	276.683	1.000	46.494	2.105	0.551	0.449	0
BFGS	275.414	1.000	46.629	2.112	0.549	0.451	0
NM	276.572	1.000	46.506	2.105	0.550	0.450	0
SANN	283.632	0.994	45.812	2.077	0.556	0.444	0
SCE	276.681	1.000	46.495	2.105	0.551	0.449	0
DE	273.941	1.000	46.772	2.120	0.548	0.452	0

Objective Function = **BL**

	cap.ave	etmult	tau_s	tau_q	v_s	v_q	delay
PORT	278.810	1.000	46.265	2.093	0.552	0.448	0
BFGS	279.864	1.000	46.149	2.088	0.553	0.447	0
NM	279.501	1.000	46.188	2.090	0.553	0.447	0
SANN	278.986	0.972	46.671	2.121	0.546	0.454	0
SCE	279.540	1.000	46.184	2.089	0.553	0.447	0
DE	279.755	0.986	46.378	2.103	0.550	0.450	0

Objective Function = **NSE^3**

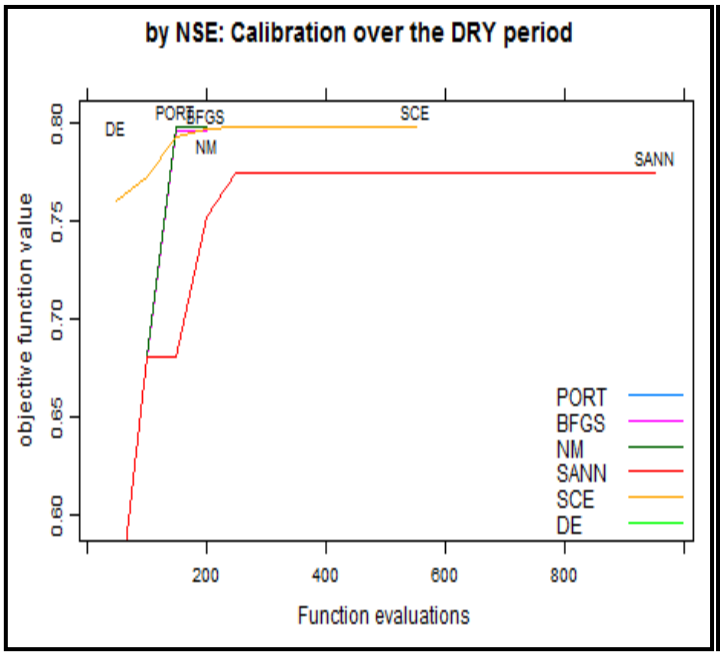
	cap.ave	etmult	tau_s	tau_q	v_s	v_q	delay
PORT	472.740	0.416	38.942	1.779	0.544	0.456	0
BFGS	472.788	0.416	38.941	1.779	0.544	0.456	0
NM	472.560	0.416	38.960	1.780	0.544	0.456	0
SANN	485.364	0.410	37.908	1.754	0.550	0.450	0
SCE	267.422	0.983	47.482	2.189	0.541	0.459	0
DE	267.166	0.985	47.479	2.189	0.541	0.459	0

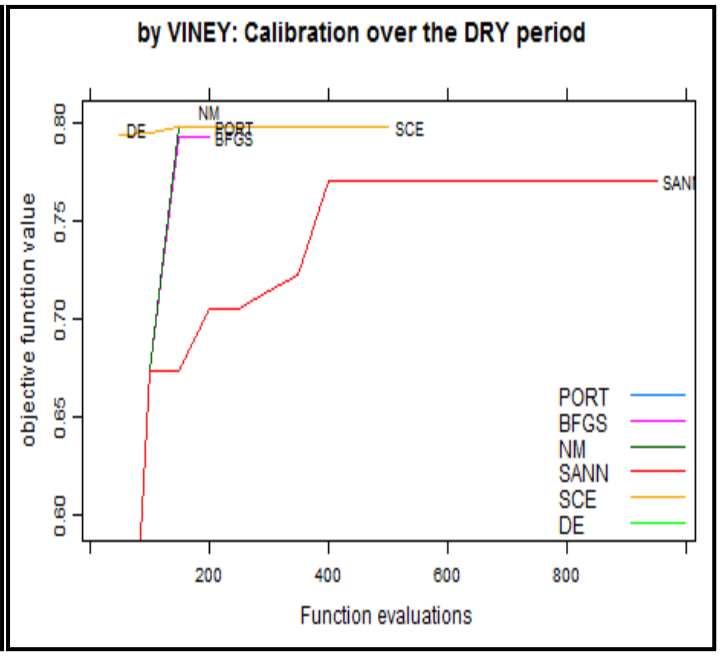
APPENDIX B: AWBM model

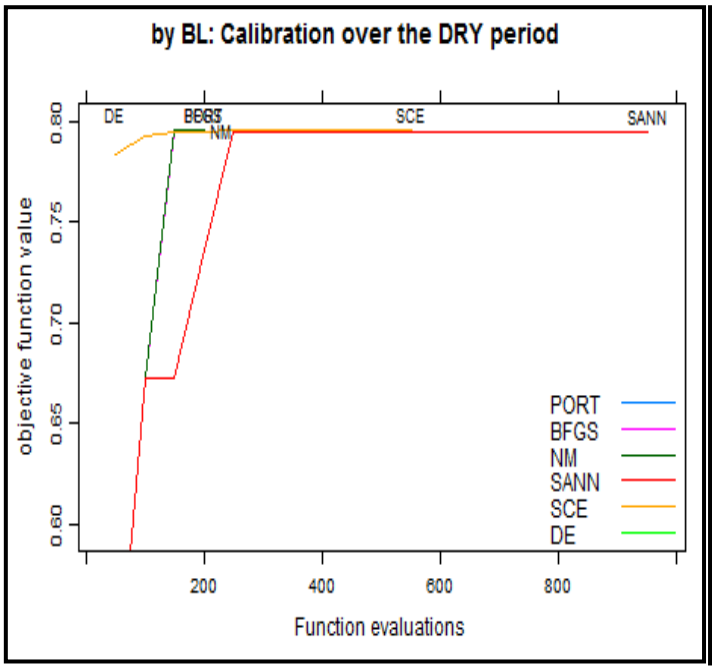
(E) Objective Function and Optimization Algorithm uncertainty

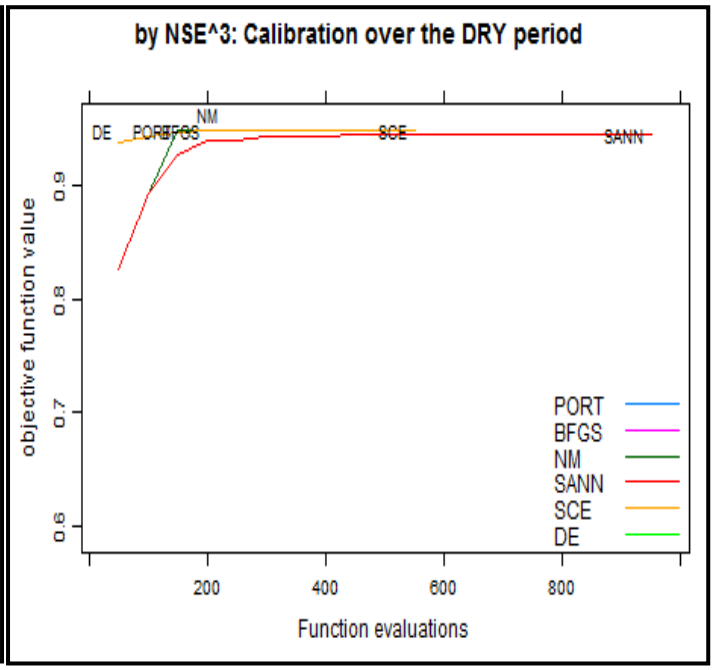
CALIBRATION: **DRY** – VALIDATION: **WET**

Optimization traces:









APPENDIX B: AWBM model

Validation FIT STATISTICS:

Objective Function = **NSE**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.078	0.802	0.741	0.734	0.733	0.908
BFGS	0.065	0.805	0.745	0.740	0.738	0.909
NM	0.077	0.802	0.741	0.734	0.734	0.908
SANN	0.102	0.798	0.724	0.709	0.713	0.896
SCE	0.078	0.802	0.741	0.734	0.733	0.908
DE	0.068	0.805	0.742	0.736	0.735	0.907

Objective Function = **Viney**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.076	0.802	0.742	0.734	0.734	0.908
BFGS	0.050	0.808	0.747	0.744	0.742	0.909
NM	0.076	0.802	0.742	0.734	0.734	0.908
SANN	0.100	0.799	0.724	0.709	0.714	0.896
SCE	0.076	0.802	0.742	0.734	0.734	0.908
DE	0.077	0.802	0.741	0.734	0.734	0.908

Objective Function = **BL**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.076	0.802	0.742	0.734	0.734	0.908
BFGS	0.070	0.804	0.744	0.738	0.737	0.909
NM	0.076	0.802	0.742	0.734	0.734	0.908
SANN	0.078	0.802	0.741	0.733	0.733	0.908
SCE	0.076	0.802	0.742	0.734	0.734	0.908
DE	0.079	0.802	0.740	0.732	0.732	0.907

Objective Function = **NSE^3**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.083	0.801	0.739	0.730	0.730	0.907
BFGS	0.084	0.801	0.738	0.729	0.730	0.907
NM	0.083	0.801	0.739	0.730	0.730	0.907
SANN	0.118	0.795	0.719	0.698	0.707	0.896
SCE	0.083	0.801	0.739	0.730	0.730	0.907
DE	0.086	0.801	0.737	0.727	0.728	0.906

APPENDIX B: AWBM model

Parameter Stability:

Objective Function = **NSE**

	cap.ave	etmult	tau_s	tau_q	v_s	v_q	delay
PORT	263.067	1.000	74.398	1.838	0.514	0.486	0
BFGS	266.566	0.999	74.276	1.830	0.514	0.486	0
NM	263.156	1.000	74.420	1.838	0.514	0.486	0
SANN	273.489	0.920	72.687	1.854	0.494	0.506	0
SCE	263.064	1.000	74.397	1.838	0.514	0.486	0
DE	270.252	0.977	73.722	1.833	0.508	0.492	0

Objective Function = **Viney**

	cap.ave	etmult	tau_s	tau_q	v_s	v_q	delay
PORT	263.349	1.000	74.469	1.837	0.514	0.486	0
BFGS	270.584	0.999	74.067	1.823	0.513	0.487	0
NM	263.350	1.000	74.469	1.837	0.514	0.486	0
SANN	274.467	0.918	72.593	1.854	0.493	0.507	0
SCE	263.349	1.000	74.469	1.837	0.514	0.486	0
DE	263.853	0.997	74.389	1.837	0.514	0.486	0

Objective Function = **BL**

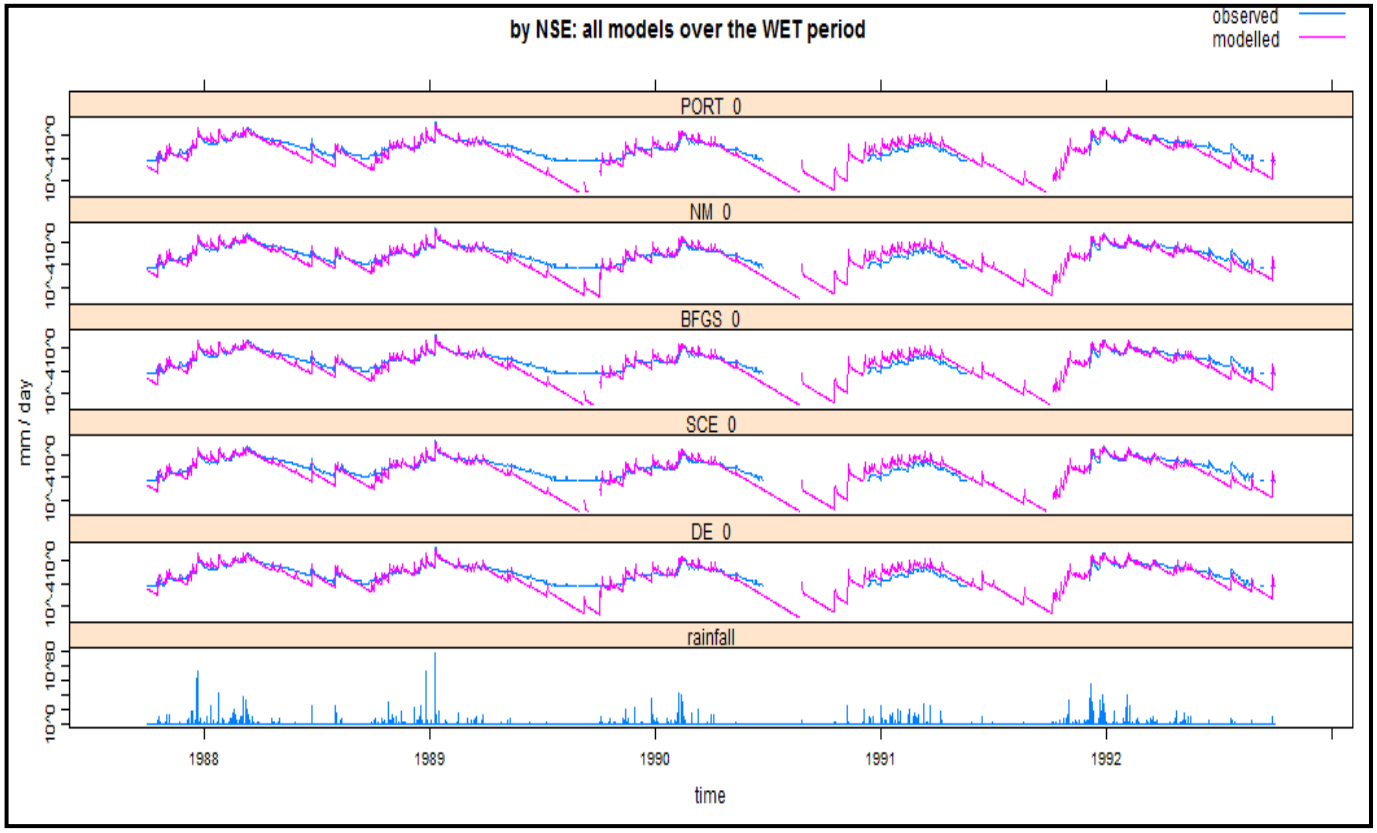
	cap.ave	etmult	tau_s	tau_q	v_s	v_q	delay
PORT	263.349	1.000	74.469	1.837	0.514	0.486	0
BFGS	265.216	1.000	74.359	1.833	0.514	0.486	0
NM	263.344	1.000	74.468	1.837	0.514	0.486	0
SANN	263.795	0.996	74.381	1.838	0.514	0.486	0
SCE	263.349	1.000	74.469	1.837	0.514	0.486	0
DE	262.543	1.000	74.265	1.840	0.513	0.487	0

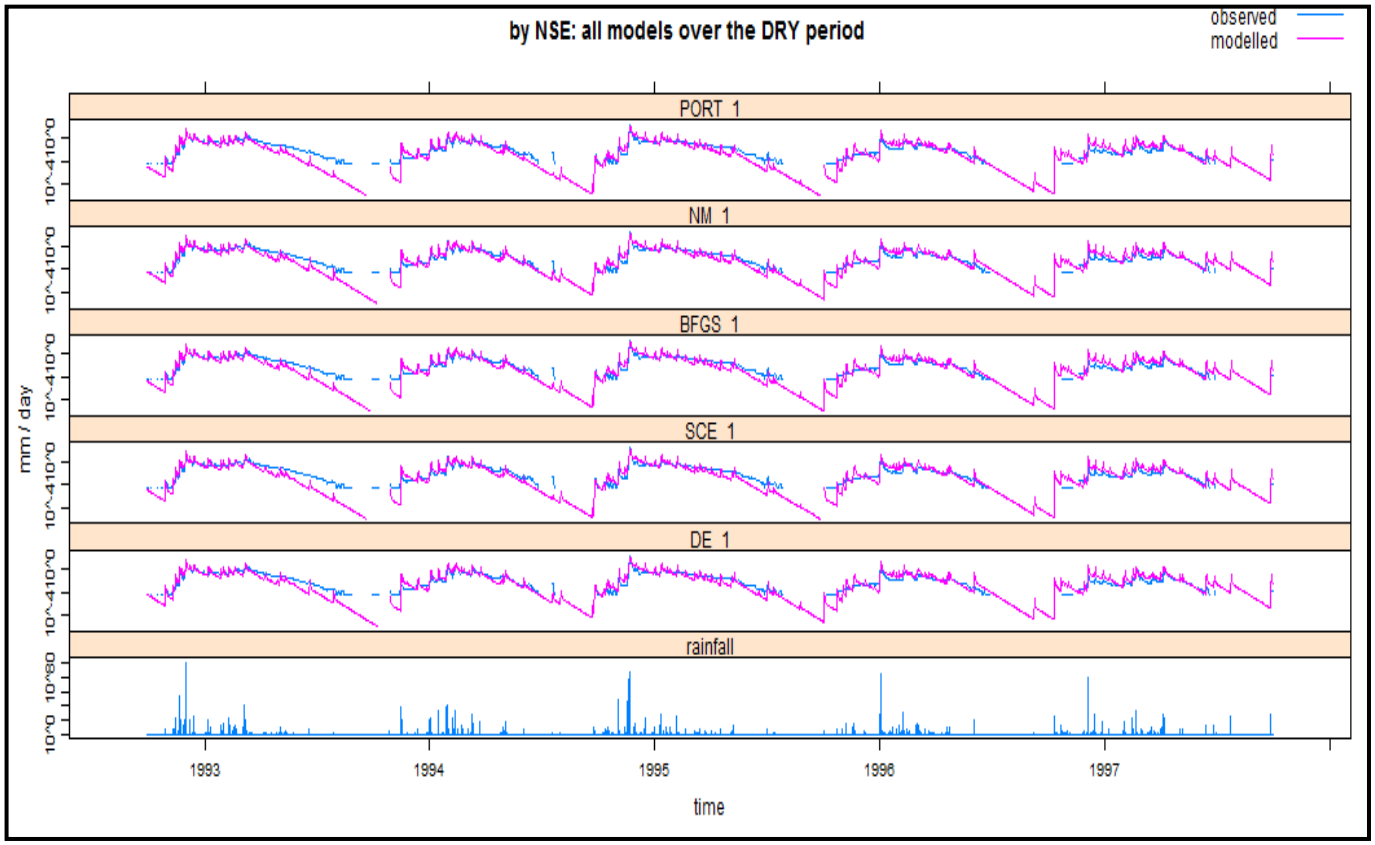
Objective Function = **NSE^3**

	cap.ave	etmult	tau_s	tau_q	v_s	v_q	delay
PORT	261.607	1.000	74.030	1.844	0.512	0.488	0
BFGS	261.594	0.999	73.989	1.844	0.512	0.488	0
NM	261.606	1.000	74.029	1.844	0.512	0.488	0
SANN	260.794	0.970	72.699	1.866	0.500	0.500	0
SCE	261.610	1.000	74.031	1.844	0.512	0.488	0
DE	262.331	0.993	73.812	1.846	0.510	0.490	0

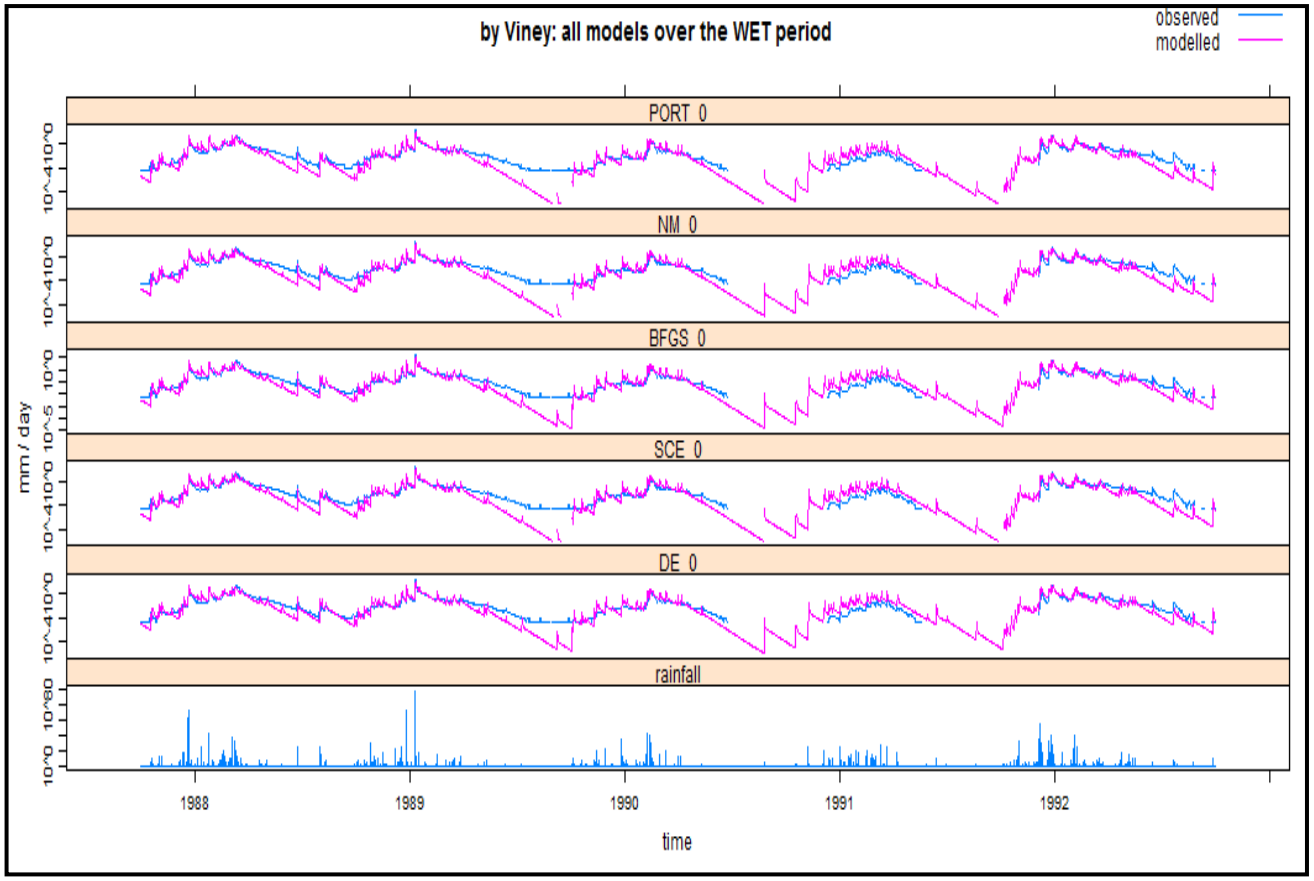
IHACRES model - CWI version:

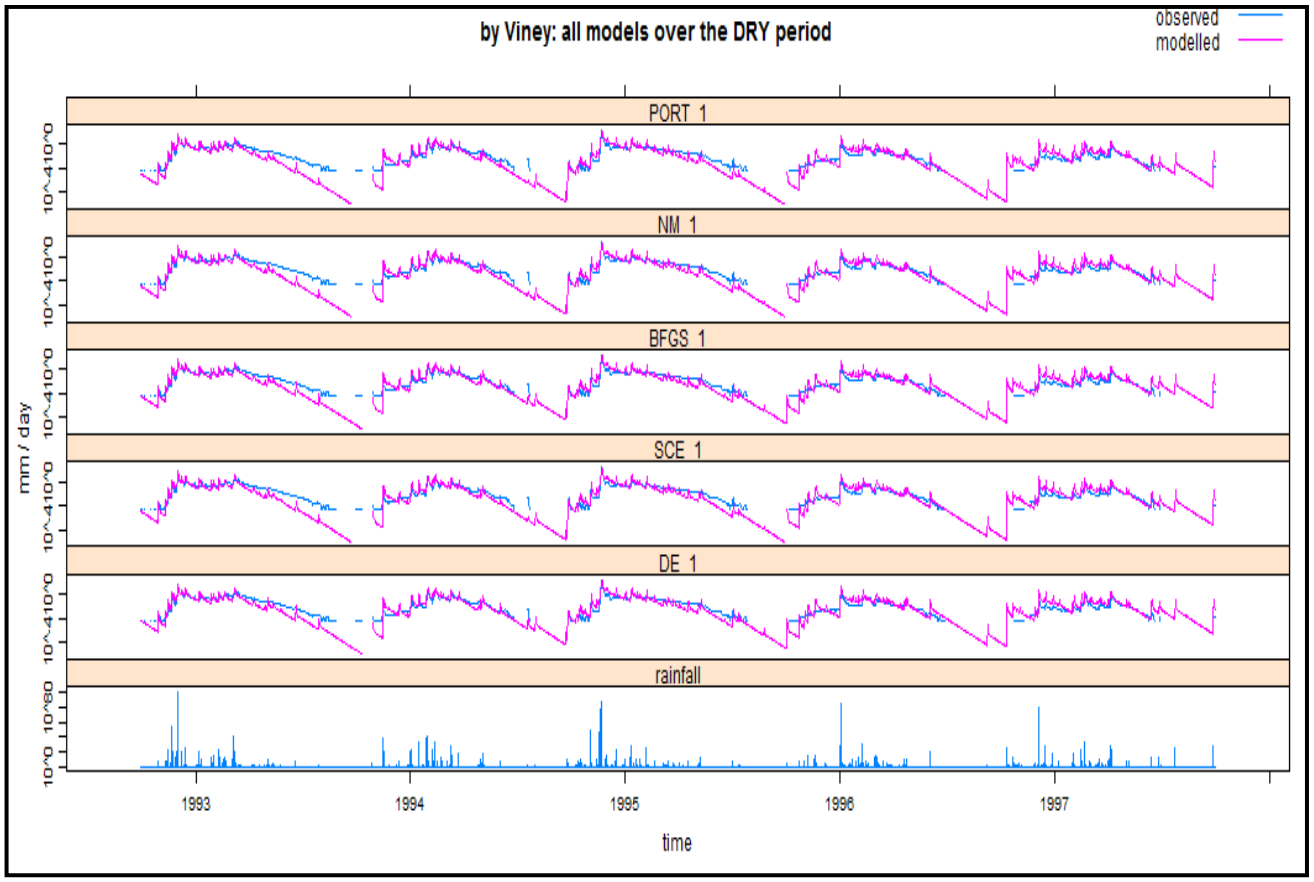
(A) Simulated streamflows (in log scale) – all models



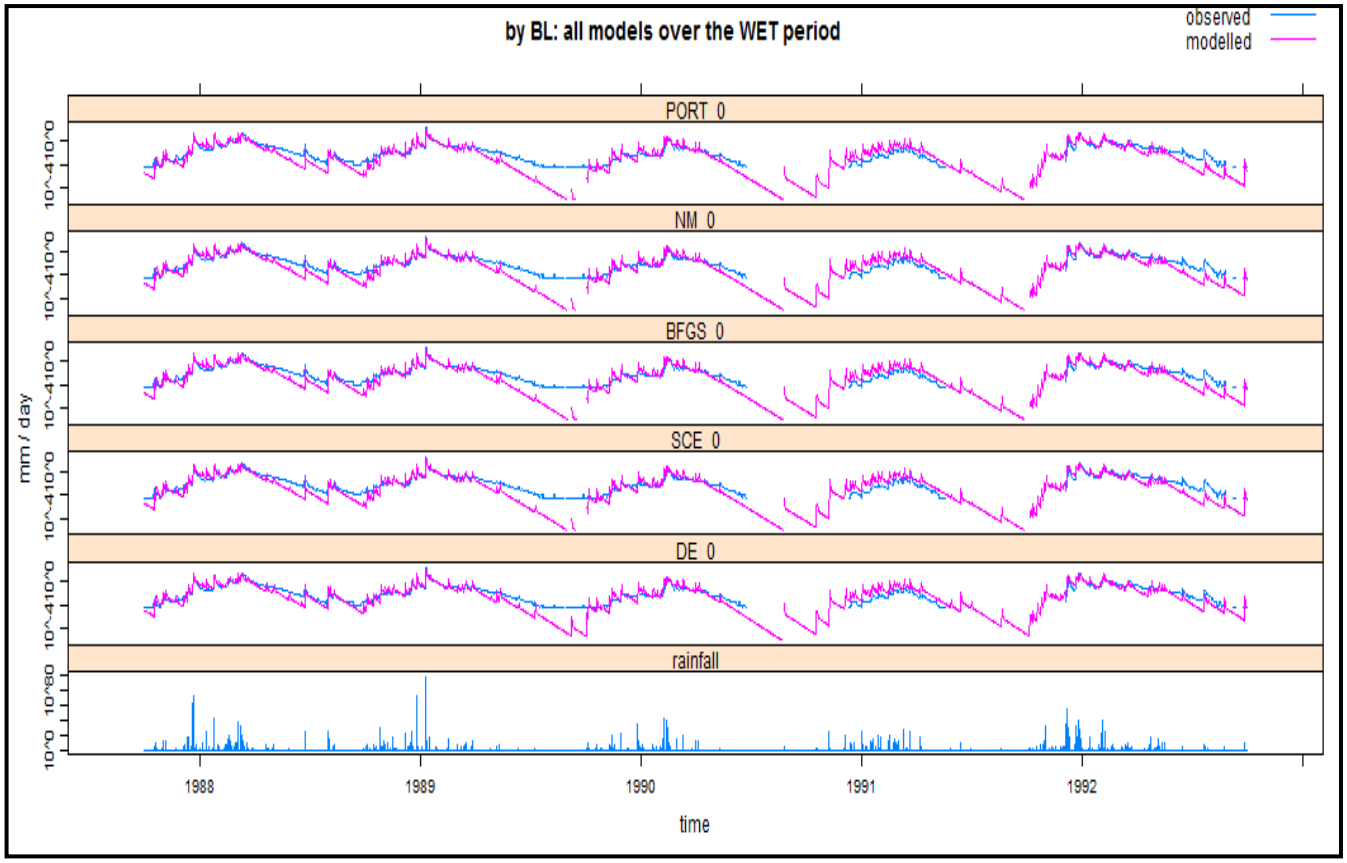


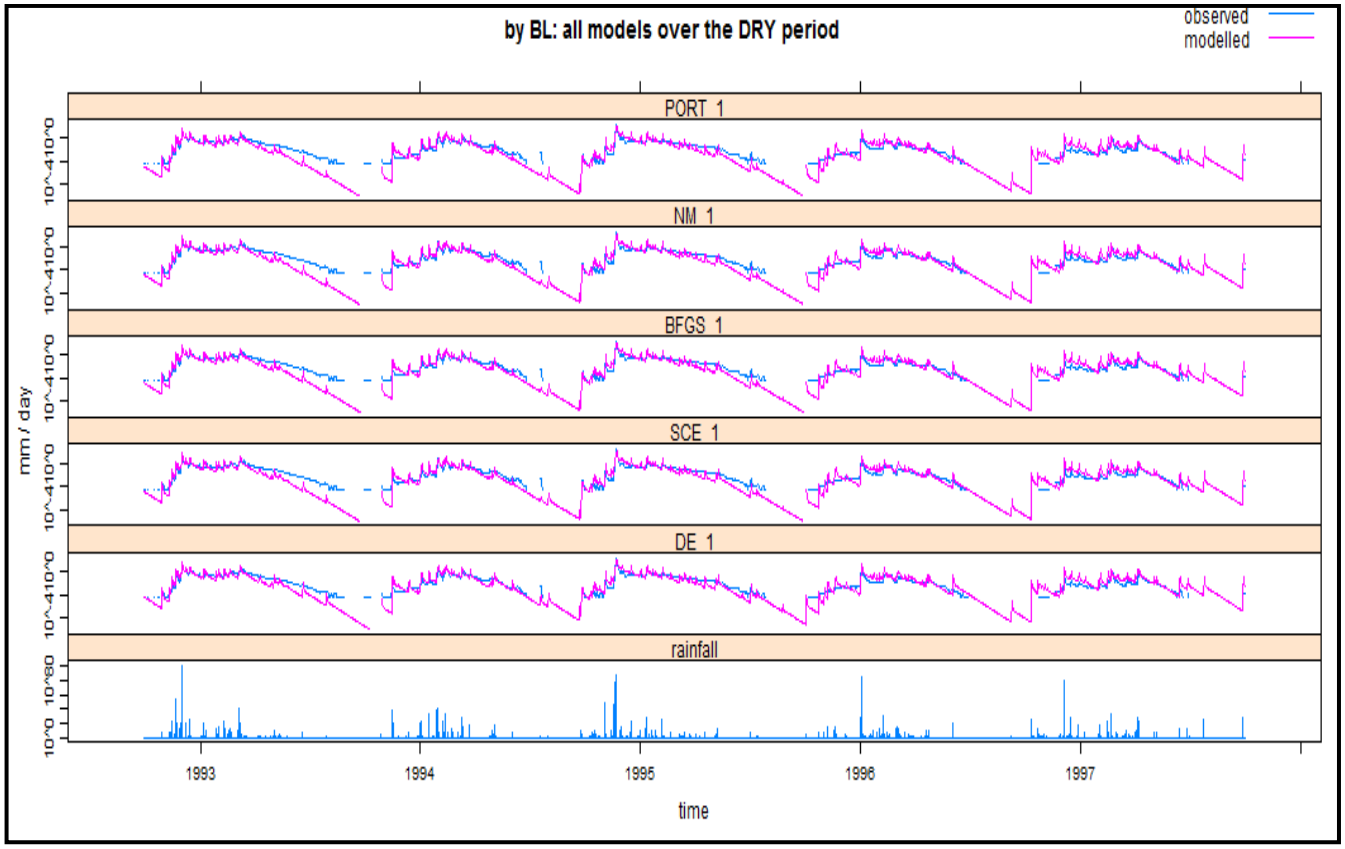
APPENDIX C



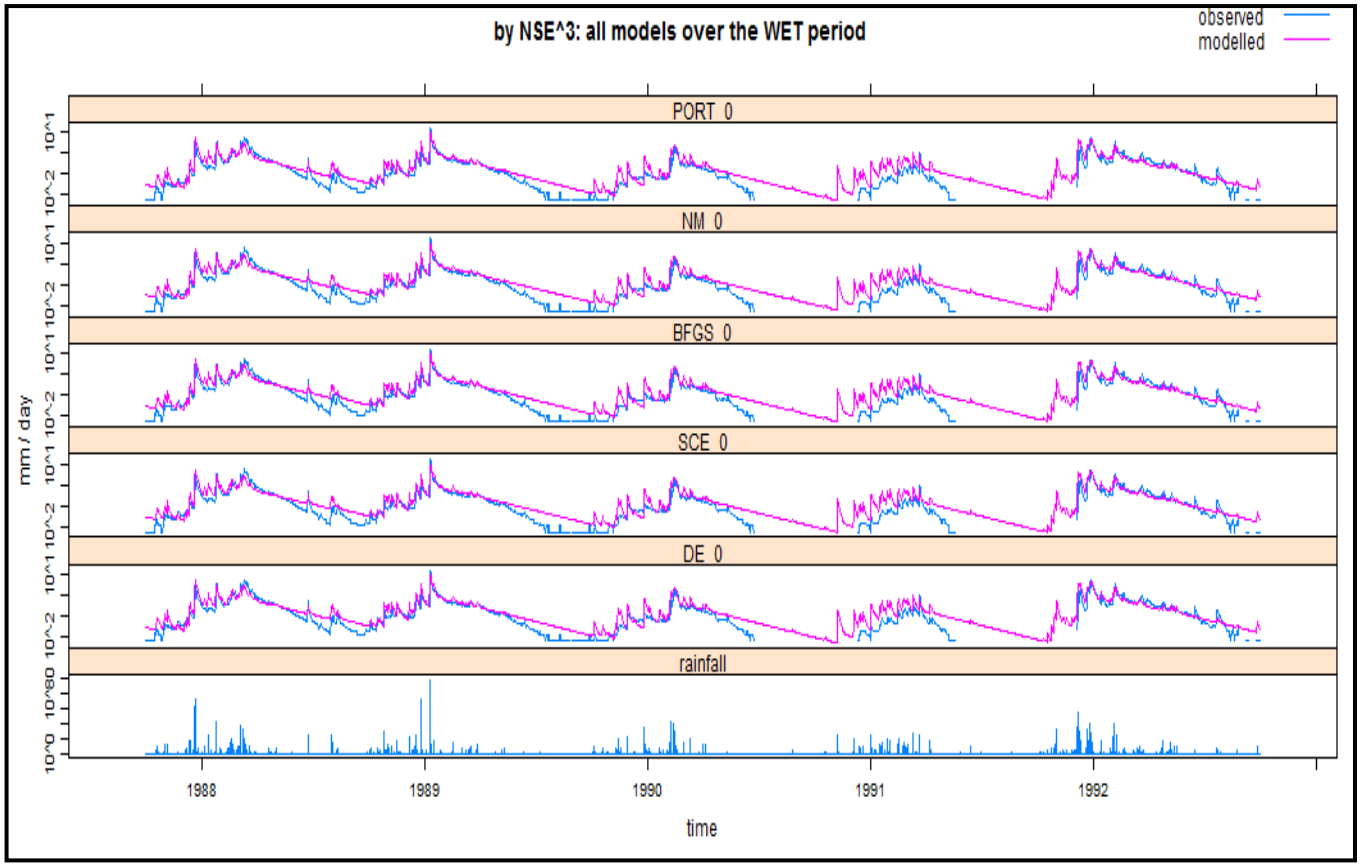


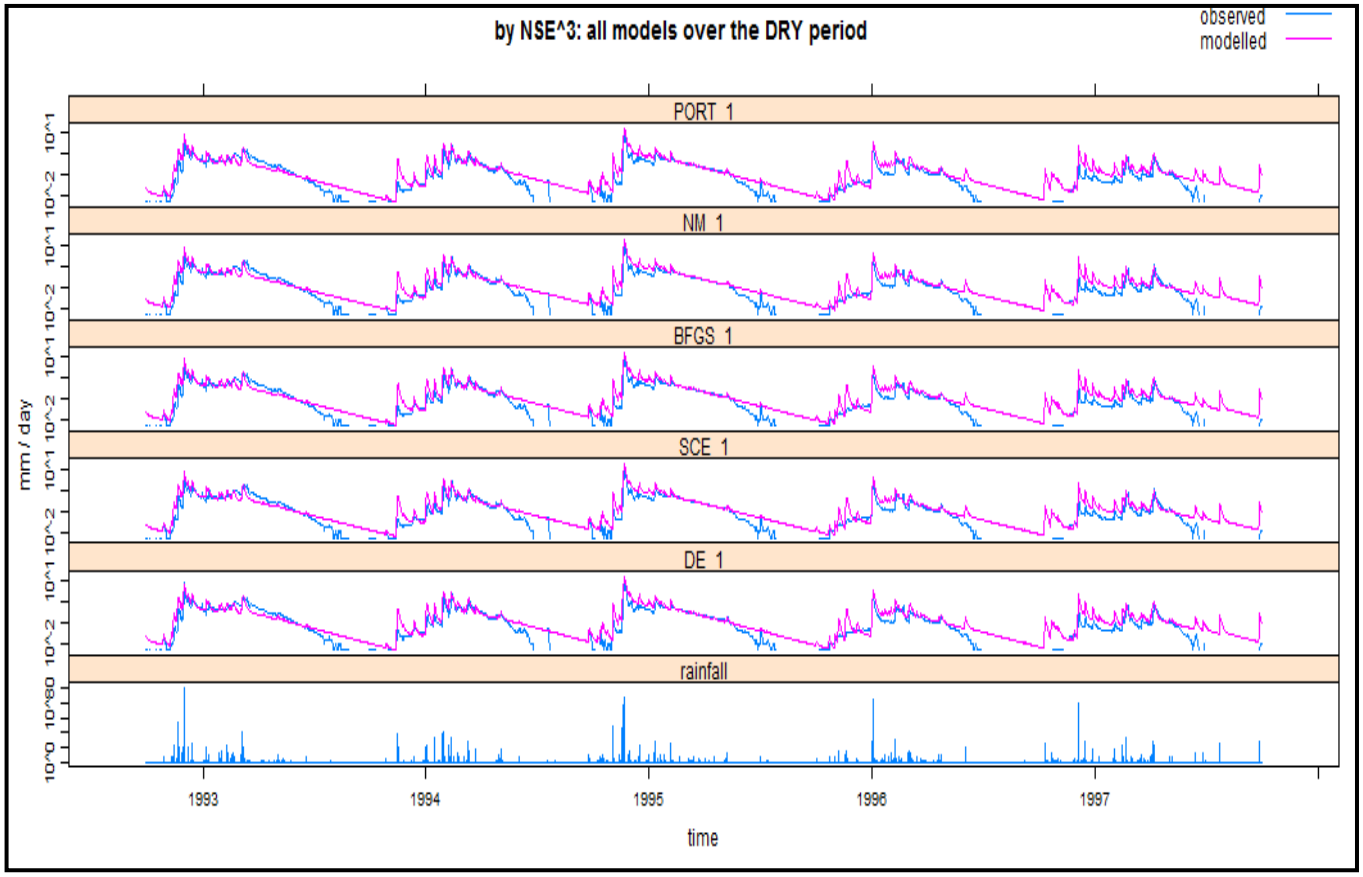
APPENDIX C





APPENDIX C

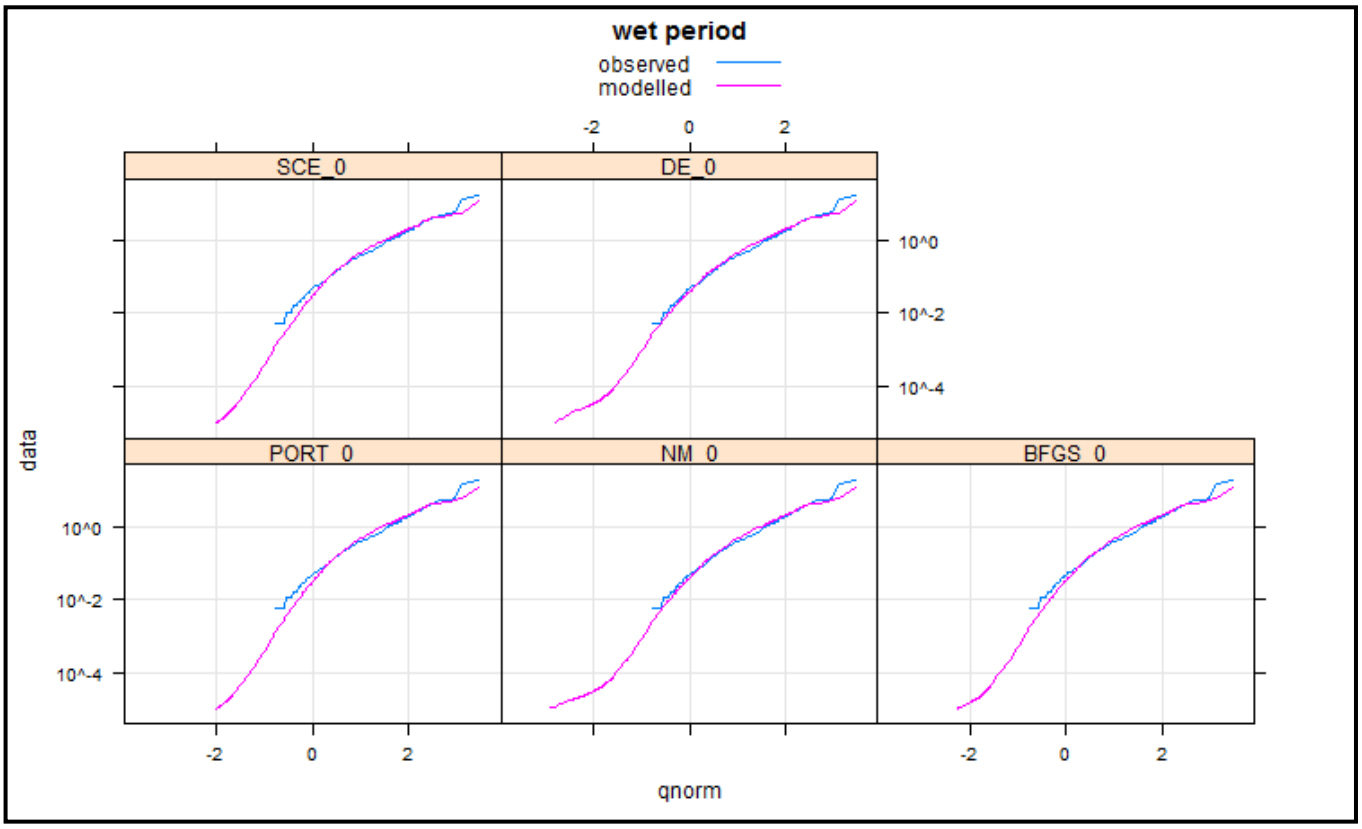


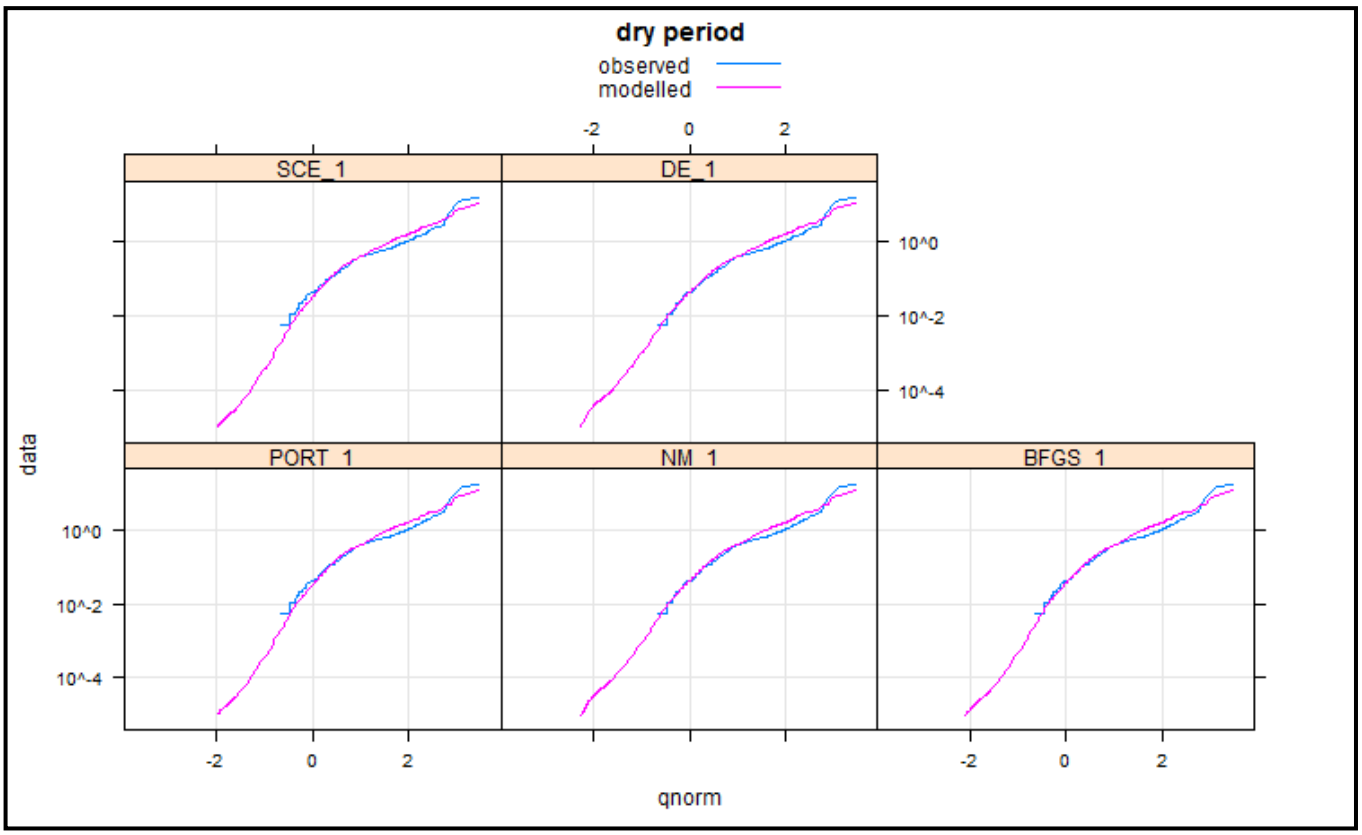


APPENDIX C

(B) Normal distribution Q-Q plot – all models

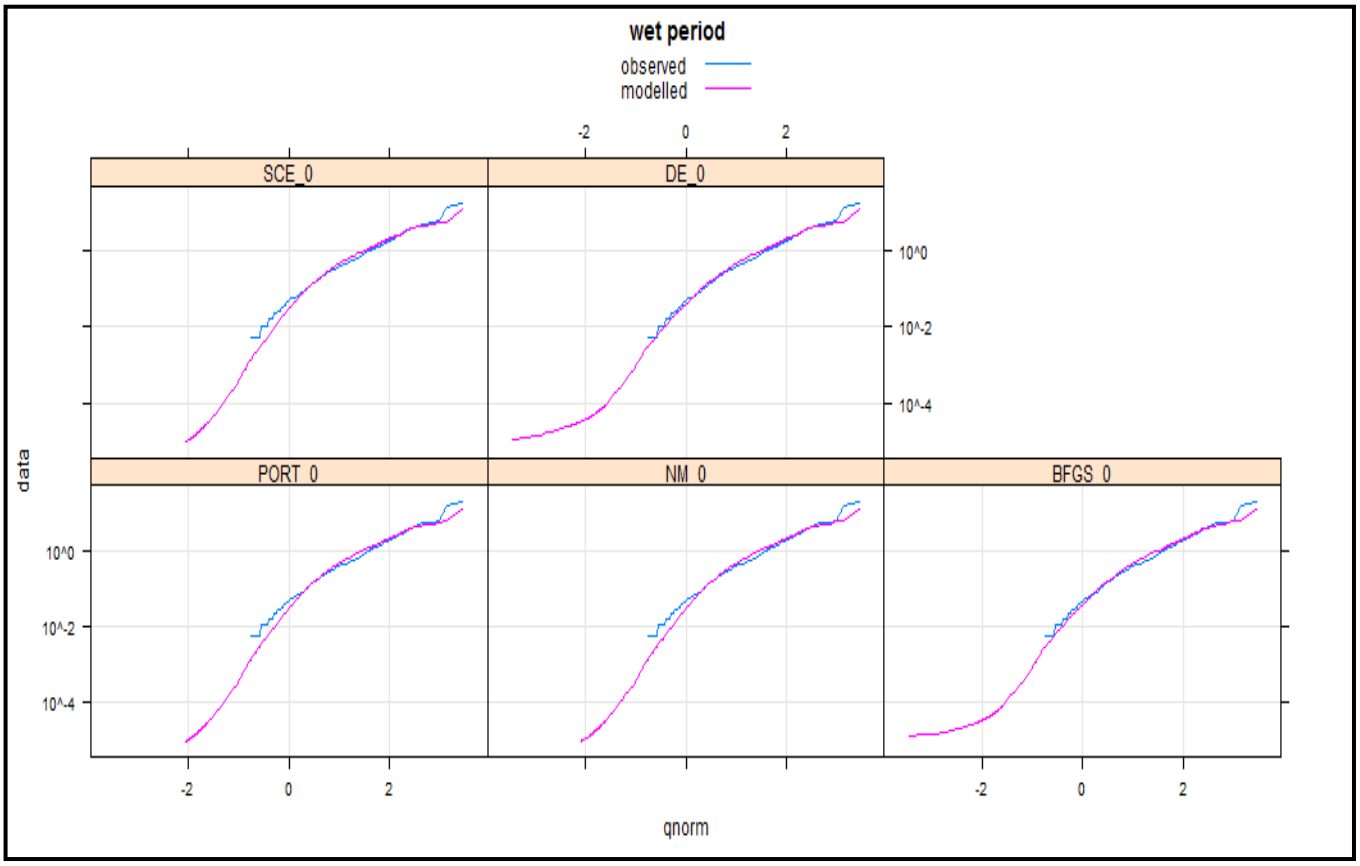
By NSE:

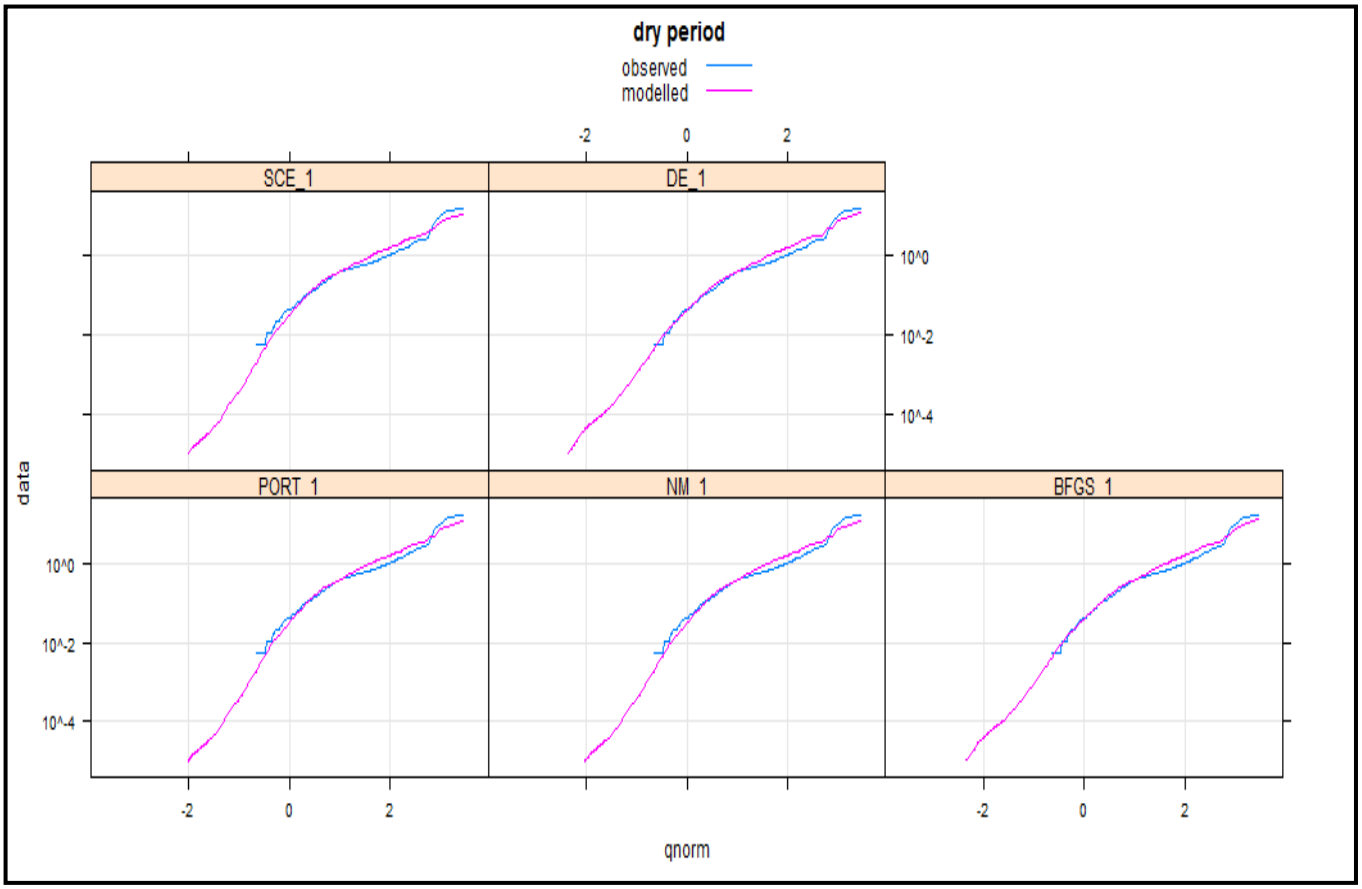




APPENDIX C

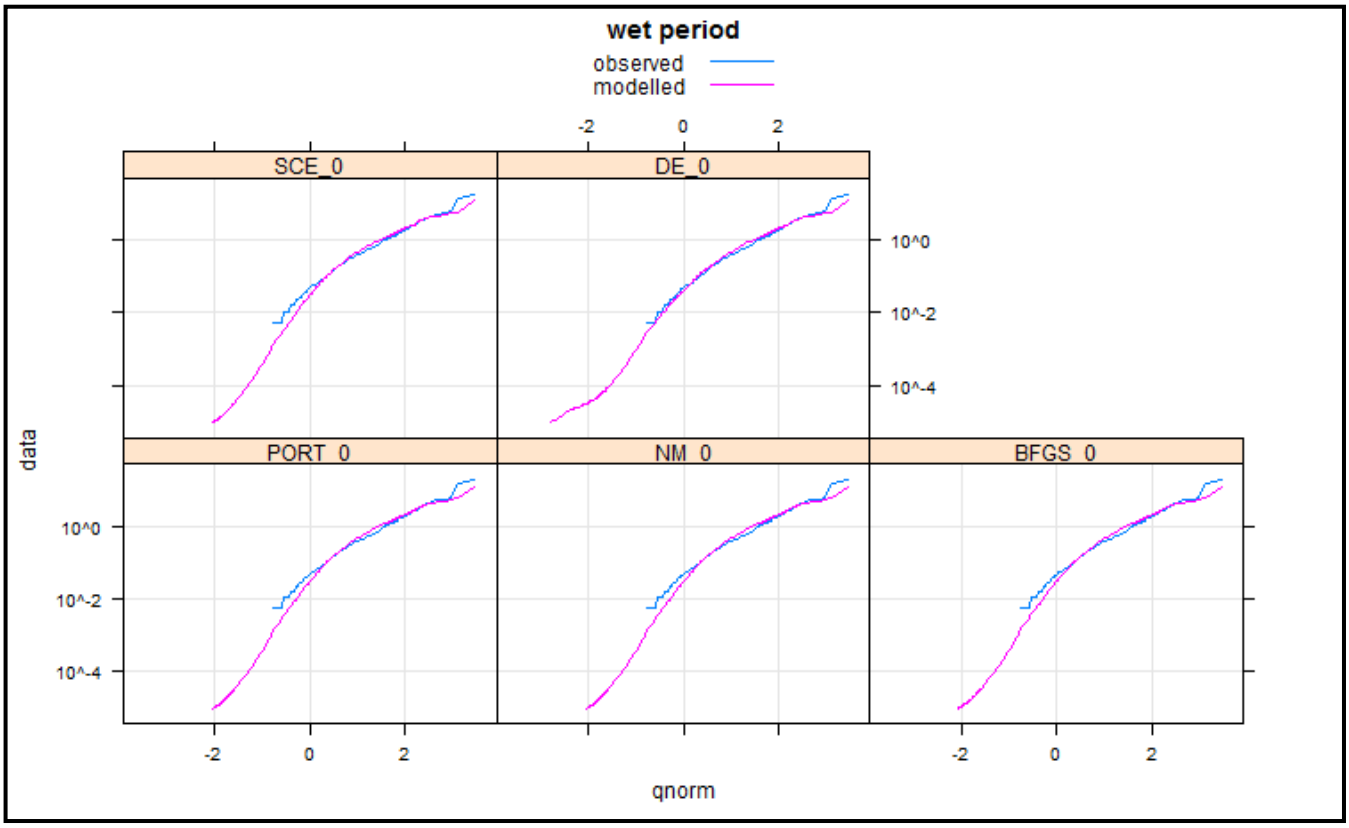
By Viney:

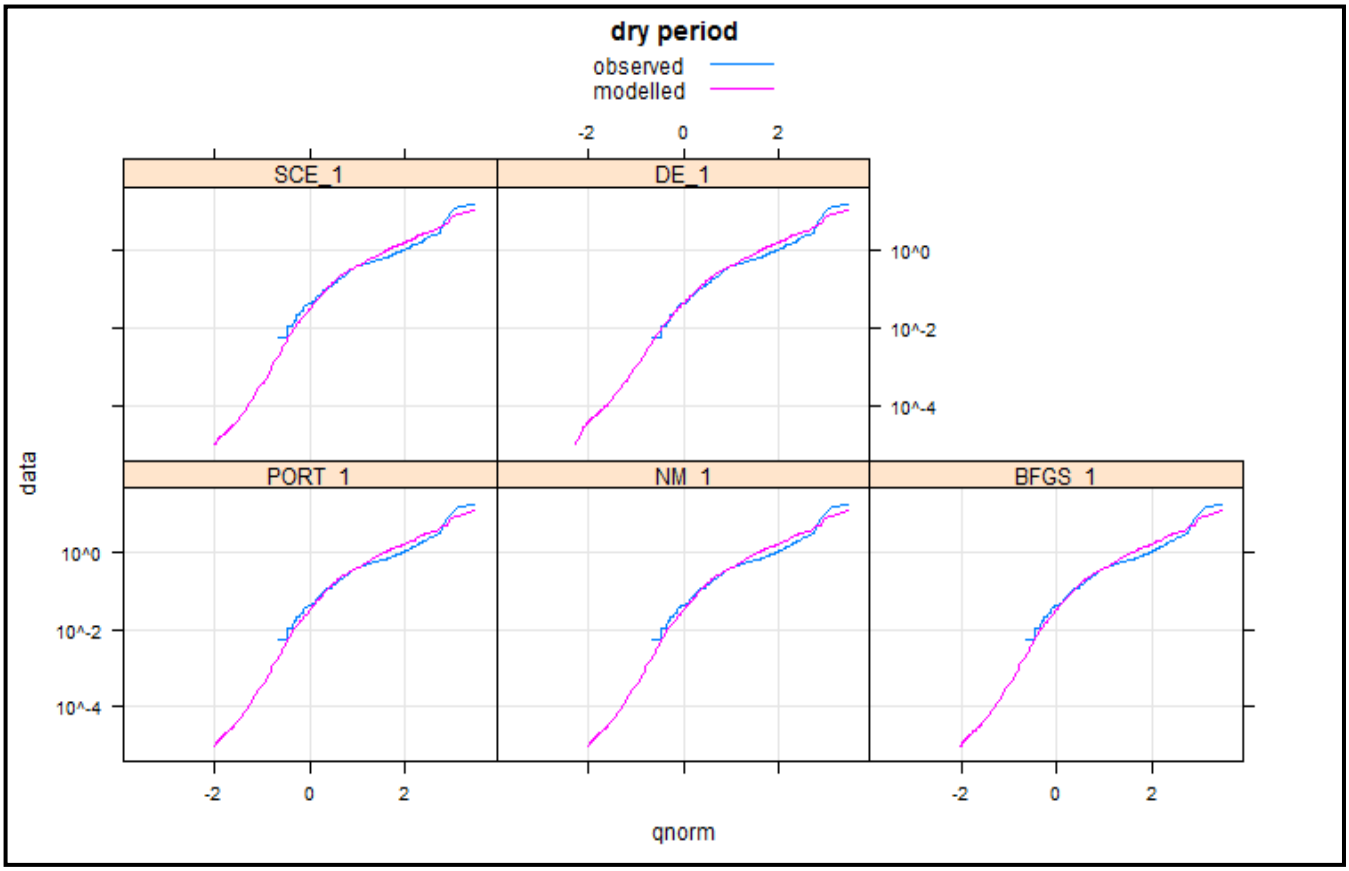




APPENDIX C

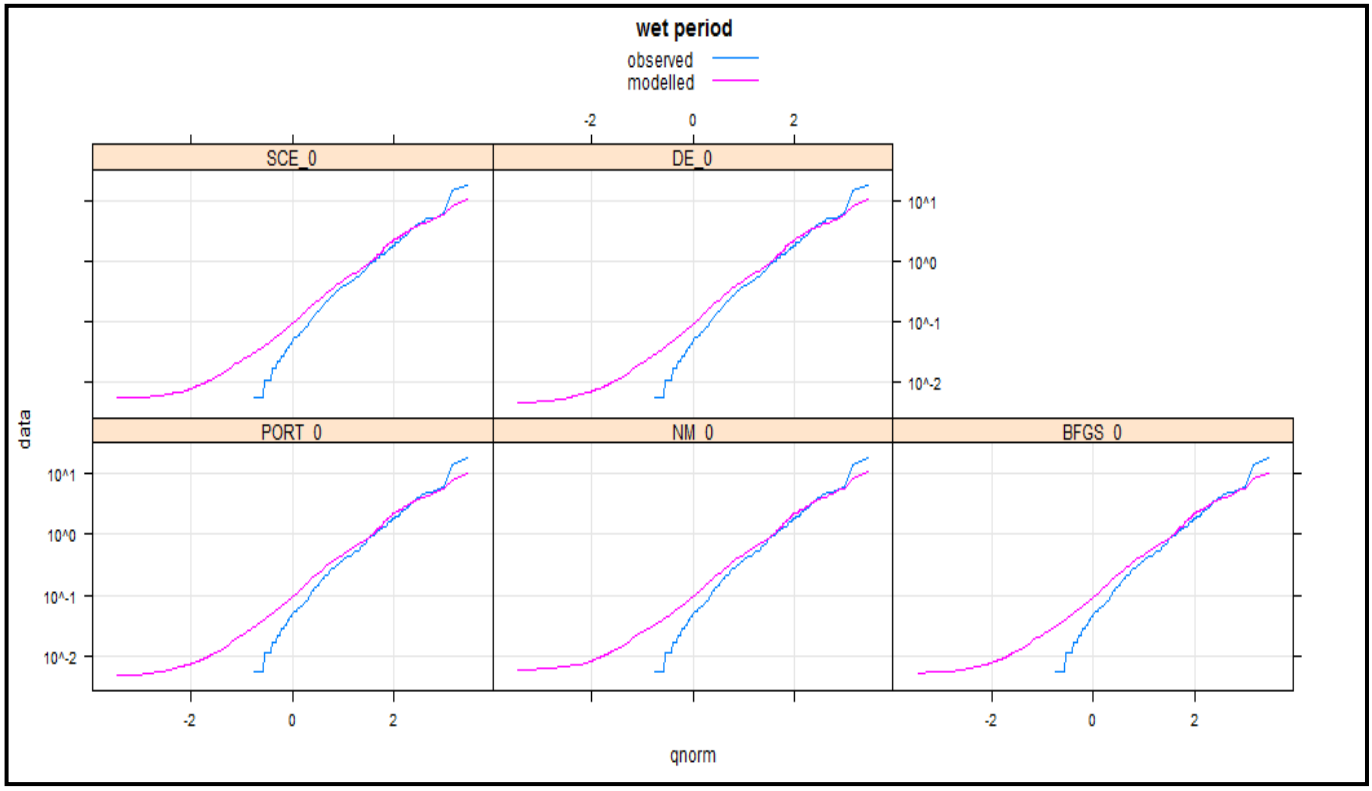
By BL:

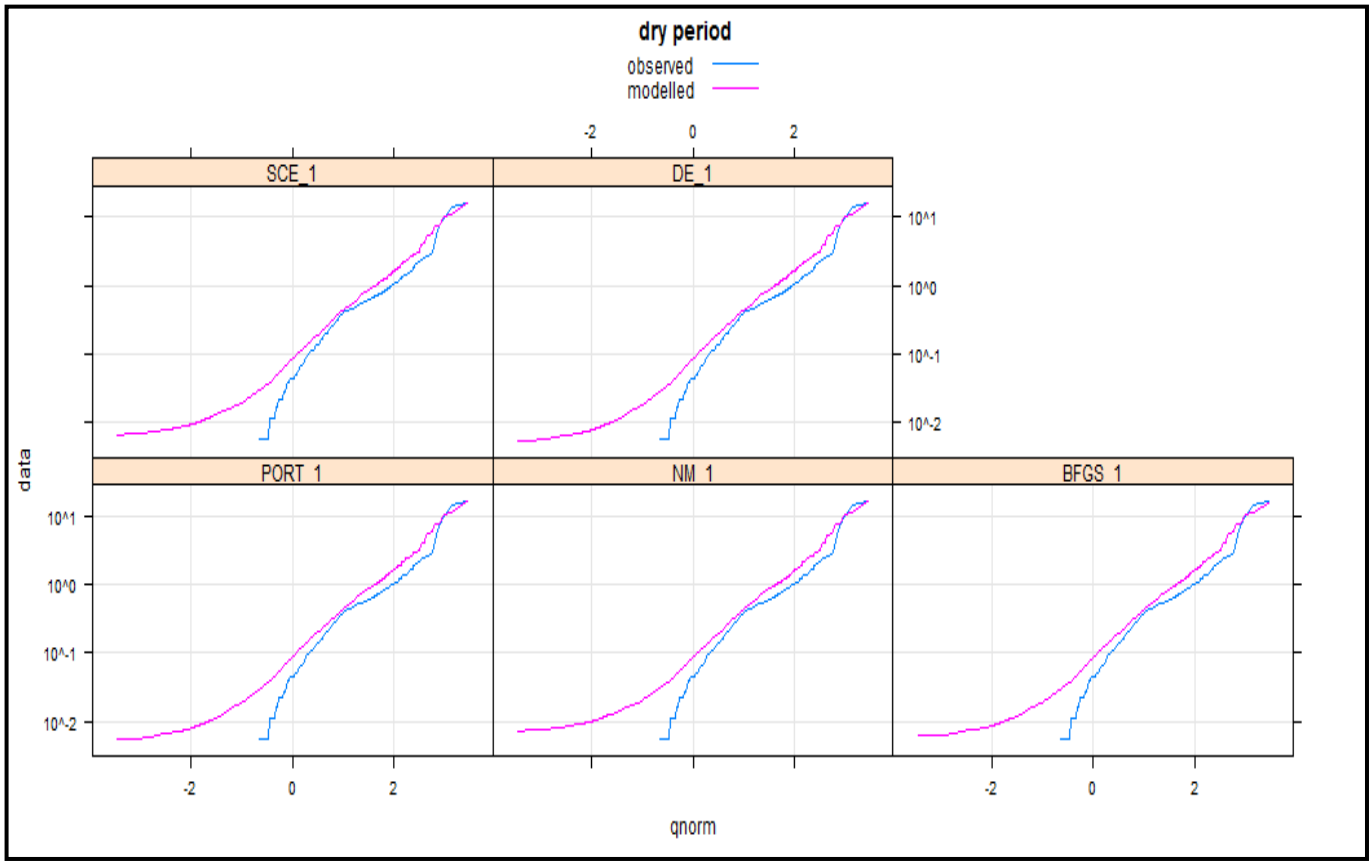




APPENDIX C

By NSE^3:

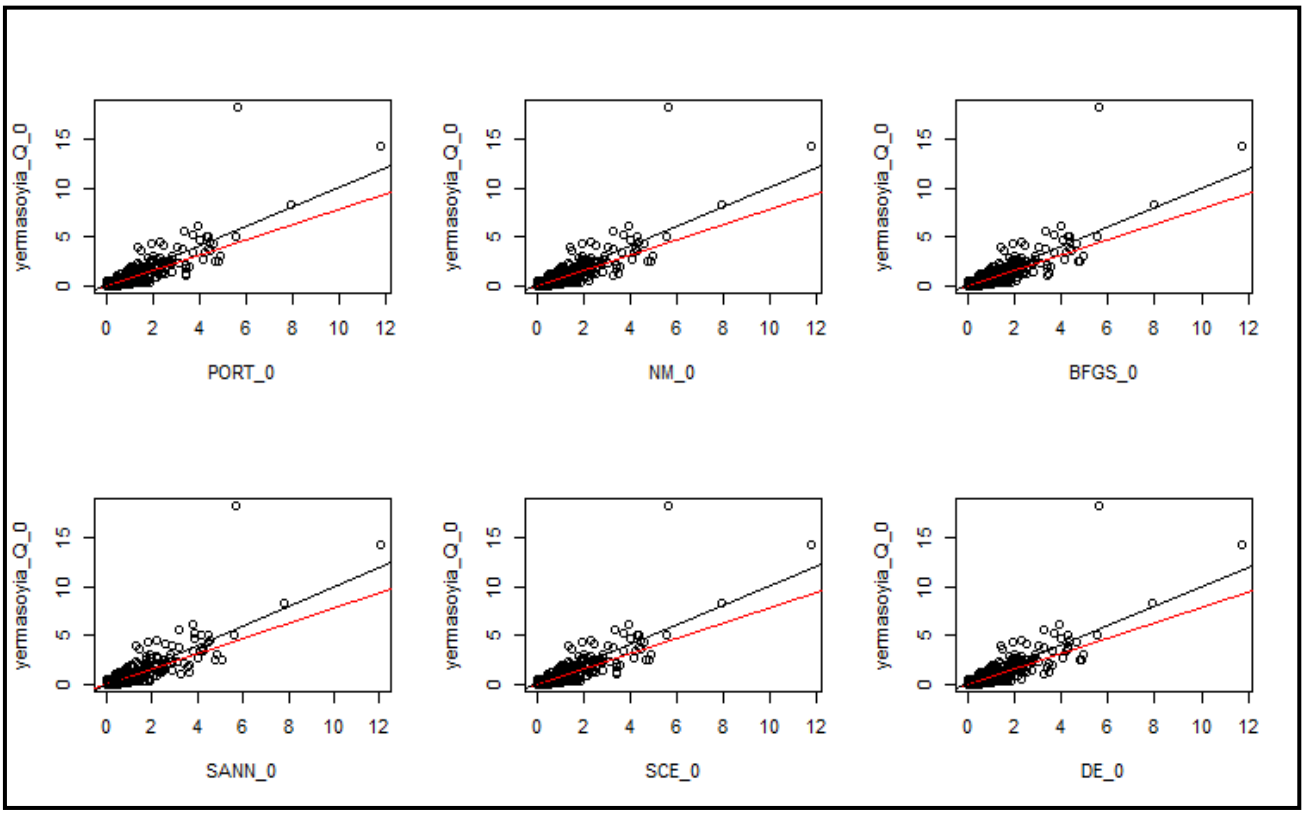




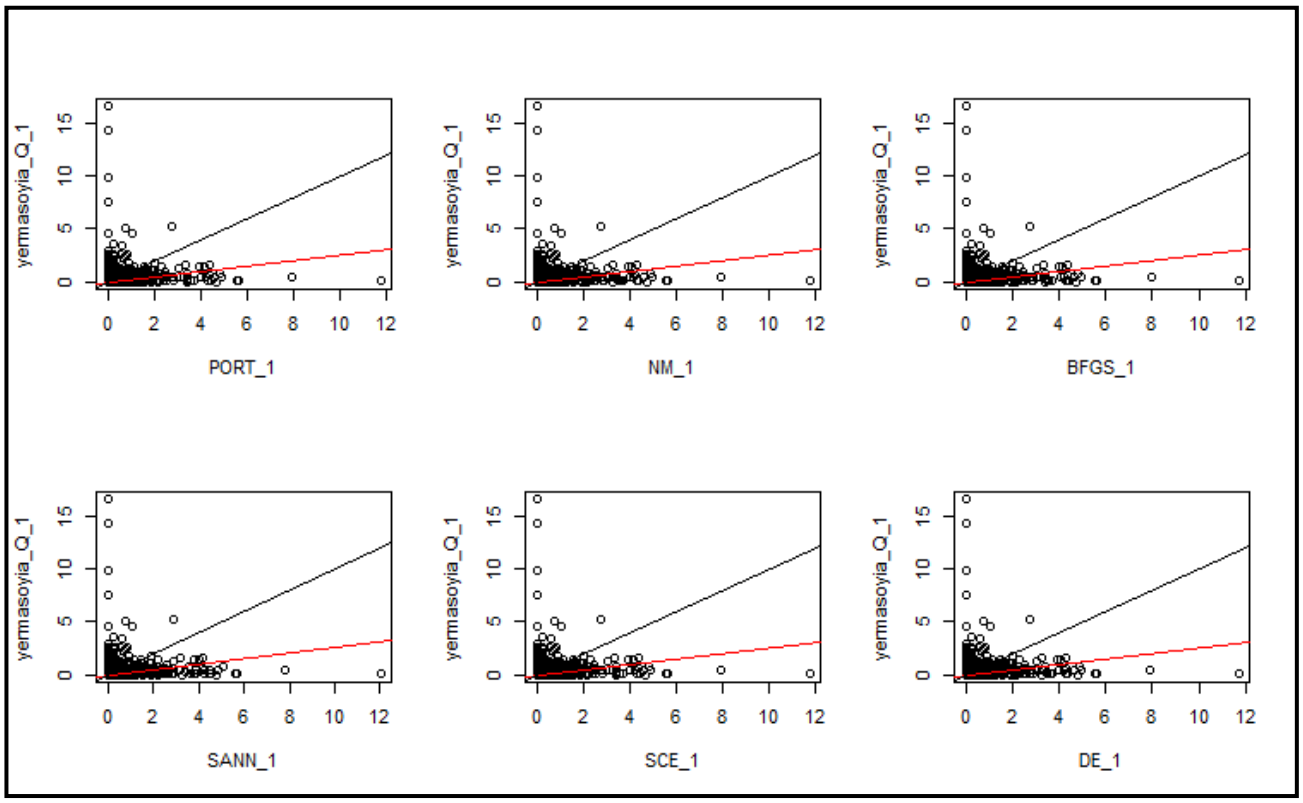
APPENDIX C

(C) Scatterplots

By NSE: wet period

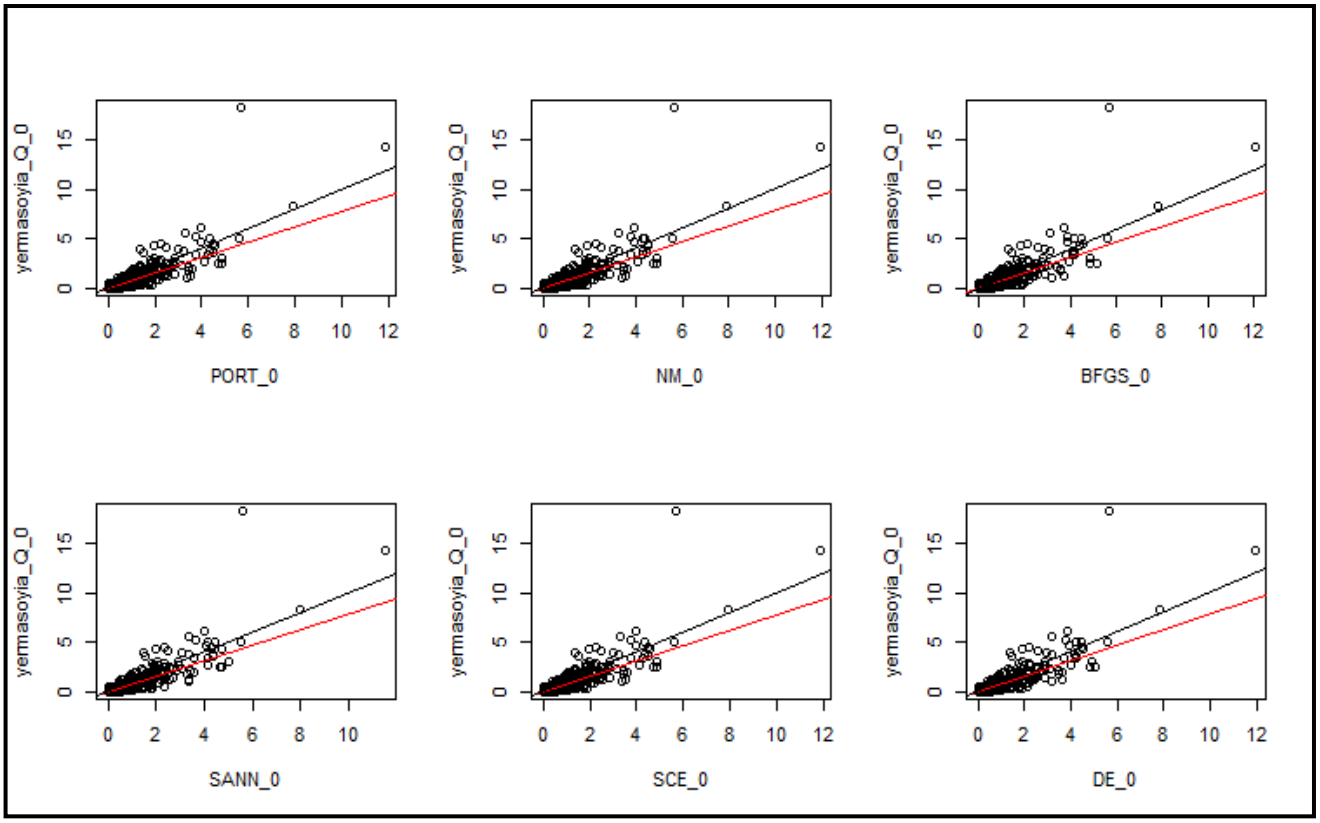


By NSE: dry period

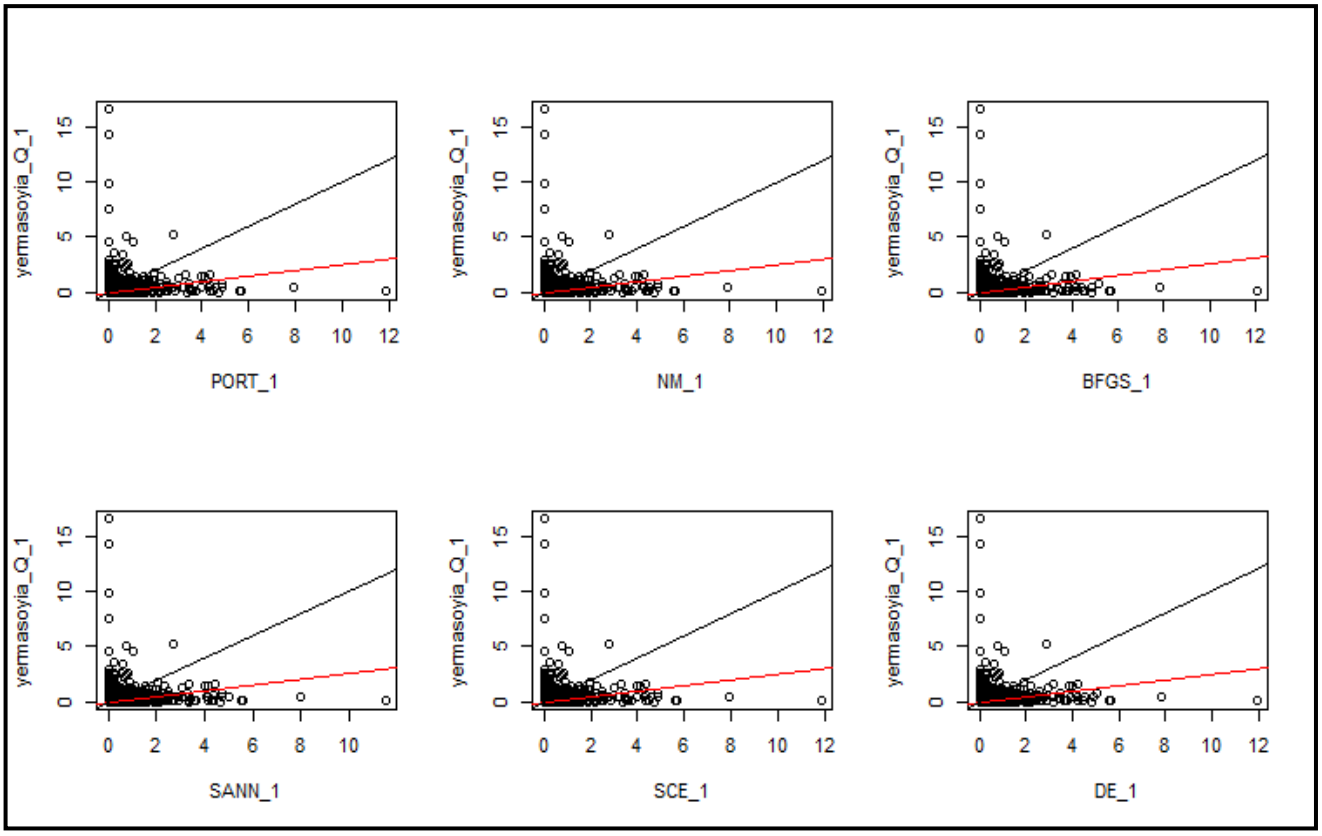


APPENDIX C

By Viney: wet period

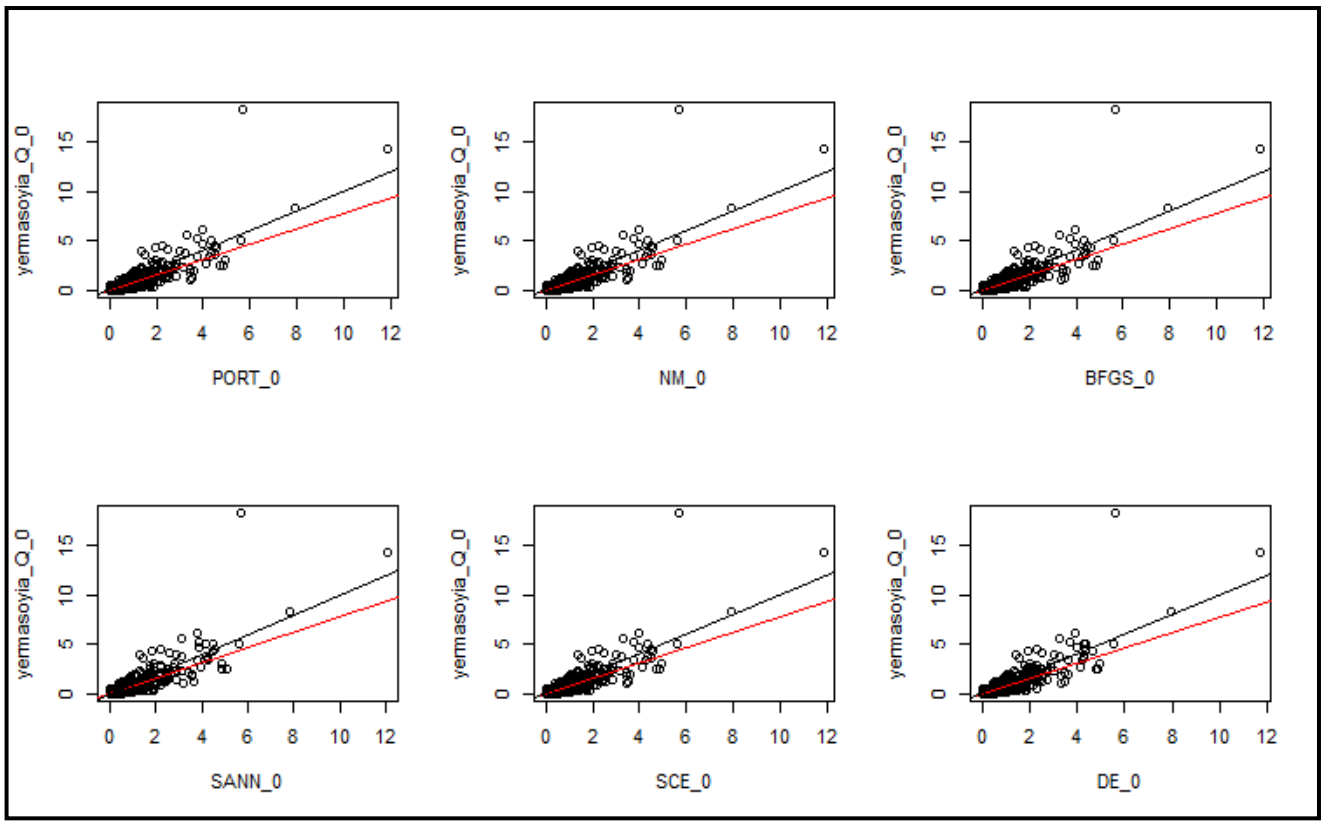


By Viney: dry period

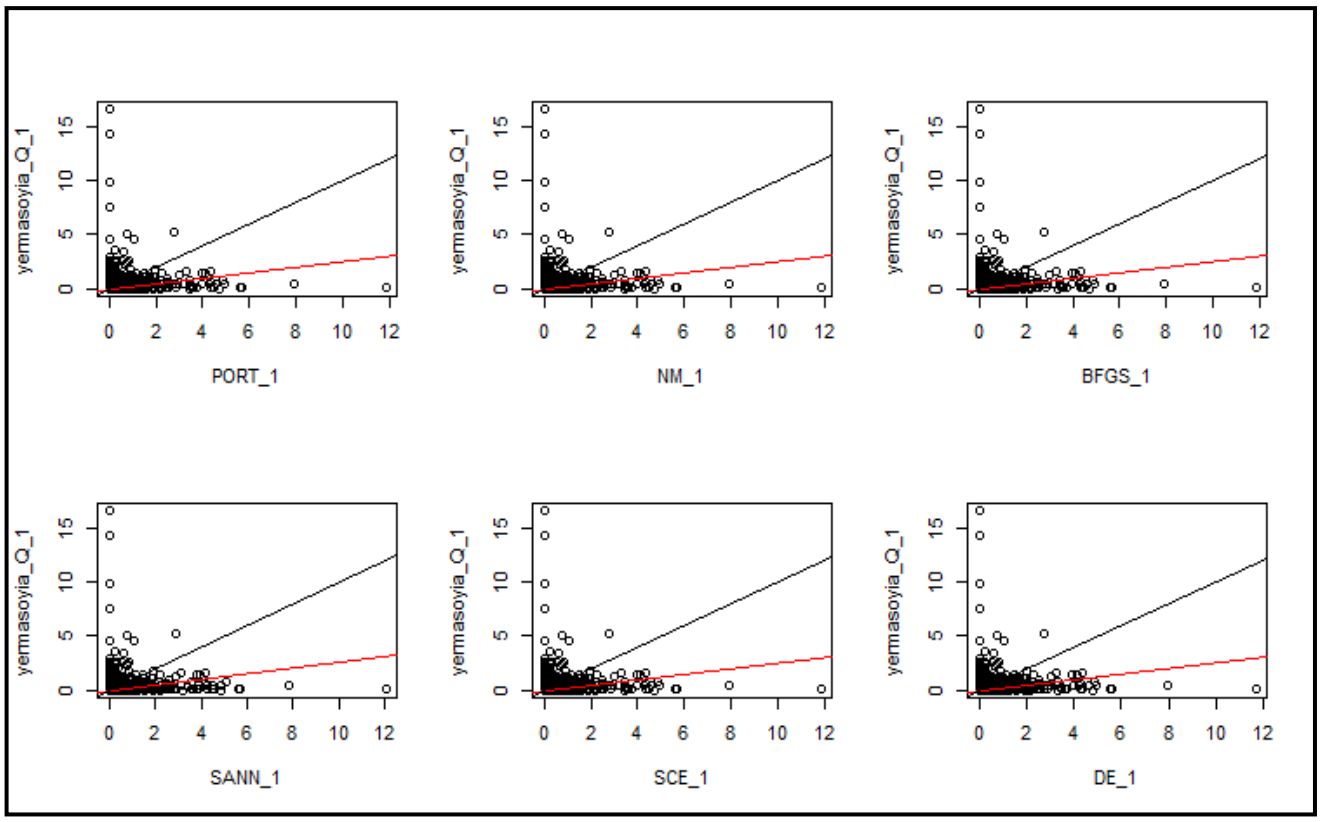


APPENDIX C

By BL: wet period

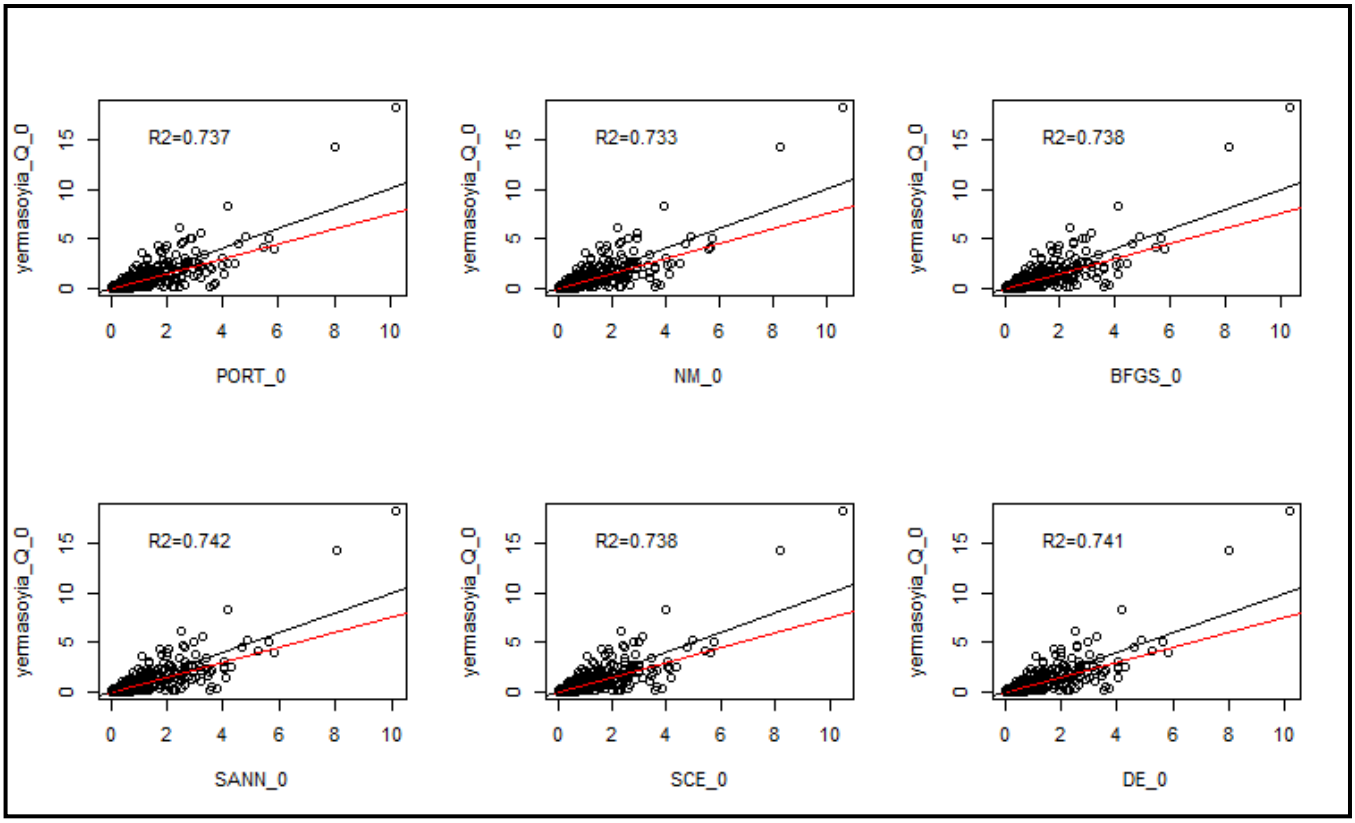


By BL: dry period

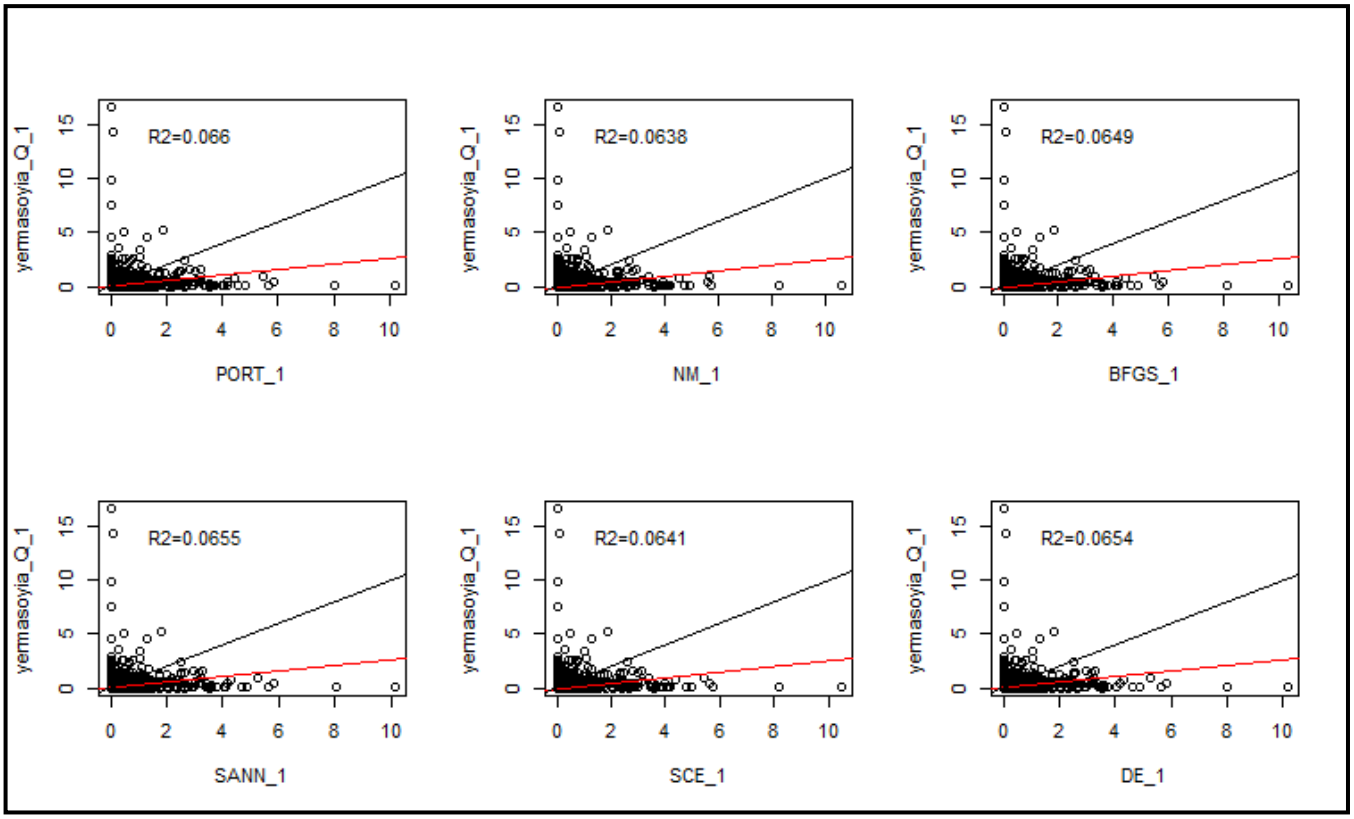


APPENDIX C

By NSE^3: wet period

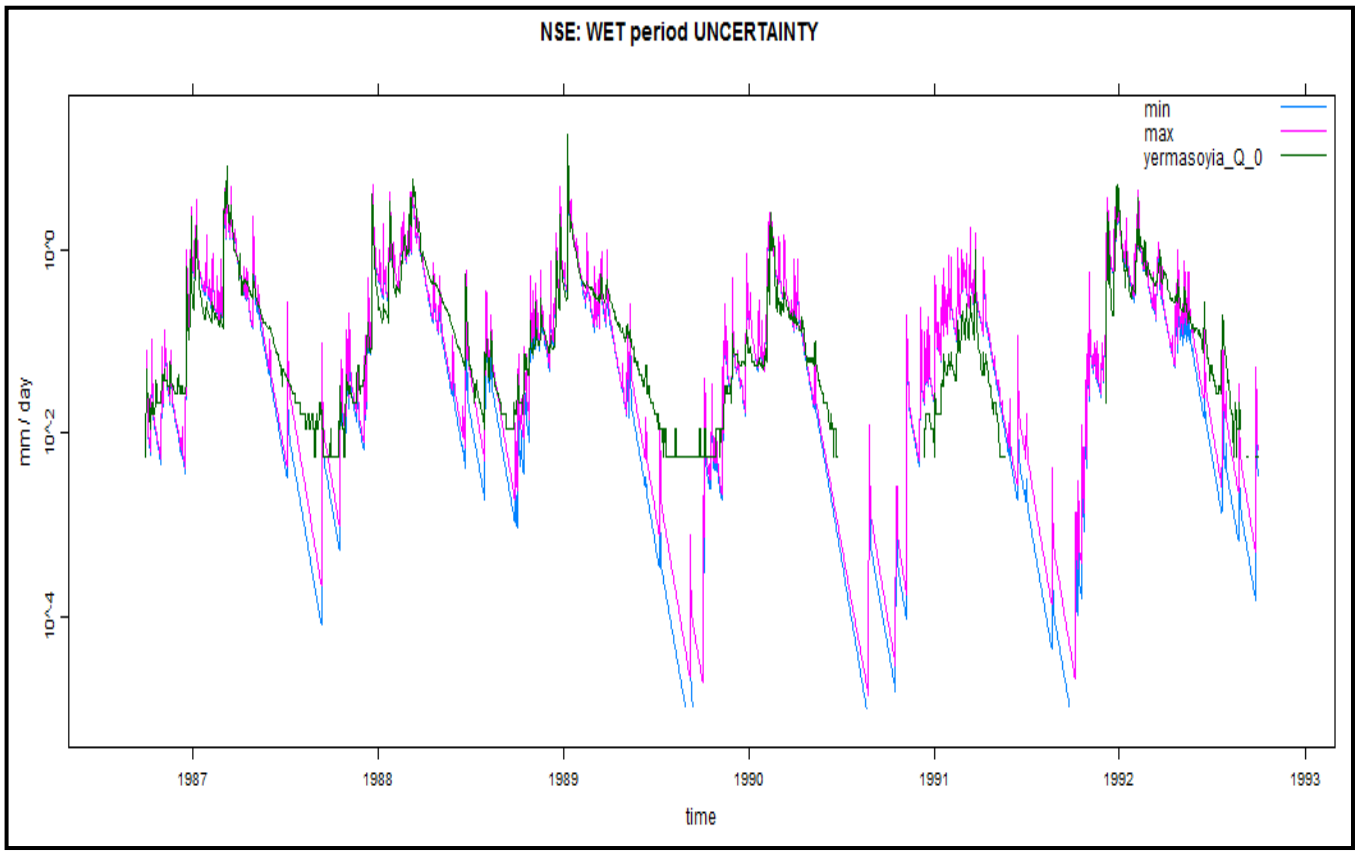


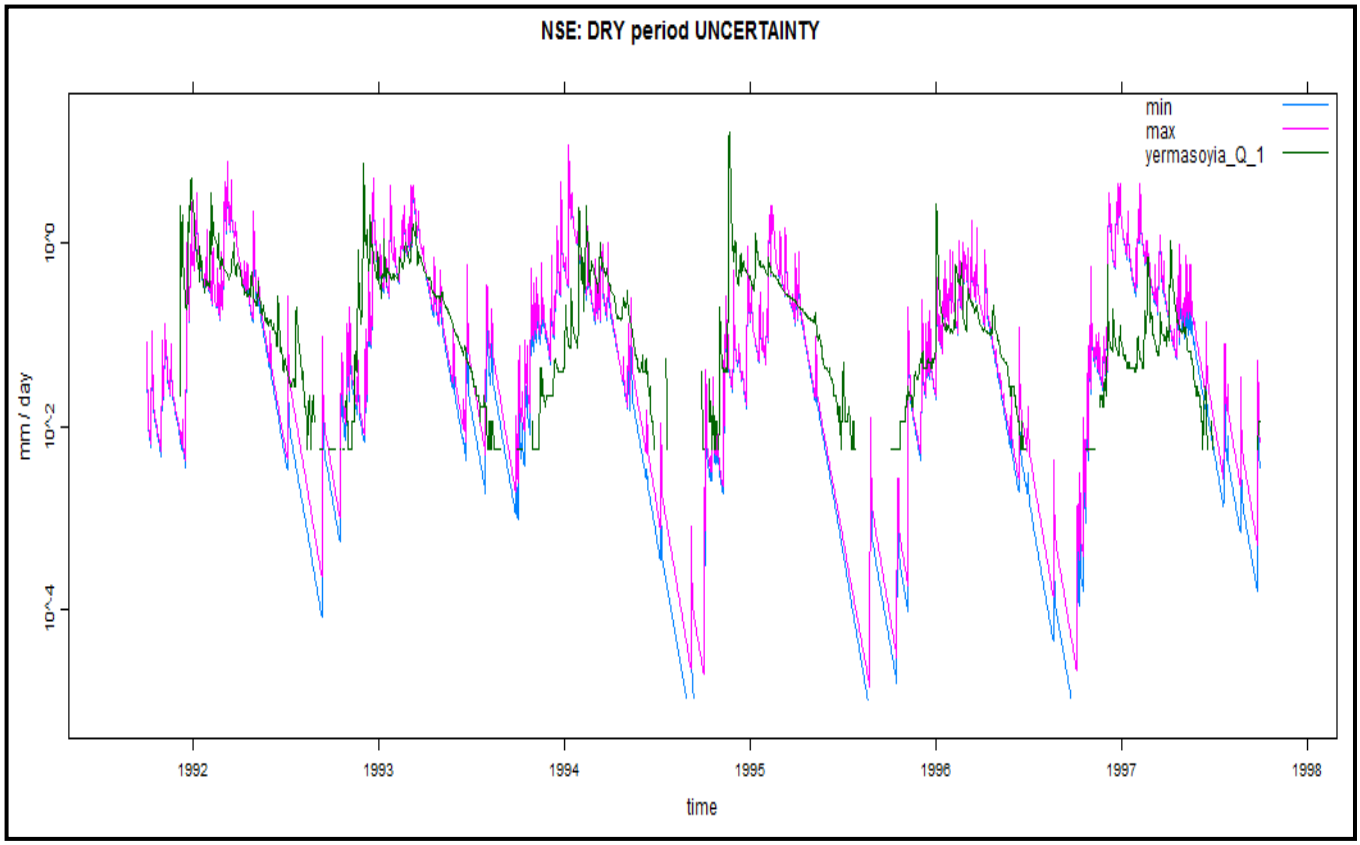
By NSE^3: dry period



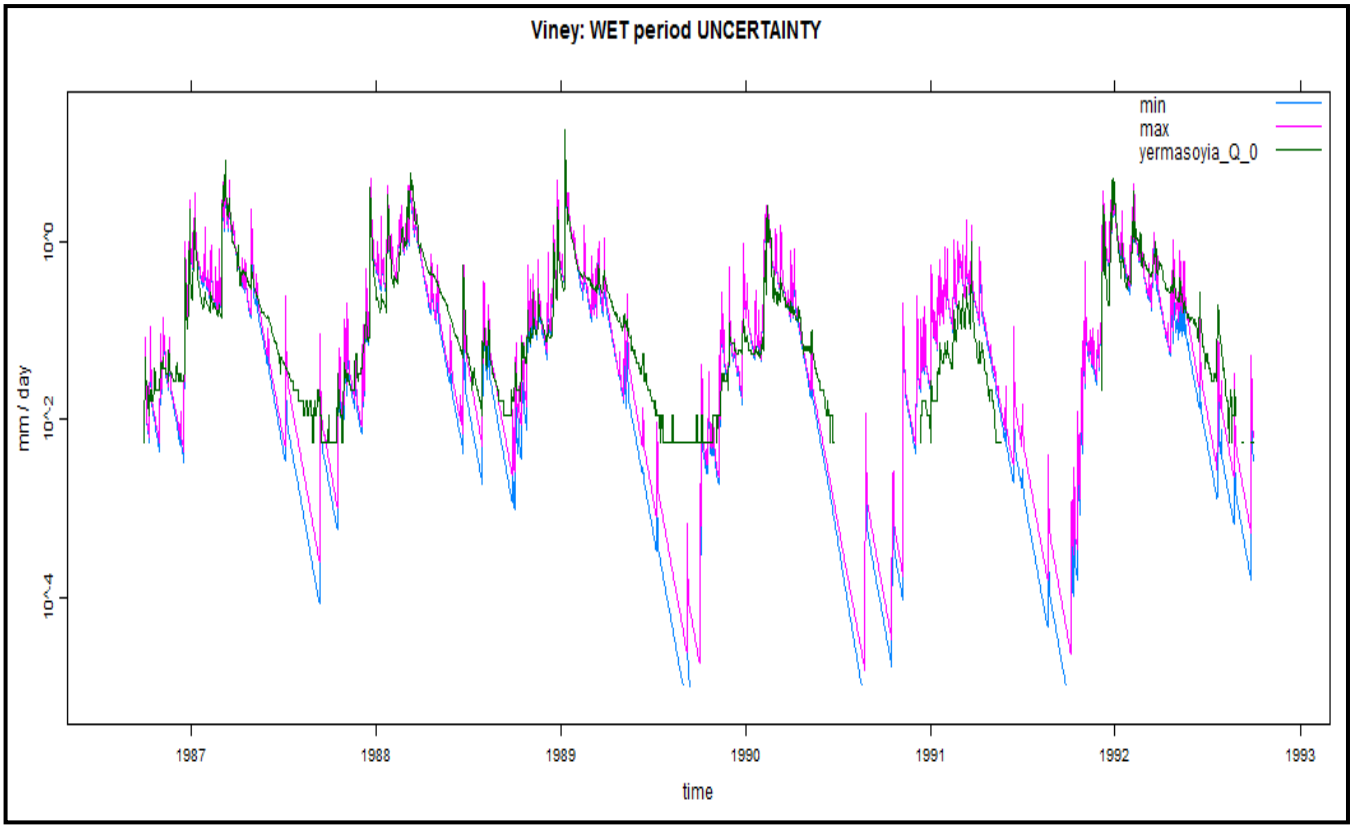
APPENDIX C

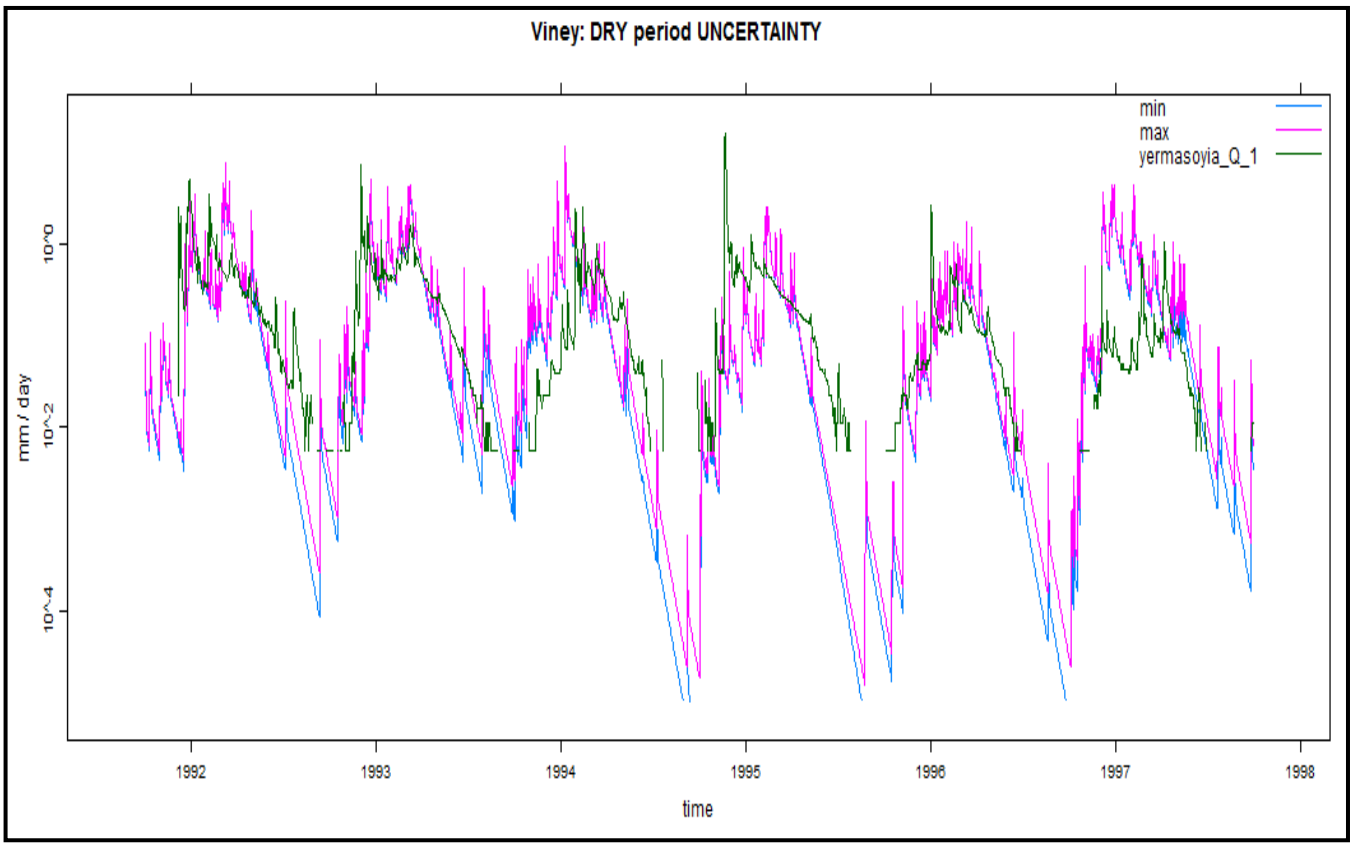
(D) Model structure uncertainty



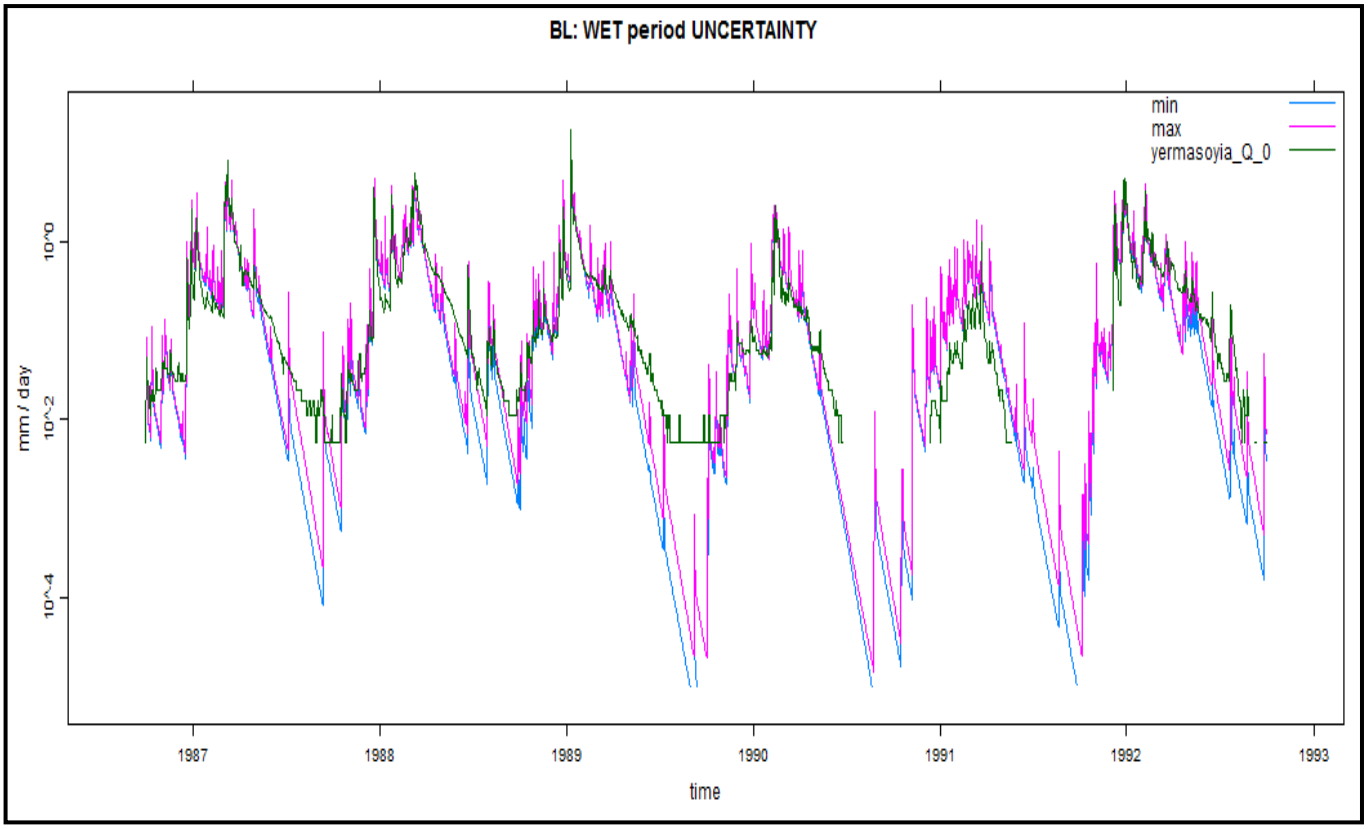


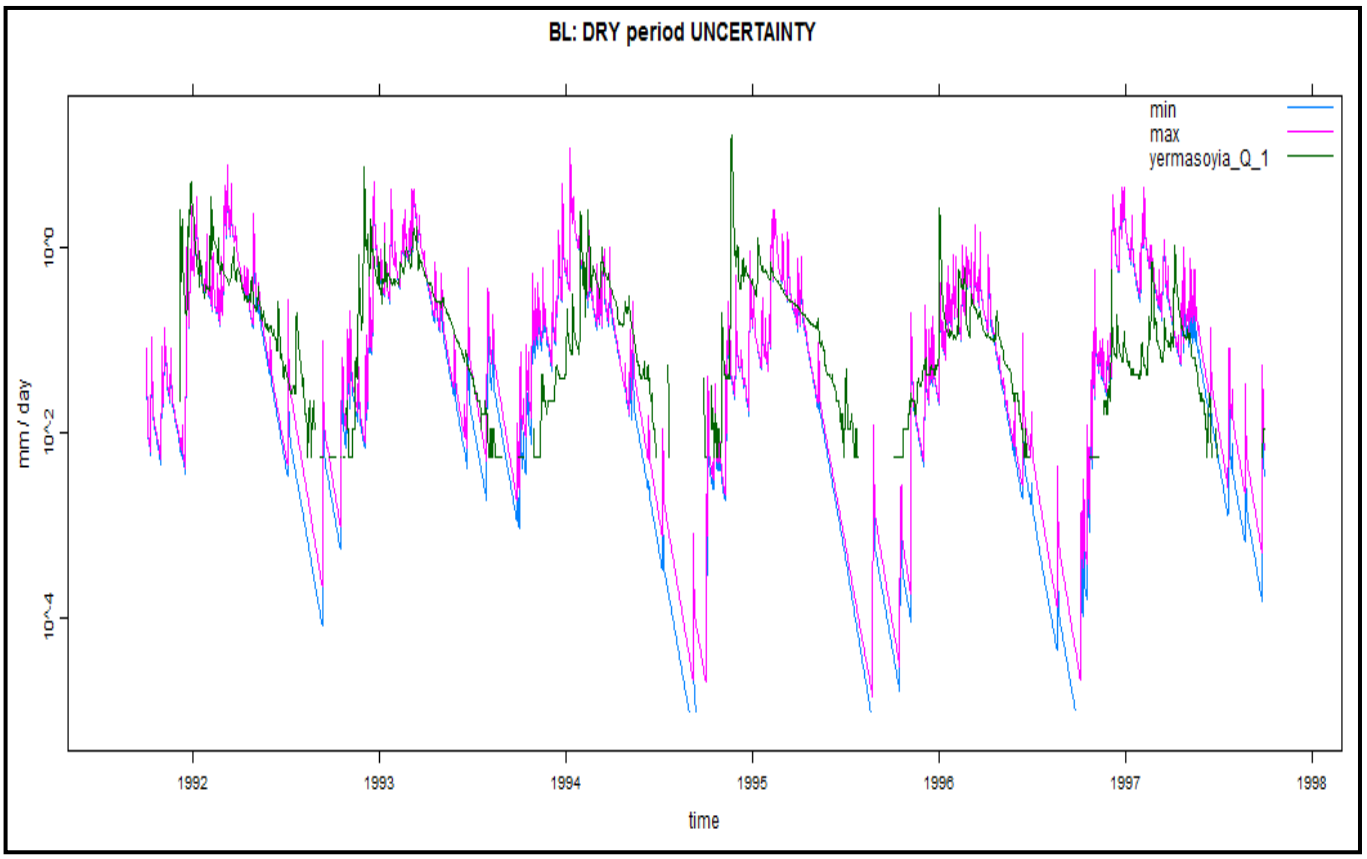
APPENDIX C



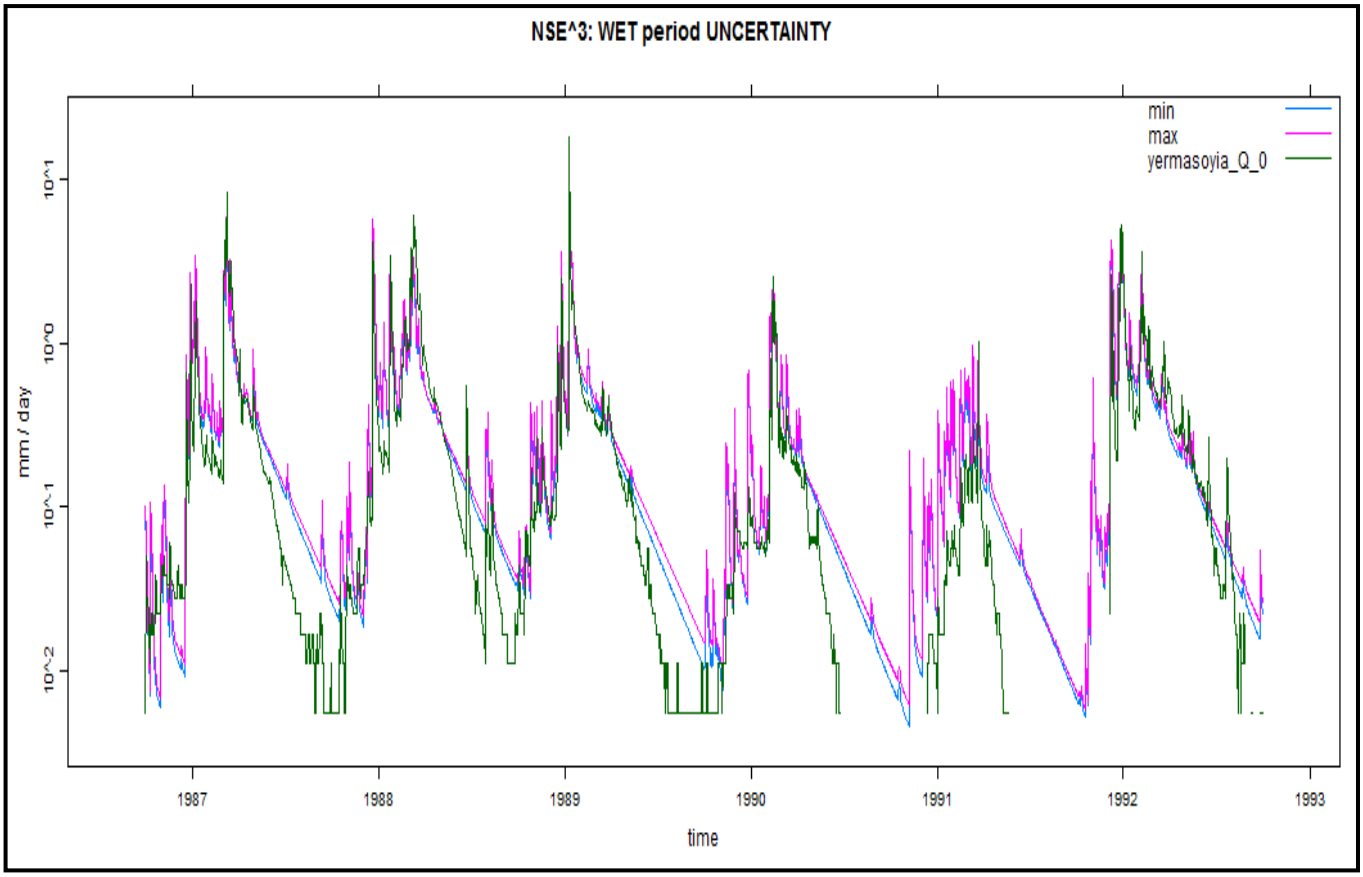


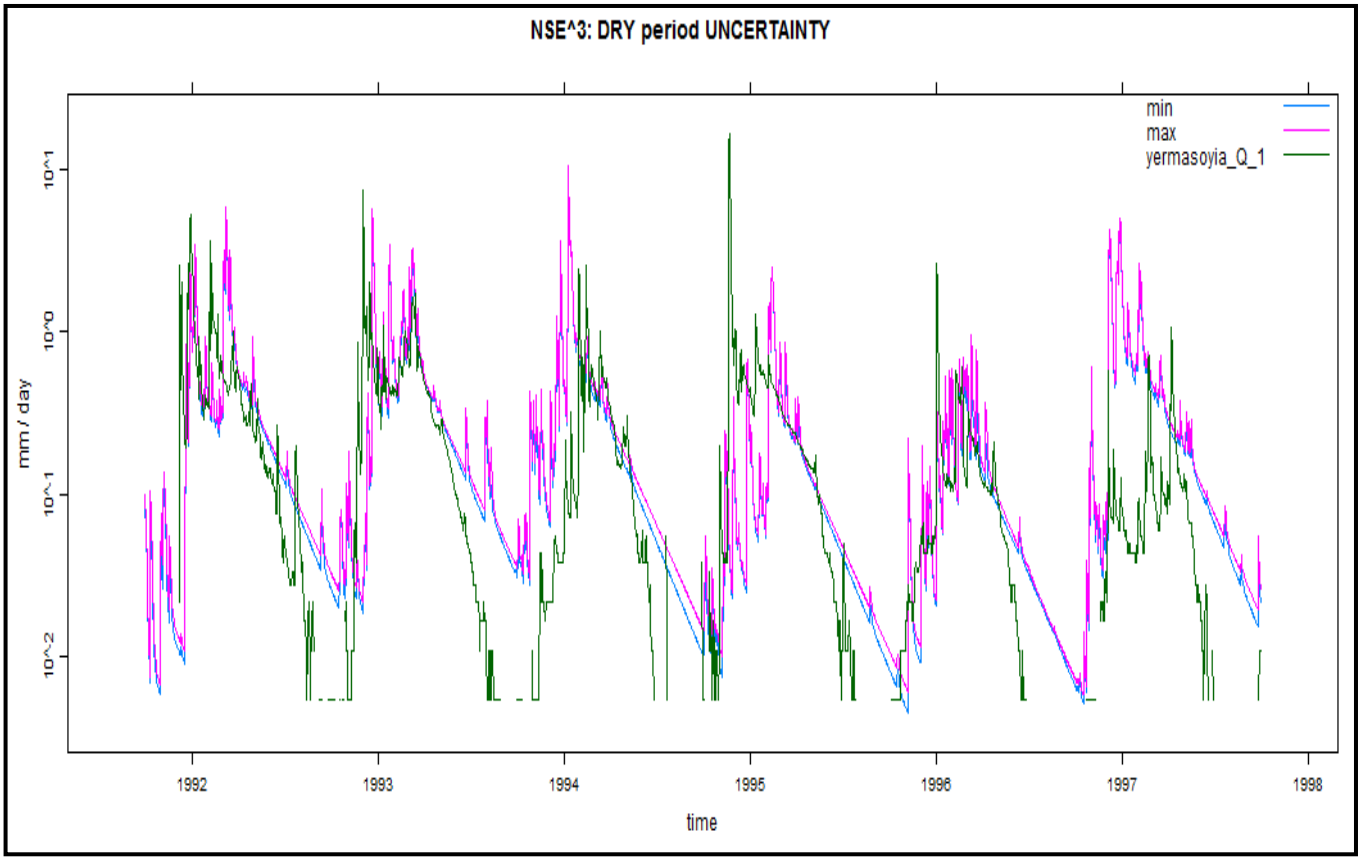
APPENDIX C





APPENDIX C



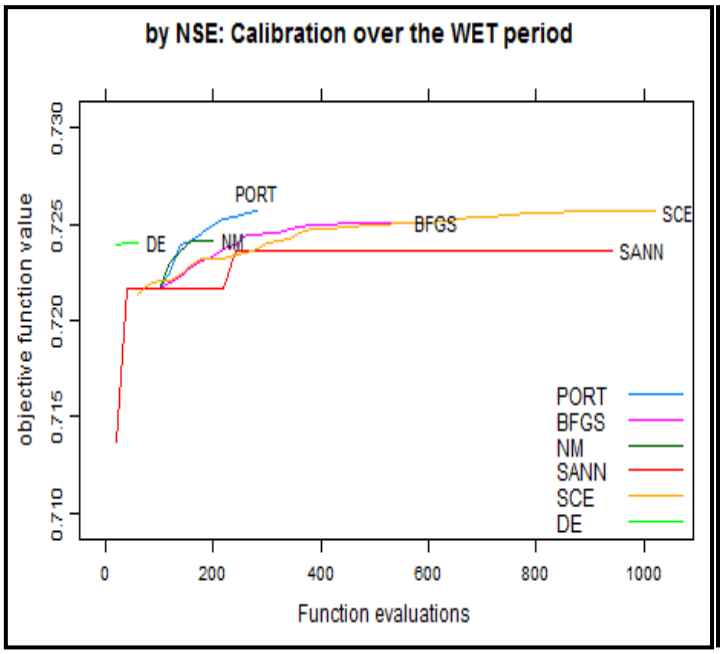


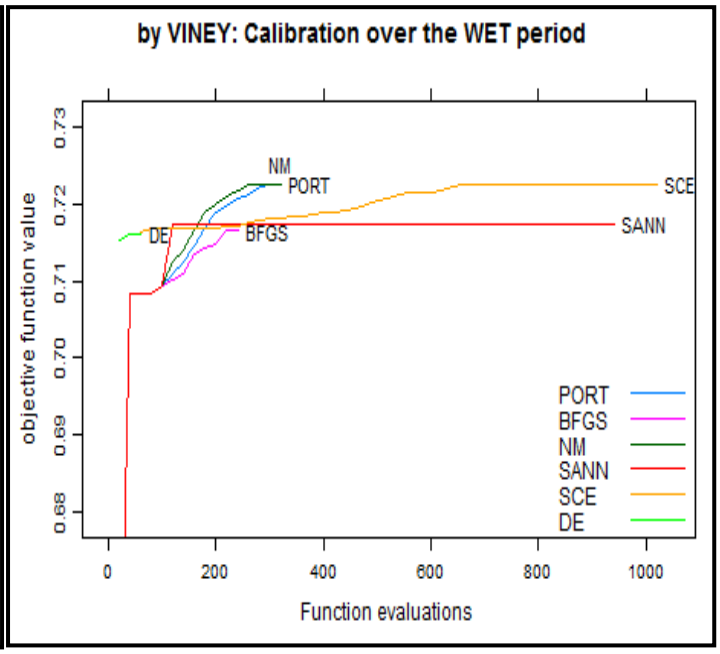
APPENDIX C

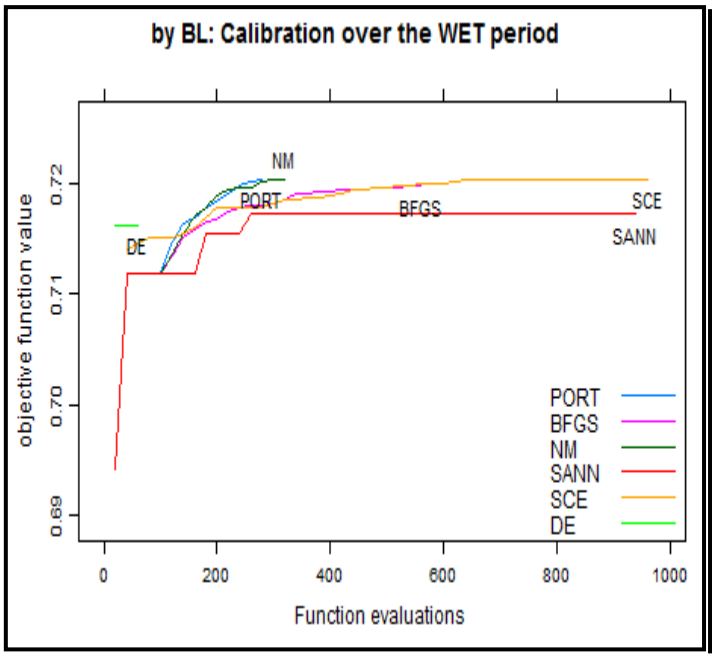
(E) Objective Function and Optimization Algorithm uncertainty

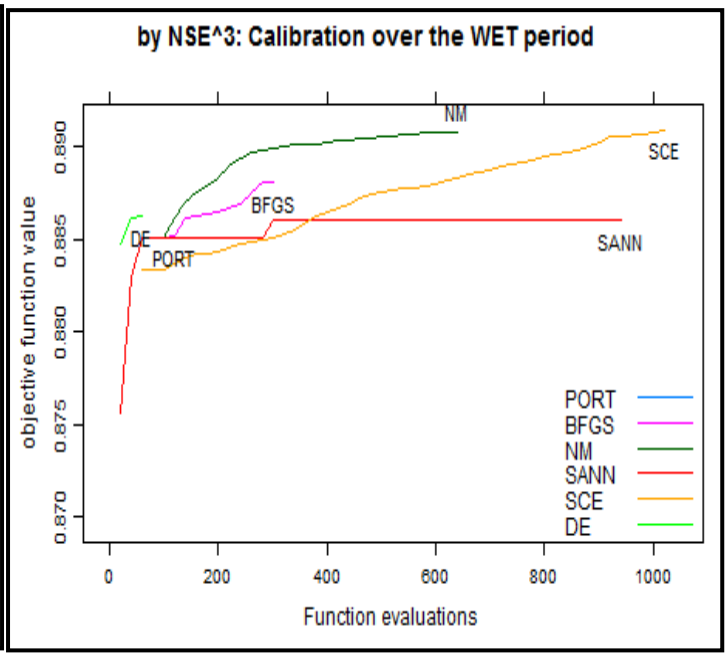
CALIBRATION: **WET** – VALIDATION: **DRY**

Optimization traces:









APPENDIX C

Validation FIT STATISTICS:

Objective Function = **NSE**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.169	0.751	0.651	0.602	0.634	0.823
BFGS	0.184	0.752	0.652	0.593	0.633	0.825
NM	0.200	0.746	0.656	0.585	0.636	0.832
SANN	0.187	0.748	0.664	0.603	0.645	0.846
SCE	0.169	0.751	0.651	0.602	0.634	0.823
DE	0.204	0.744	0.654	0.580	0.634	0.830

Objective Function = **Viney**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.168	0.750	0.654	0.606	0.637	0.827
BFGS	0.202	0.742	0.666	0.594	0.646	0.852
NM	0.167	0.749	0.656	0.609	0.639	0.831
SANN	0.198	0.748	0.646	0.576	0.626	0.816
SCE	0.168	0.750	0.654	0.606	0.637	0.827
DE	0.205	0.742	0.663	0.588	0.642	0.846

Objective Function = **BL**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.168	0.750	0.653	0.605	0.636	0.827
BFGS	0.174	0.751	0.654	0.603	0.637	0.829
NM	0.168	0.750	0.653	0.606	0.636	0.827
SANN	0.189	0.747	0.664	0.602	0.645	0.847
SCE	0.168	0.751	0.653	0.605	0.636	0.827
DE	0.204	0.744	0.653	0.578	0.633	0.828

Objective Function = **NSE^3**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.422	0.685	0.746	0.378	0.704	0.946
BFGS	0.408	0.693	0.753	0.410	0.712	0.948
NM	0.401	0.694	0.754	0.424	0.714	0.947
SANN	0.400	0.698	0.757	0.428	0.717	0.950
SCE	0.388	0.703	0.761	0.454	0.722	0.950
DE	0.400	0.698	0.756	0.427	0.716	0.950

APPENDIX C

Parameter Stability:

Objective Function = **NSE**

	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	delay
PORT	0.015	8.000	0.001	0	1	20	13.424	0.689	0.619	0.381	1
BFGS	0.121	6.133	0.001	0	1	20	13.064	0.684	0.622	0.378	1
NM	0.870	4.298	0.001	0	1	20	13.059	0.677	0.624	0.376	1
SANN	0.318	4.988	0.001	0	1	20	14.113	0.681	0.619	0.381	1
SCE	0.015	8.000	0.001	0	1	20	13.425	0.689	0.619	0.381	1
DE	1.067	4.156	0.001	0	1	20	12.914	0.677	0.625	0.375	1

Objective Function = **Viney**

	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	delay
PORT	0.014	8.000	0.001	0	1	20	13.718	0.689	0.618	0.382	1
BFGS	1.841	3.357	0.001	0	1	20	14.148	0.677	0.621	0.379	1
NM	0.013	8.000	0.001	0	1	20	13.951	0.689	0.617	0.383	1
SANN	0.358	5.309	0.001	0	1	20	12.554	0.682	0.626	0.374	1
SCE	0.014	8.000	0.001	0	1	20	13.717	0.689	0.618	0.382	1
DE	2.084	3.343	0.001	0	1	20	13.630	0.675	0.623	0.377	1

Objective Function = **BL**

	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	delay
PORT	0.014	8.000	0.001	0	1	20	13.654	0.689	0.618	0.382	1
BFGS	0.038	7.092	0.001	0	1	20	13.535	0.686	0.619	0.381	1
NM	0.014	8.000	0.001	0	1	20	13.671	0.689	0.618	0.382	1
SANN	0.435	4.691	0.001	0	1	20	14.160	0.680	0.619	0.381	1
SCE	0.014	8.000	0.001	0	1	20	13.650	0.689	0.618	0.382	1
DE	1.067	4.175	0.001	0	1	20	12.858	0.677	0.625	0.375	1

Objective Function = **NSE^3**

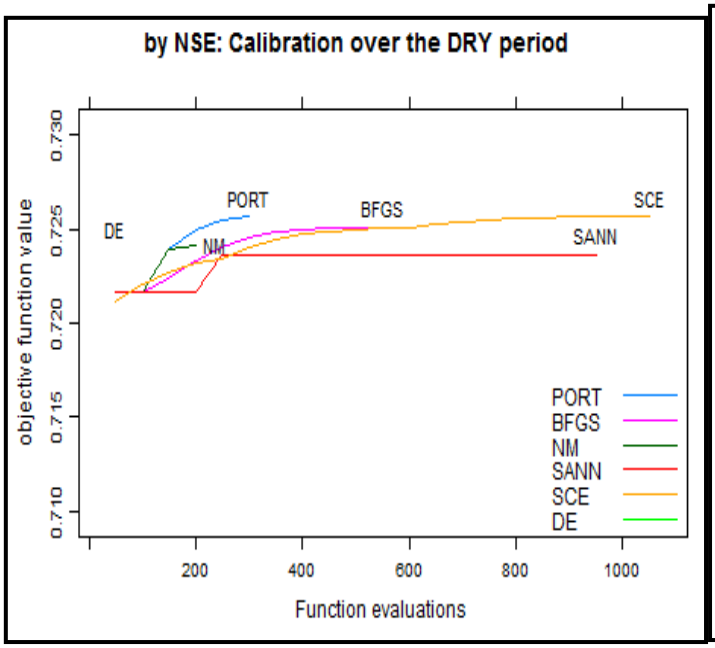
	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	Delay
PORT	3.028	1.817	0.002	0	1	20	57.070	2.505	0.482	0.518	0
BFGS	0.177	4.340	0.002	0	1	20	58.931	2.482	0.491	0.509	0
NM	0.003	7.999	0.002	0	1	20	61.528	2.406	0.515	0.485	0
SANN	0.959	2.983	0.002	0	1	20	56.843	2.557	0.467	0.533	0
SCE	0.005	7.488	0.002	0	1	20	60.319	2.465	0.496	0.504	0
DE	0.827	3.104	0.002	0	1	20	57.034	2.551	0.469	0.531	0

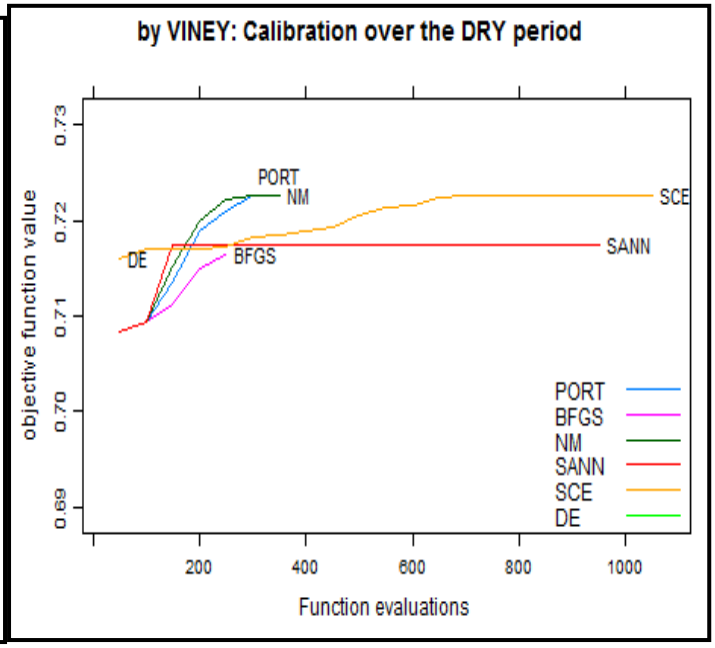
APPENDIX C

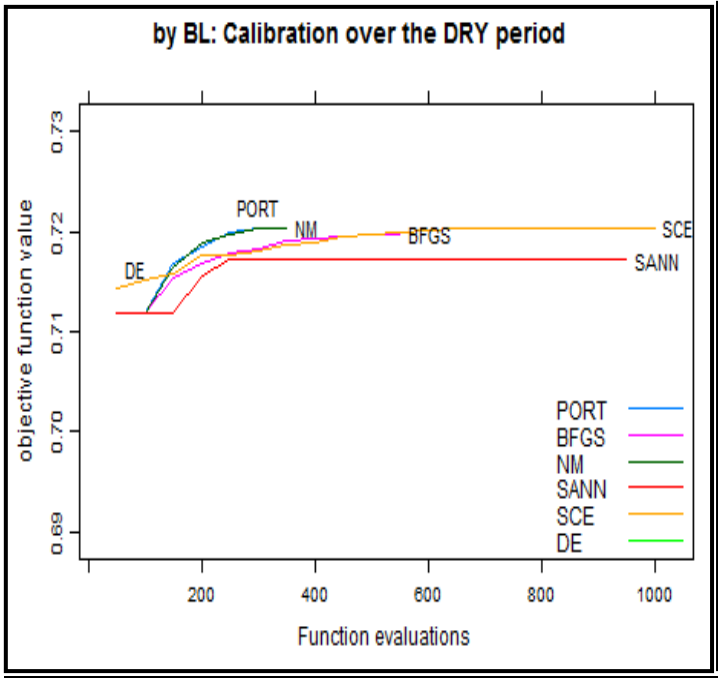
(F) Objective Function and Optimization Algorithm uncertainty

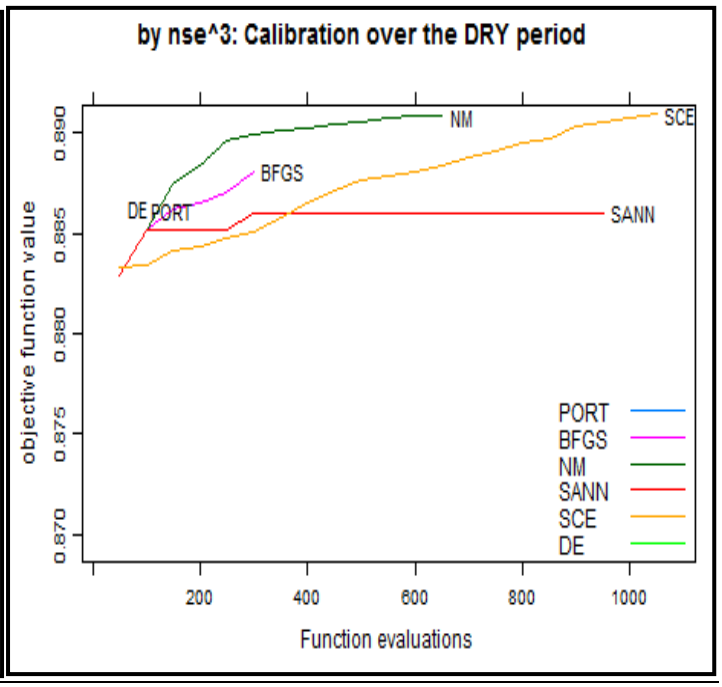
CALIBRATION: **DRY** – VALIDATION: **WET**

Optimization traces:









APPENDIX C

Validation FIT STATISTICS:

Objective Function = **NSE**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.053	0.802	0.726	0.723	0.720	0.788
BFGS	0.062	0.806	0.725	0.721	0.719	0.787
NM	0.073	0.805	0.724	0.718	0.717	0.788
SANN	0.061	0.802	0.724	0.719	0.718	0.790
SCE	0.053	0.802	0.726	0.723	0.720	0.788
DE	0.077	0.805	0.724	0.717	0.716	0.788

Objective Function = **Viney**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.052	0.801	0.726	0.723	0.720	0.789
BFGS	0.069	0.802	0.722	0.716	0.715	0.790
NM	0.052	0.800	0.725	0.723	0.720	0.790
SANN	0.074	0.808	0.724	0.718	0.717	0.786
SCE	0.052	0.801	0.726	0.723	0.720	0.789
DE	0.073	0.803	0.723	0.716	0.716	0.789

Objective Function = **BL**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.053	0.801	0.726	0.723	0.720	0.789
BFGS	0.056	0.802	0.725	0.722	0.720	0.789
NM	0.053	0.801	0.726	0.723	0.720	0.789
SANN	0.062	0.802	0.723	0.719	0.717	0.790
SCE	0.053	0.801	0.726	0.723	0.720	0.789
DE	0.078	0.805	0.724	0.716	0.716	0.788

Objective Function = **NSE^3**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.202	0.747	0.707	0.634	0.686	0.885
BFGS	0.202	0.749	0.708	0.636	0.688	0.888
NM	0.200	0.749	0.706	0.635	0.686	0.891
SANN	0.202	0.753	0.711	0.639	0.691	0.886
SCE	0.201	0.753	0.711	0.639	0.691	0.891
DE	0.202	0.753	0.711	0.638	0.691	0.886

APPENDIX C

Parameter Stability:

Objective Function = **NSE**

Parameters calibrated over **Dry period**:

	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	delay
PORT	0.015	8.000	0.001	0	1	20	13.424	0.689	0.619	0.381	1
BFGS	0.121	6.133	0.001	0	1	20	13.064	0.684	0.622	0.378	1
NM	0.870	4.298	0.001	0	1	20	13.059	0.677	0.624	0.376	1
SANN	0.318	4.988	0.001	0	1	20	14.113	0.681	0.619	0.381	1
SCE	0.015	8.000	0.001	0	1	20	13.425	0.689	0.619	0.381	1
DE	1.067	4.156	0.001	0	1	20	12.914	0.677	0.625	0.375	1

Objective Function = **Viney**

	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	delay
PORT	0.014	8.000	0.001	0	1	20	13.718	0.689	0.618	0.382	1
BFGS	1.841	3.357	0.001	0	1	20	14.148	0.677	0.621	0.379	1
NM	0.013	8.000	0.001	0	1	20	13.951	0.689	0.617	0.383	1
SANN	0.358	5.309	0.001	0	1	20	12.554	0.682	0.626	0.374	1
SCE	0.014	8.000	0.001	0	1	20	13.717	0.689	0.618	0.382	1
DE	2.084	3.343	0.001	0	1	20	13.630	0.675	0.623	0.377	1

Objective Function = **BL**

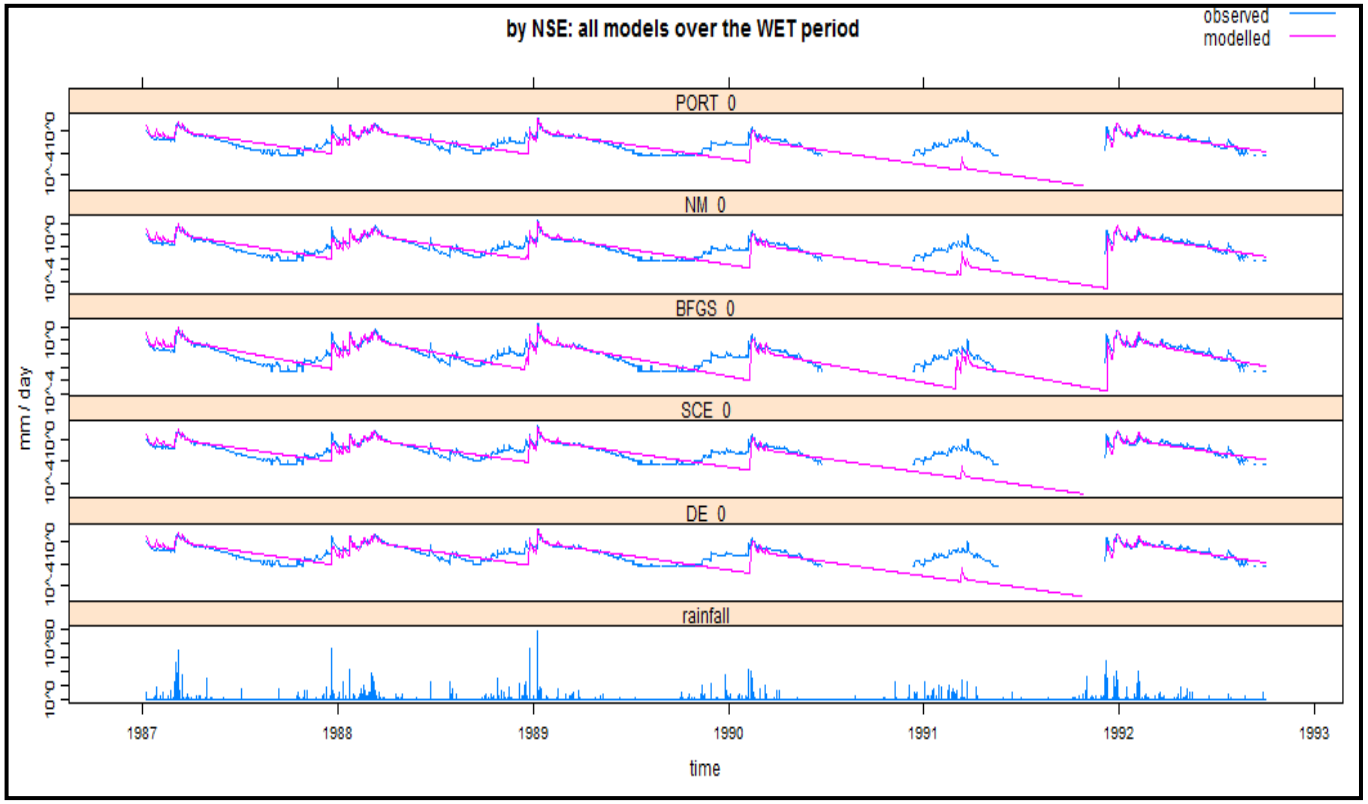
	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	delay
PORT	0.014	8.000	0.001	0	1	20	13.654	0.689	0.618	0.382	1
BFGS	0.038	7.092	0.001	0	1	20	13.535	0.686	0.619	0.381	1
NM	0.014	8.000	0.001	0	1	20	13.671	0.689	0.618	0.382	1
SANN	0.435	4.691	0.001	0	1	20	14.160	0.680	0.619	0.381	1
SCE	0.014	8.000	0.001	0	1	20	13.650	0.689	0.618	0.382	1
DE	1.067	4.175	0.001	0	1	20	12.858	0.677	0.625	0.375	1

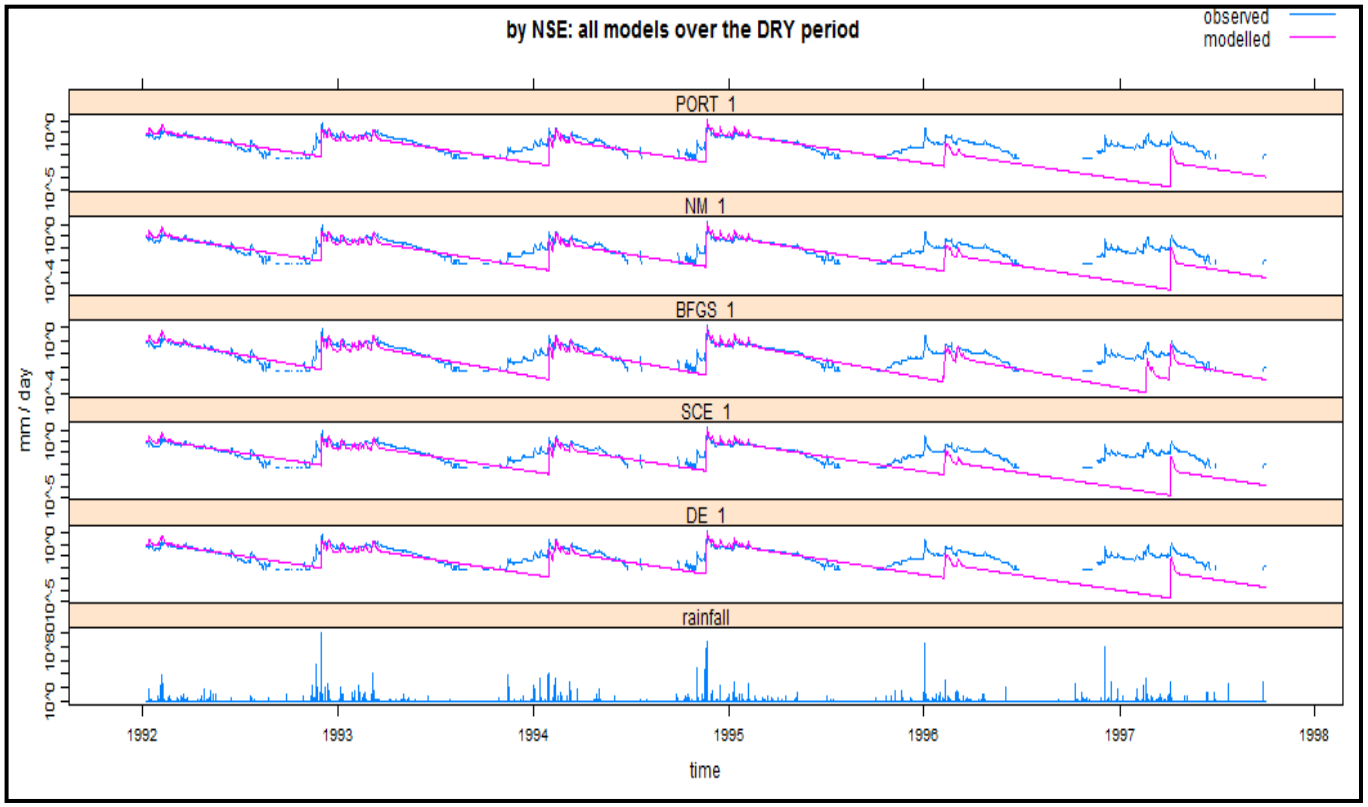
Objective Function = **NSE^3**

	tw	f	scale	l	p	t_ref	tau_s	tau_q	v_s	v_q	Delay
PORT	3.028	1.817	0.002	0	1	20	57.070	2.505	0.482	0.518	0
BFGS	0.177	4.340	0.002	0	1	20	58.931	2.482	0.491	0.509	0
NM	0.003	7.999	0.002	0	1	20	61.528	2.406	0.515	0.485	0
SANN	0.959	2.983	0.002	0	1	20	56.843	2.557	0.467	0.533	0
SCE	0.005	7.488	0.002	0	1	20	60.319	2.465	0.496	0.504	0
DE	0.827	3.104	0.002	0	1	20	57.034	2.551	0.469	0.531	0

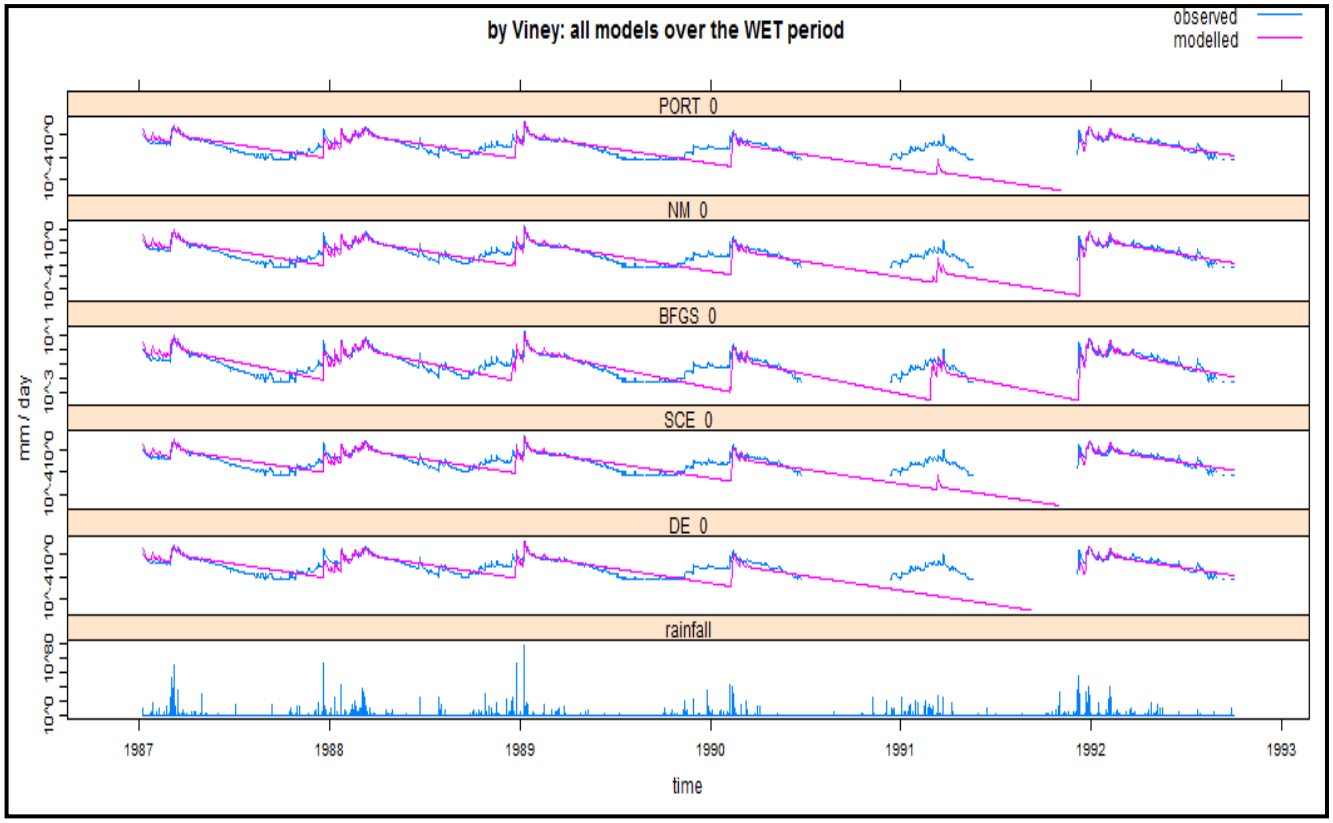
IHACRES model – CMD version

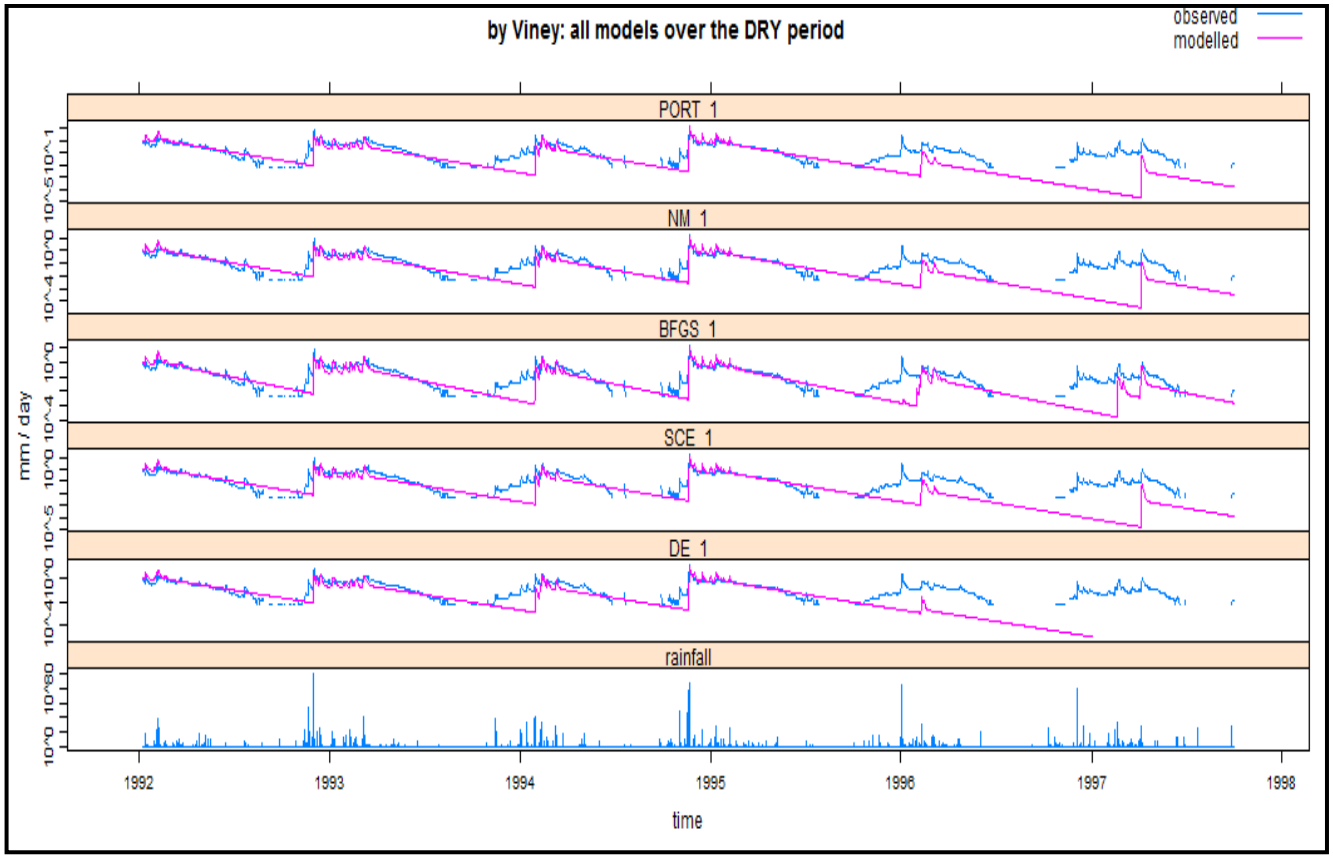
(A) Simulated streamflows (in log scale) – all models



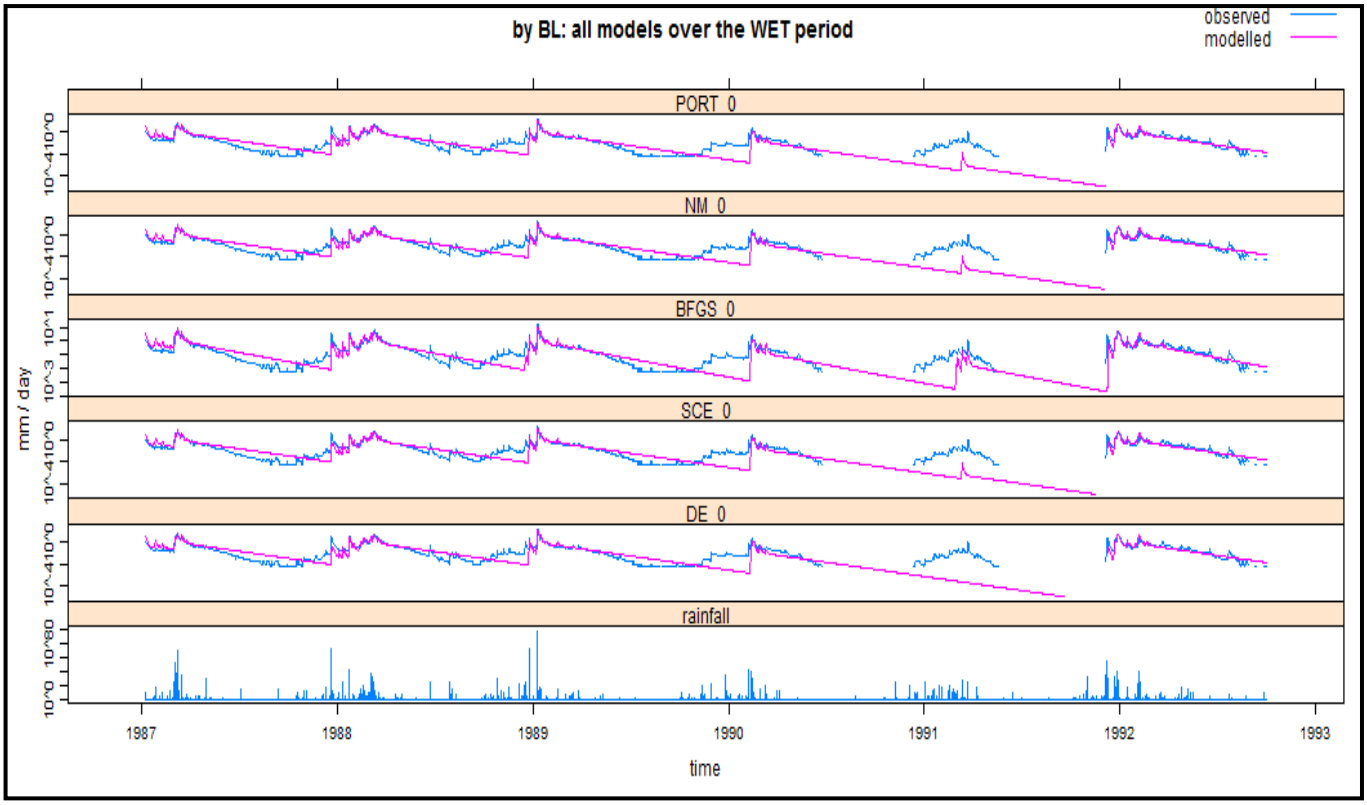


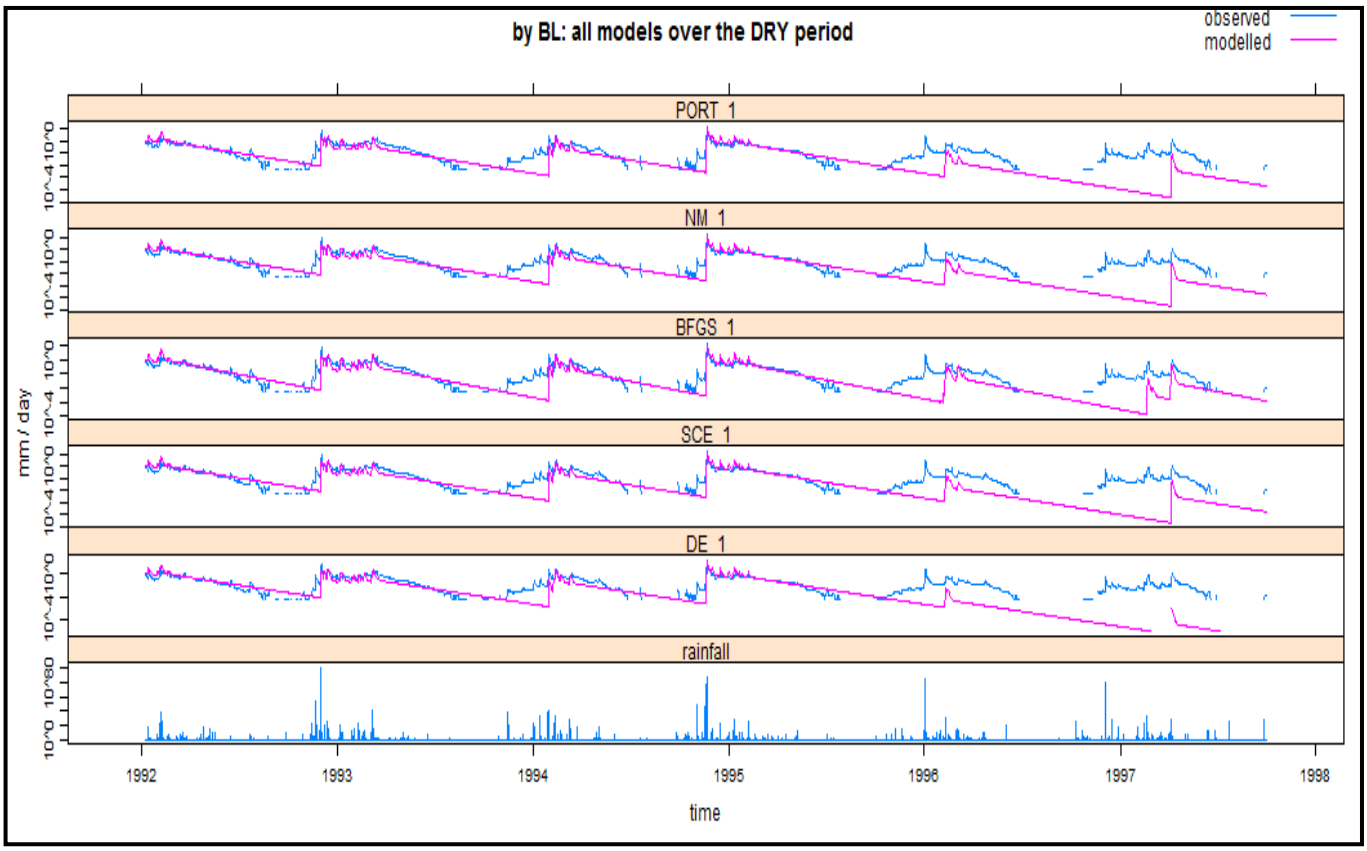
APPENDIX D



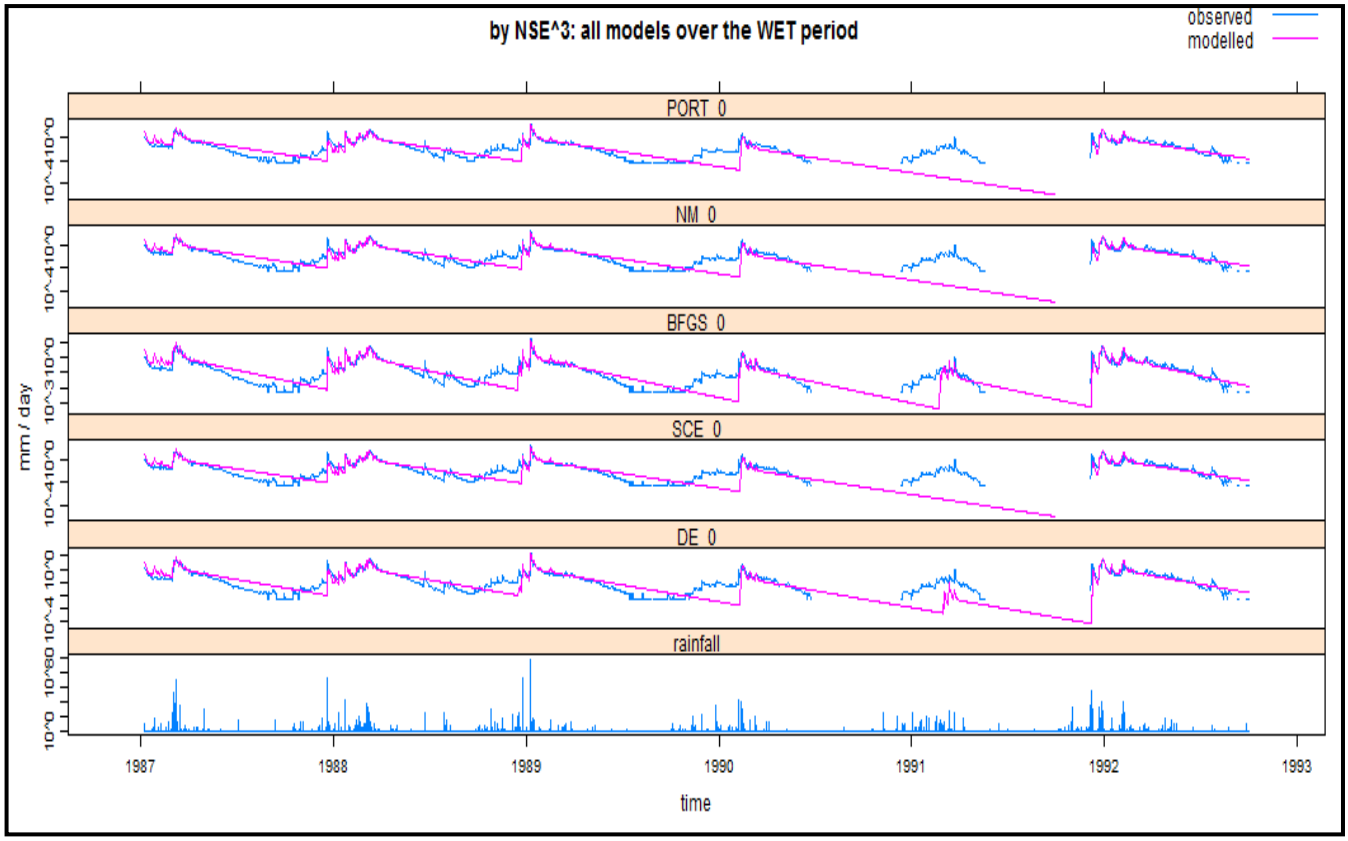


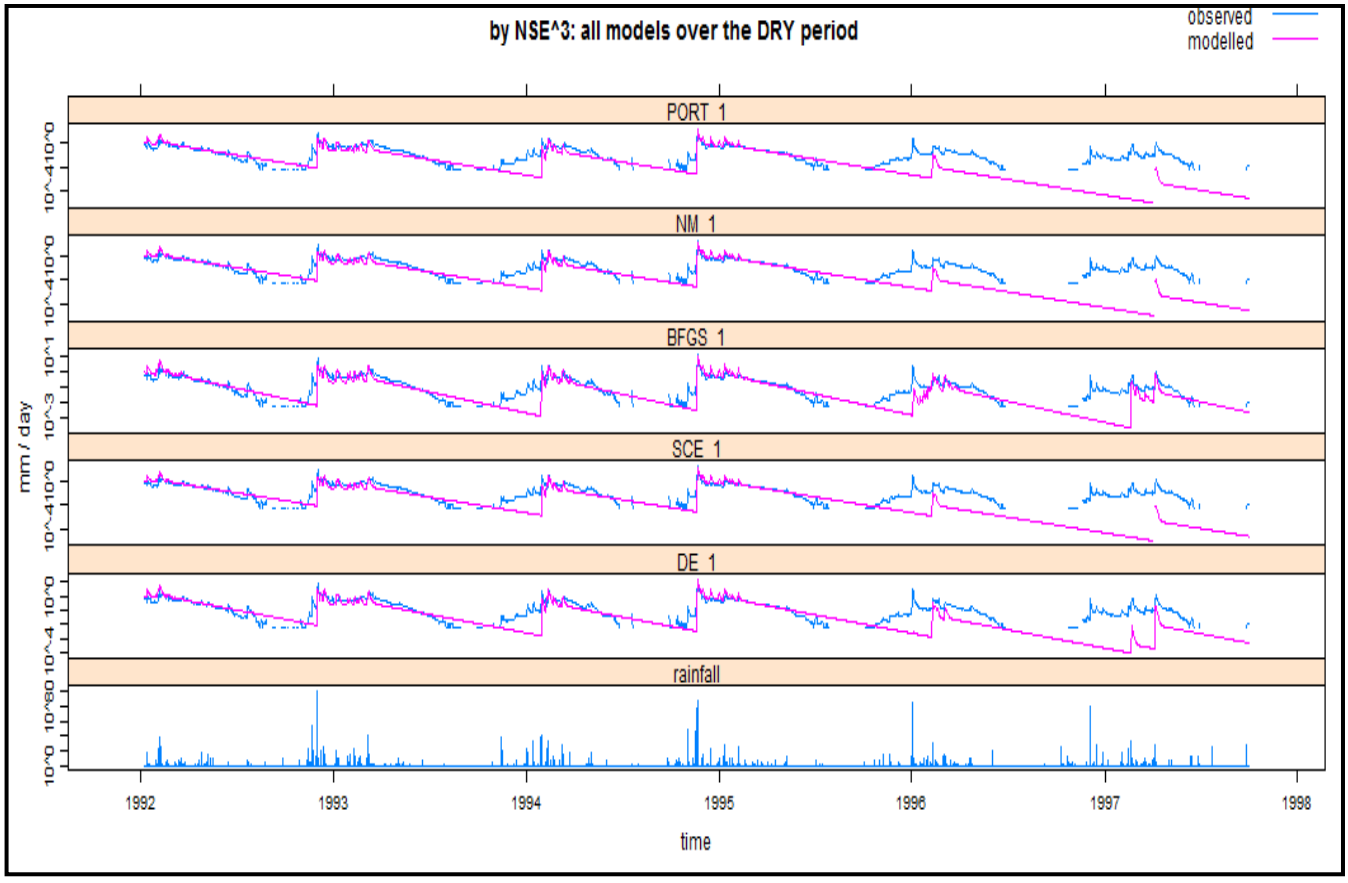
APPENDIX D





APPENDIX D

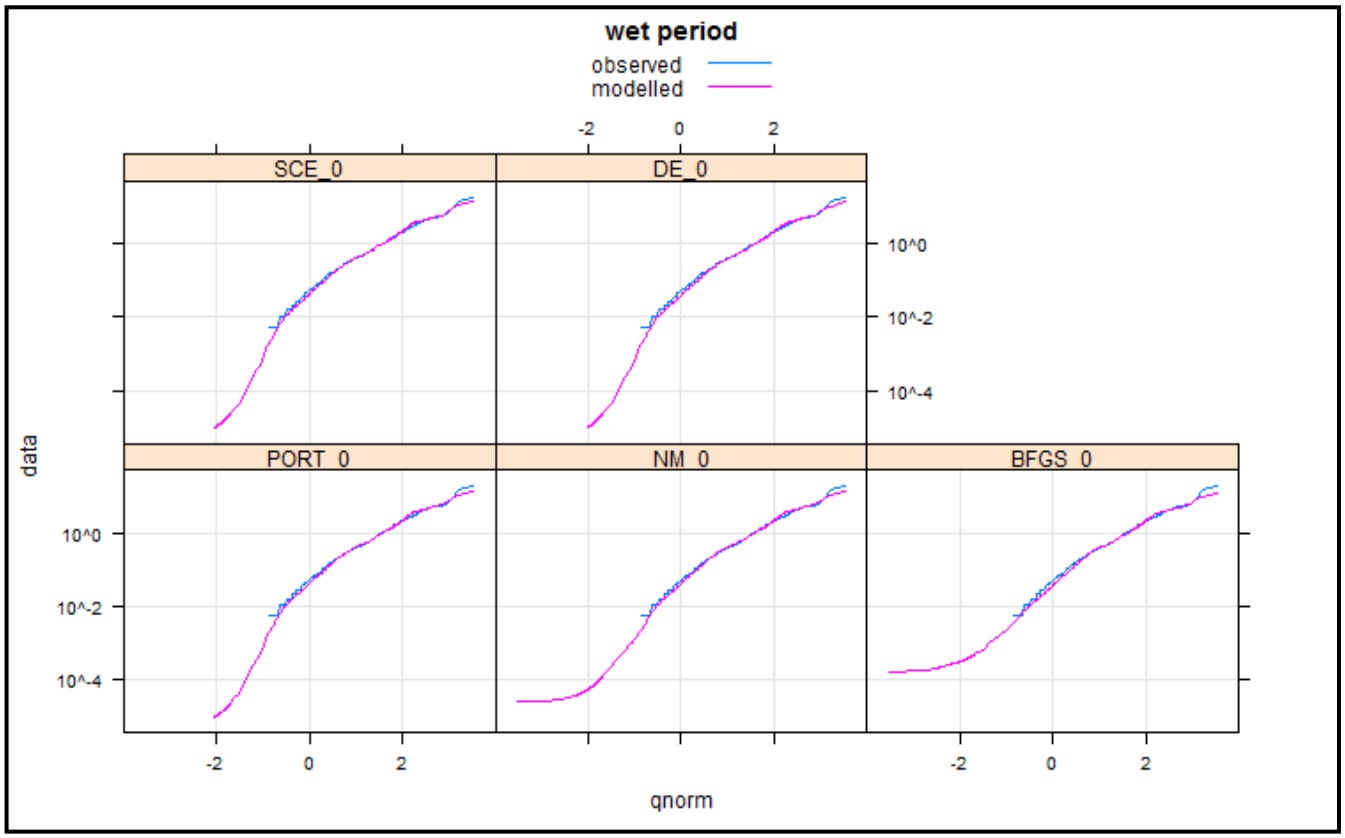


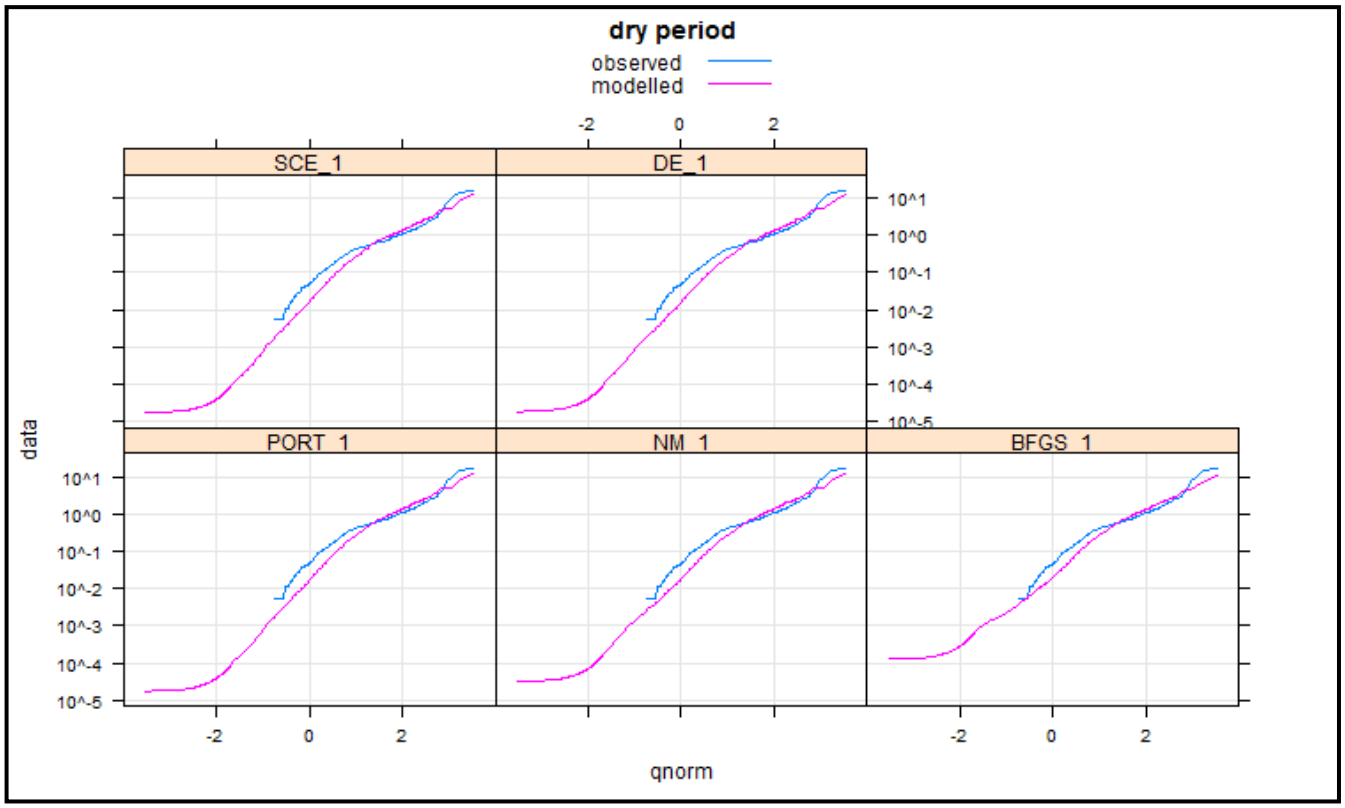


APPENDIX D

(B) Normal distribution Q-Q plot – all models

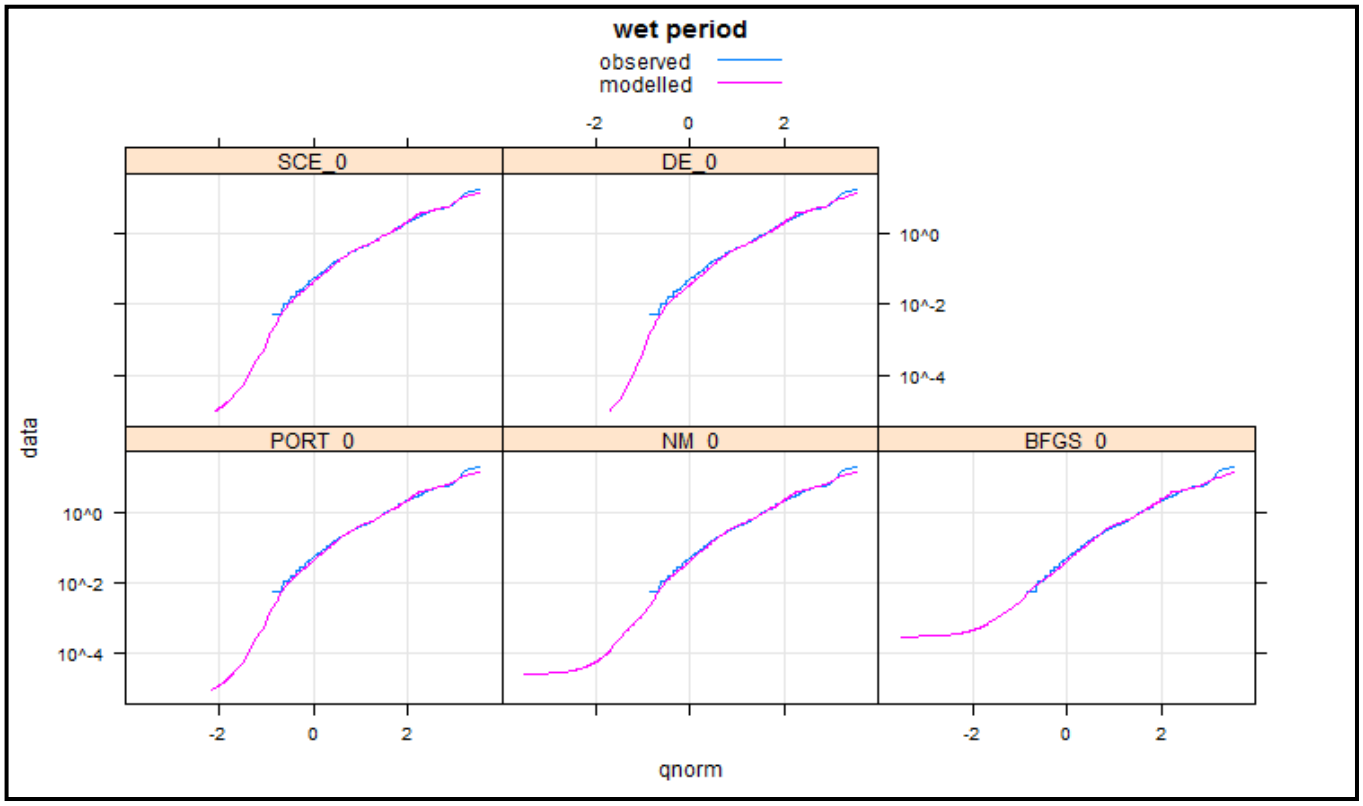
By NSE:

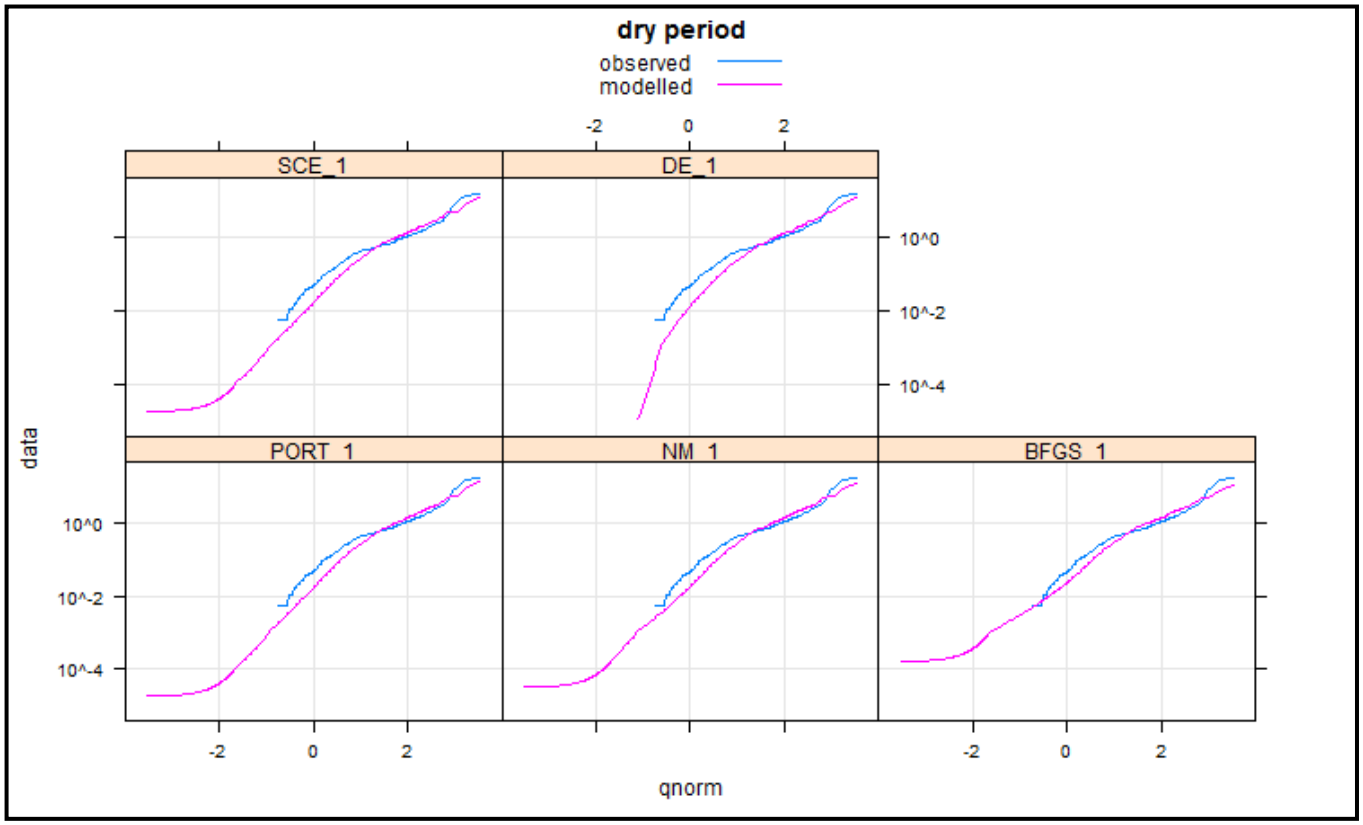




APPENDIX D

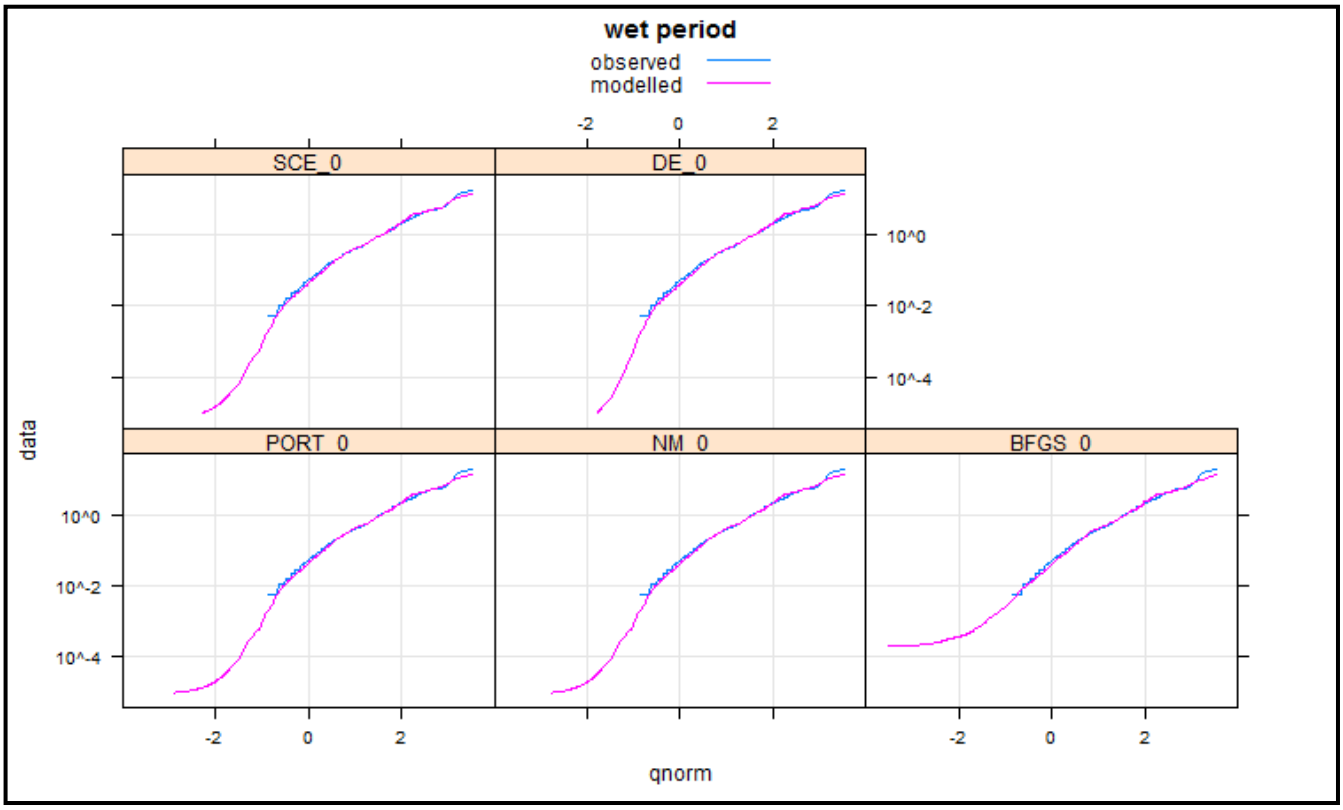
By Viney:

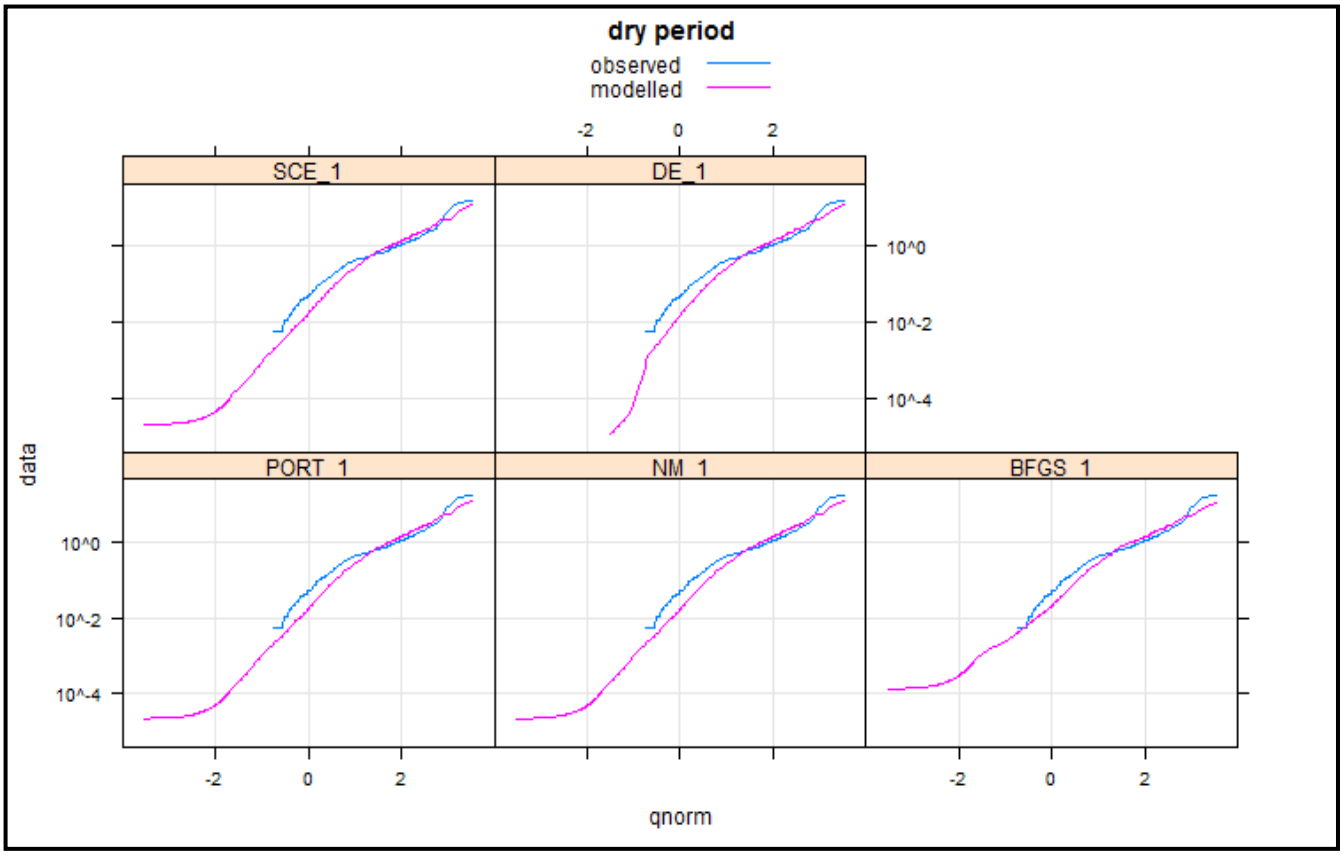




APPENDIX D

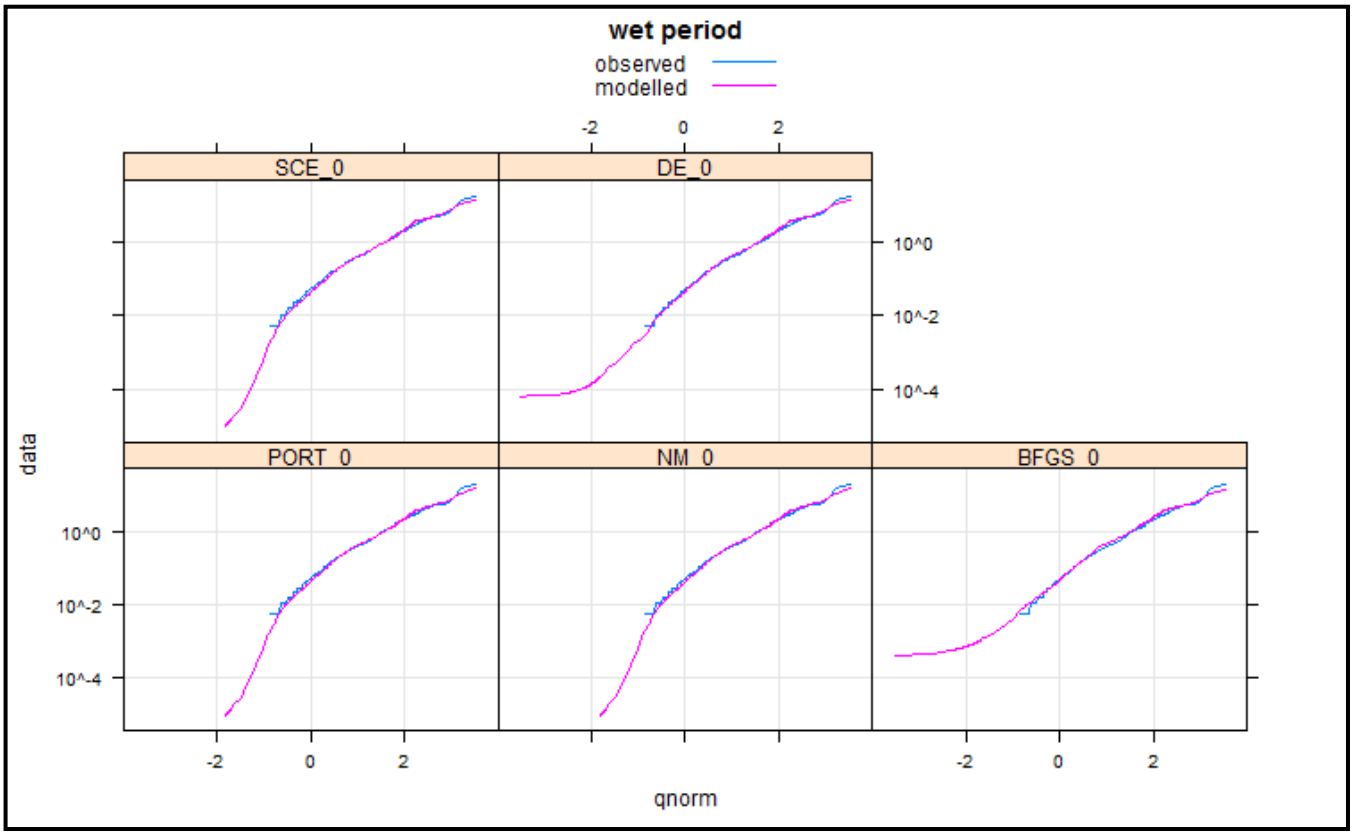
By BL:

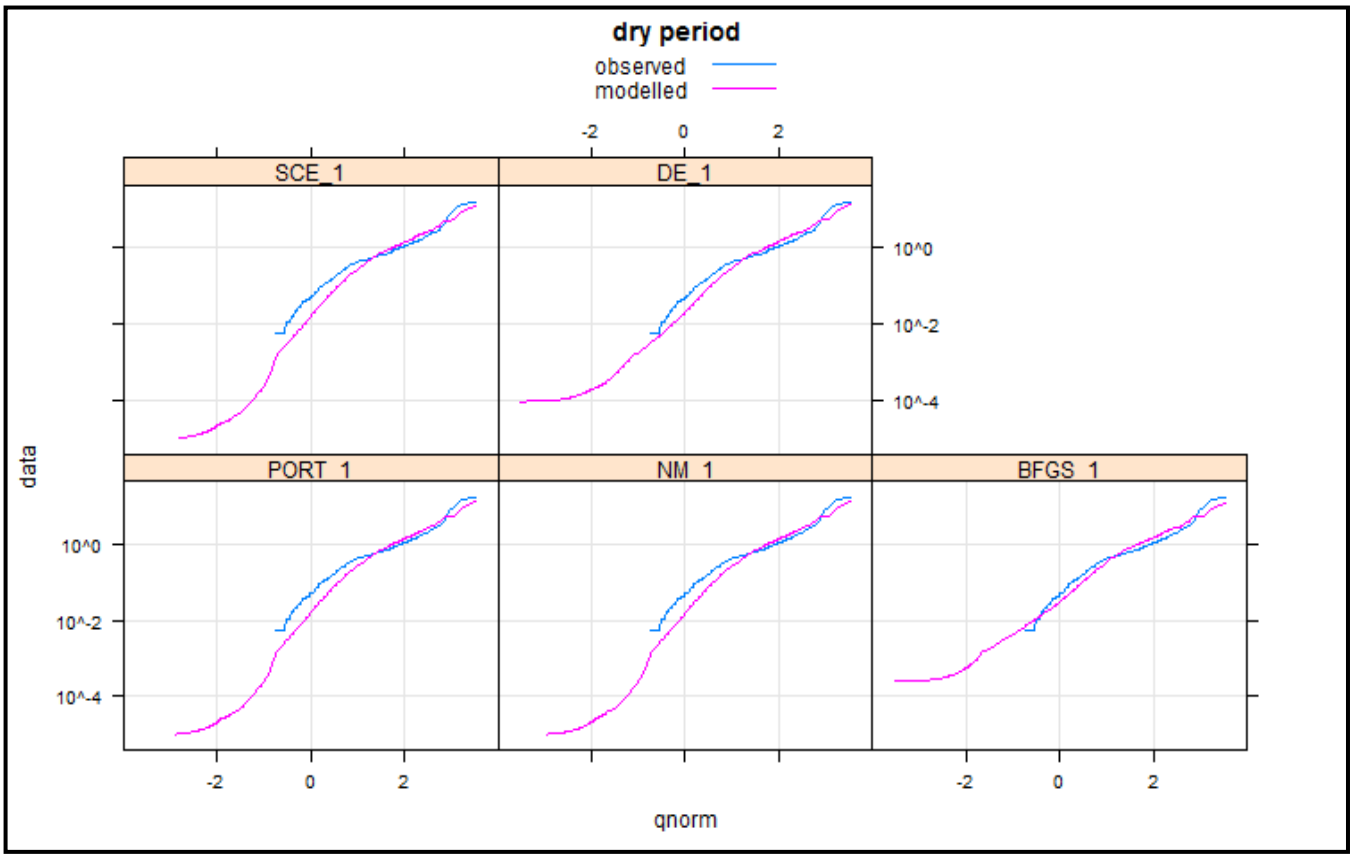




APPENDIX D

By NSE^3:

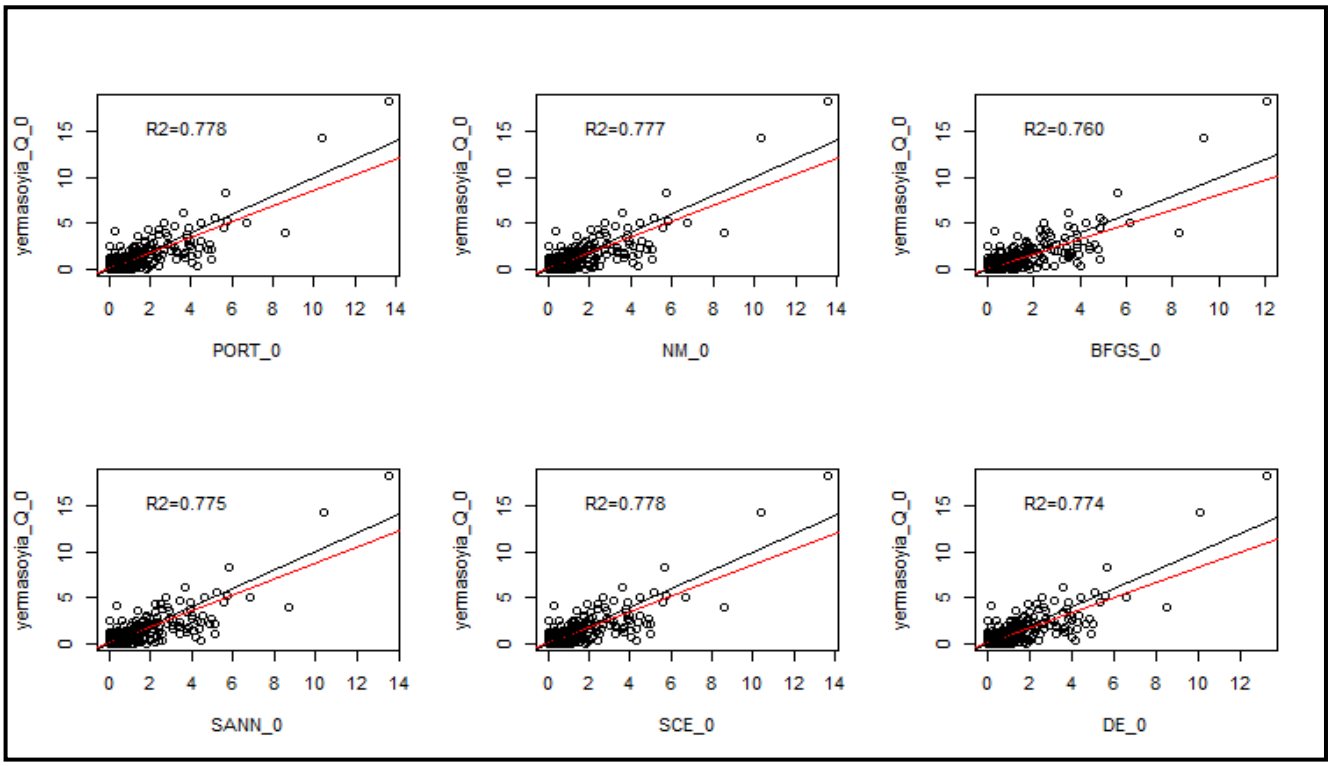




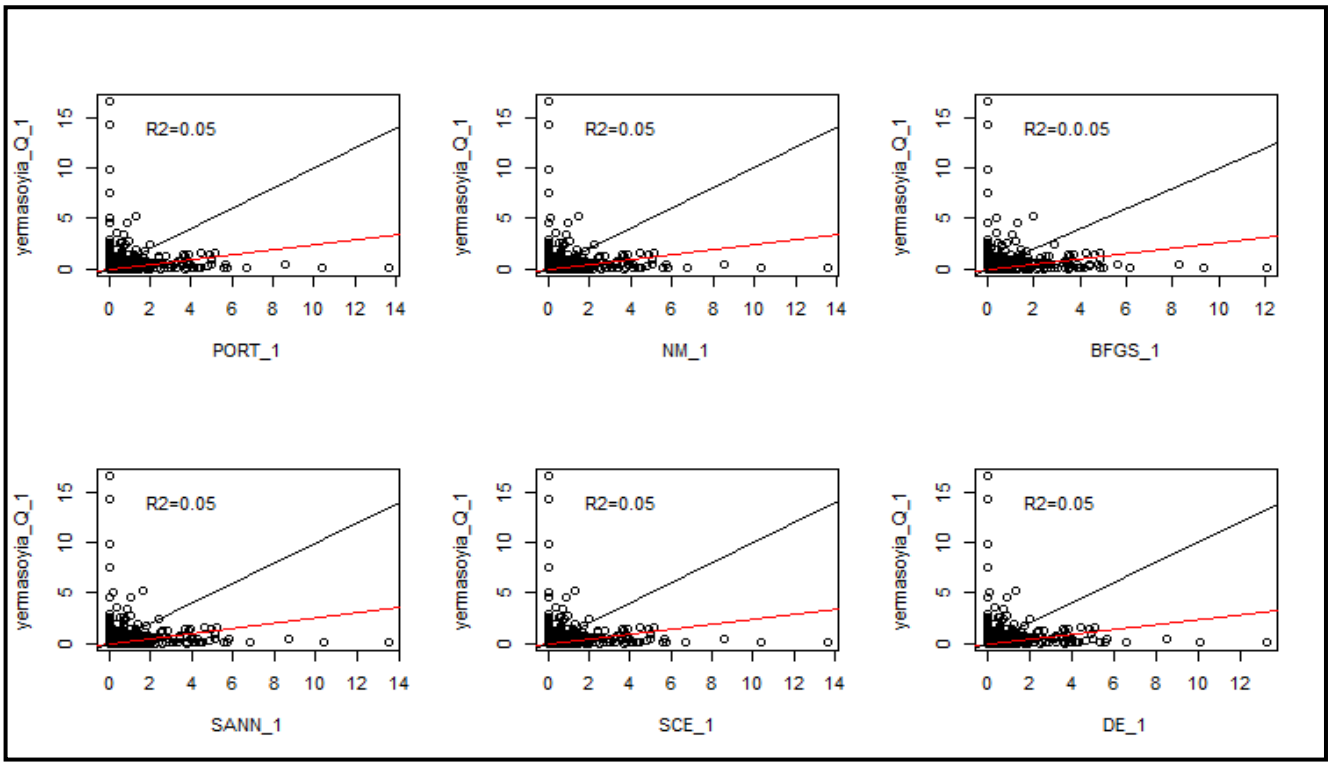
APPENDIX D

(C) Scatterplots

By NSE: wet period

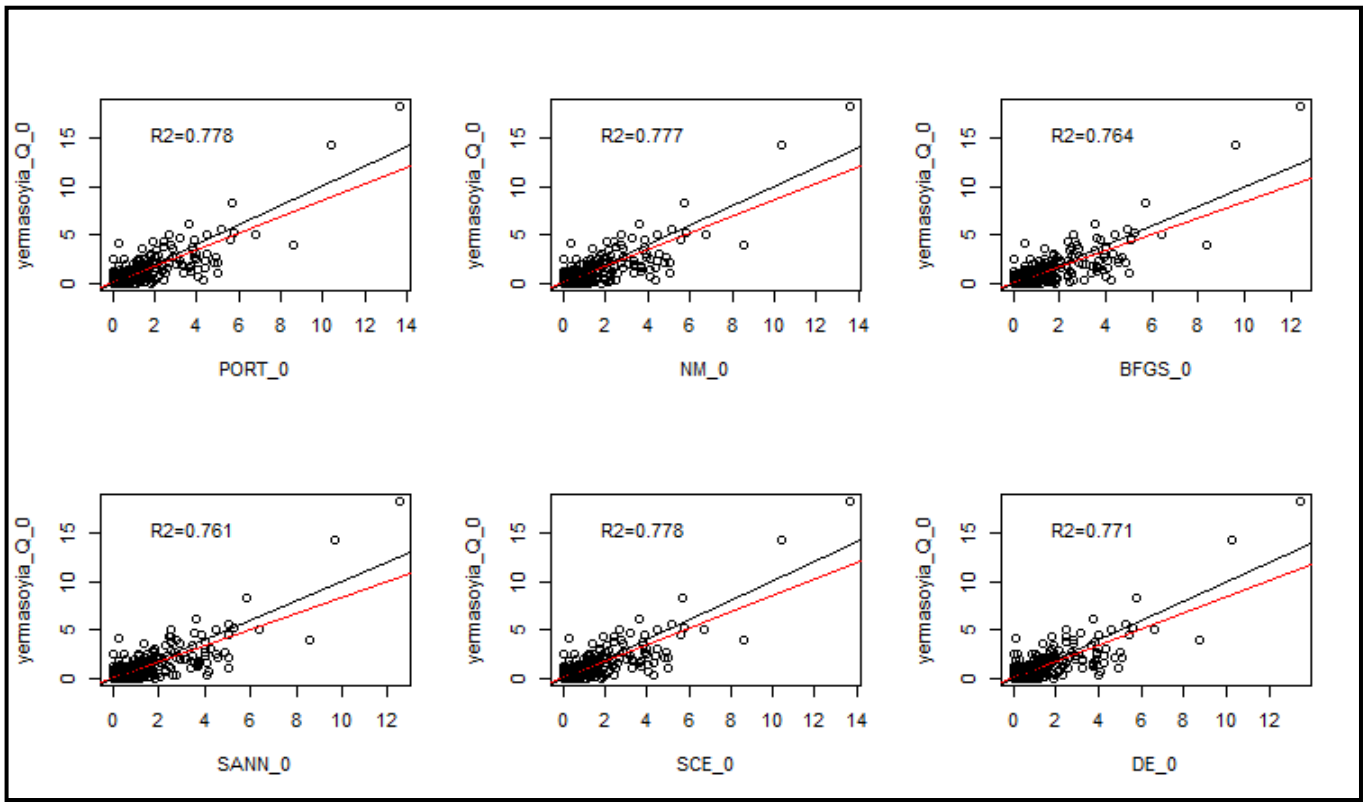


By NSE: dry period

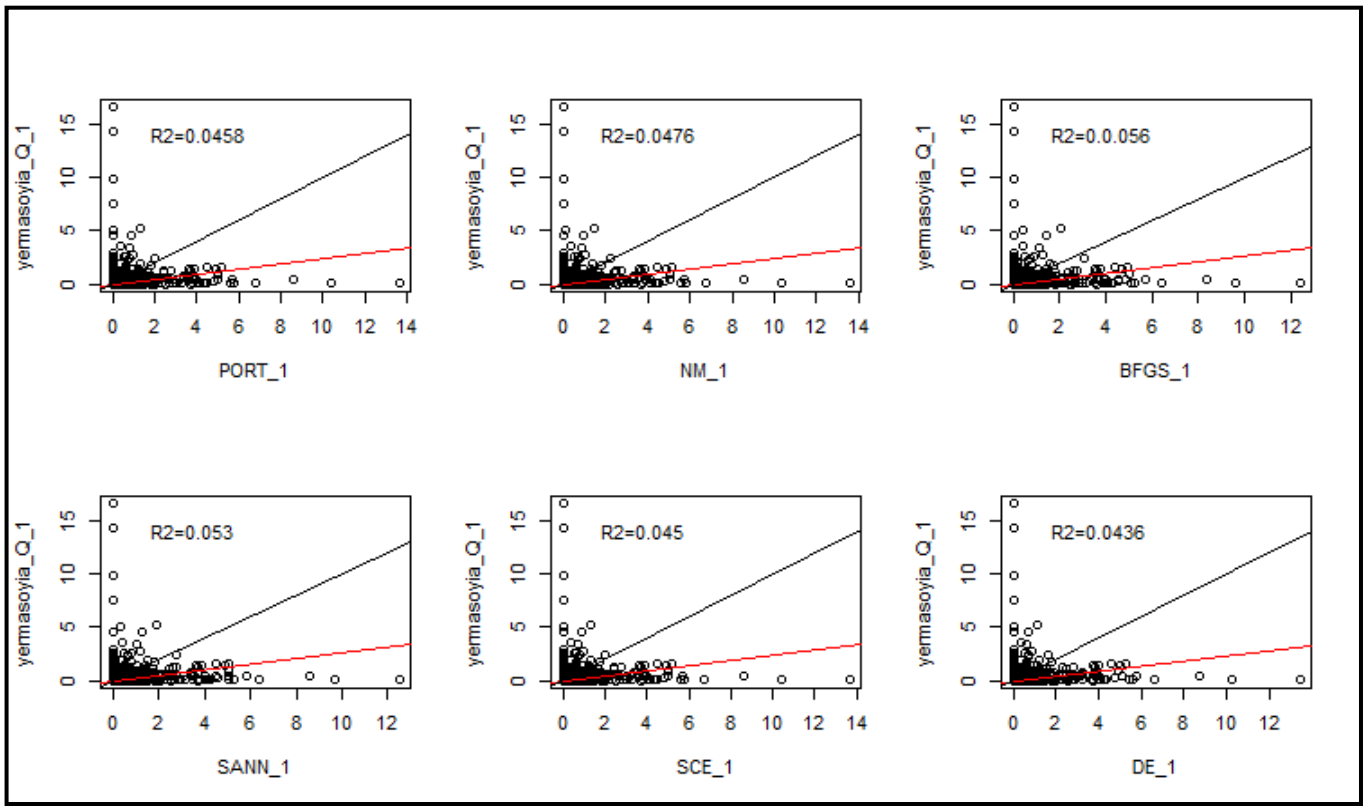


APPENDIX D

By Viney: wet period

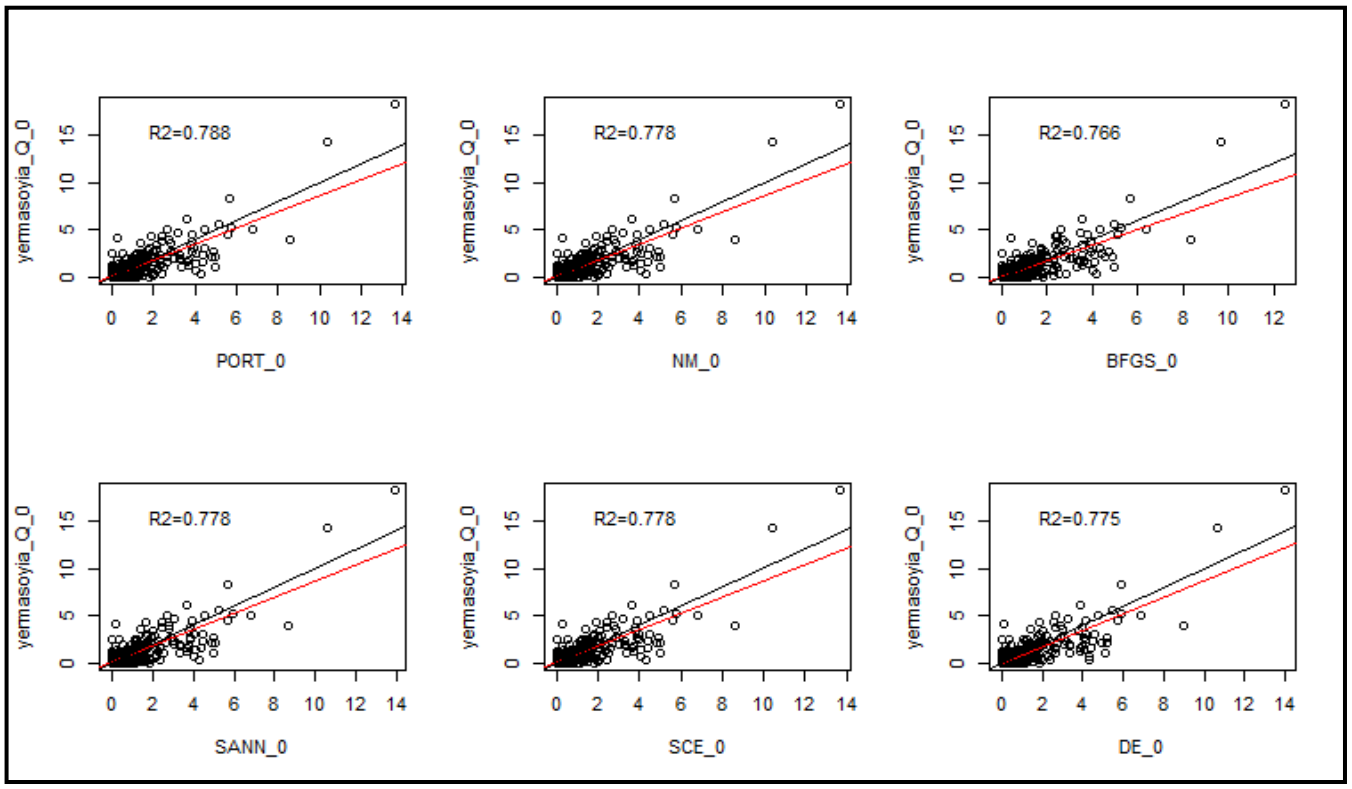


By Viney: dry period

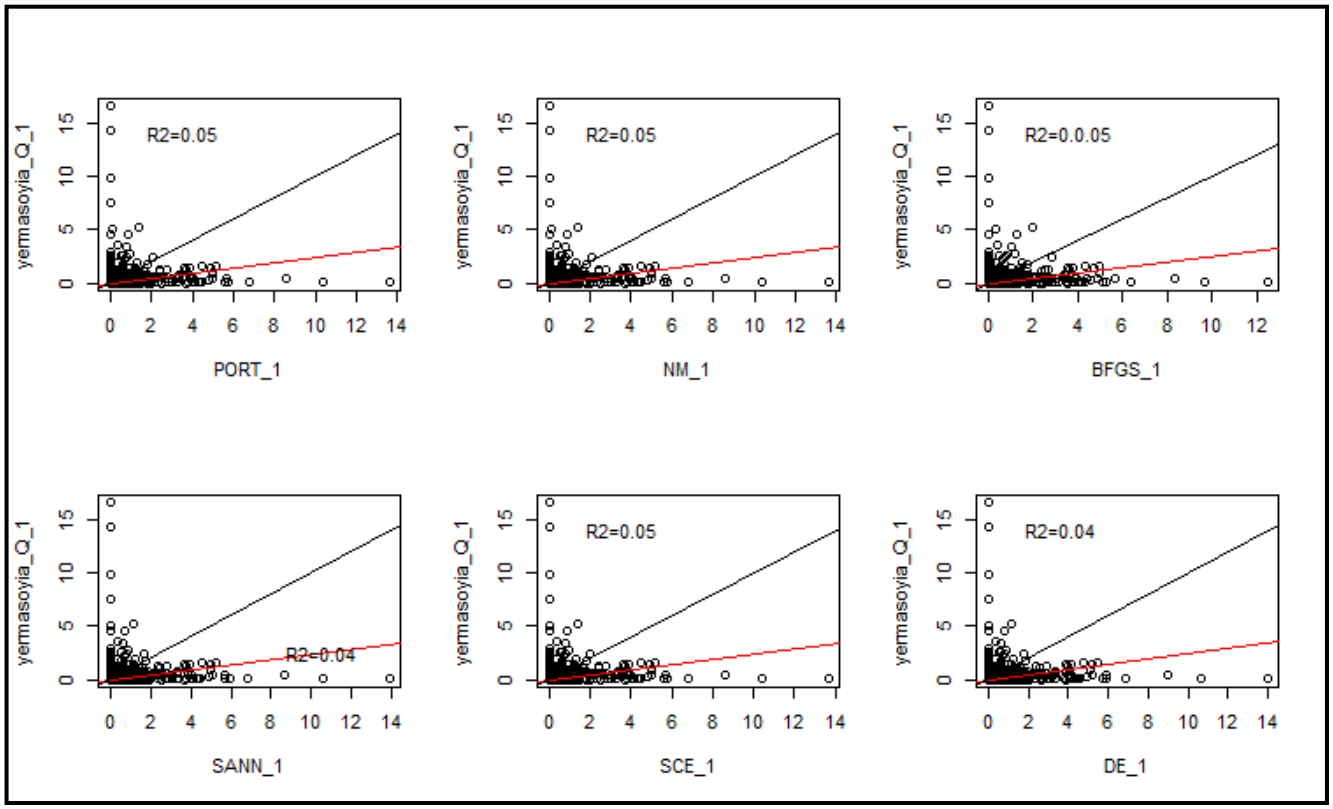


APPENDIX D

By BL: wet period

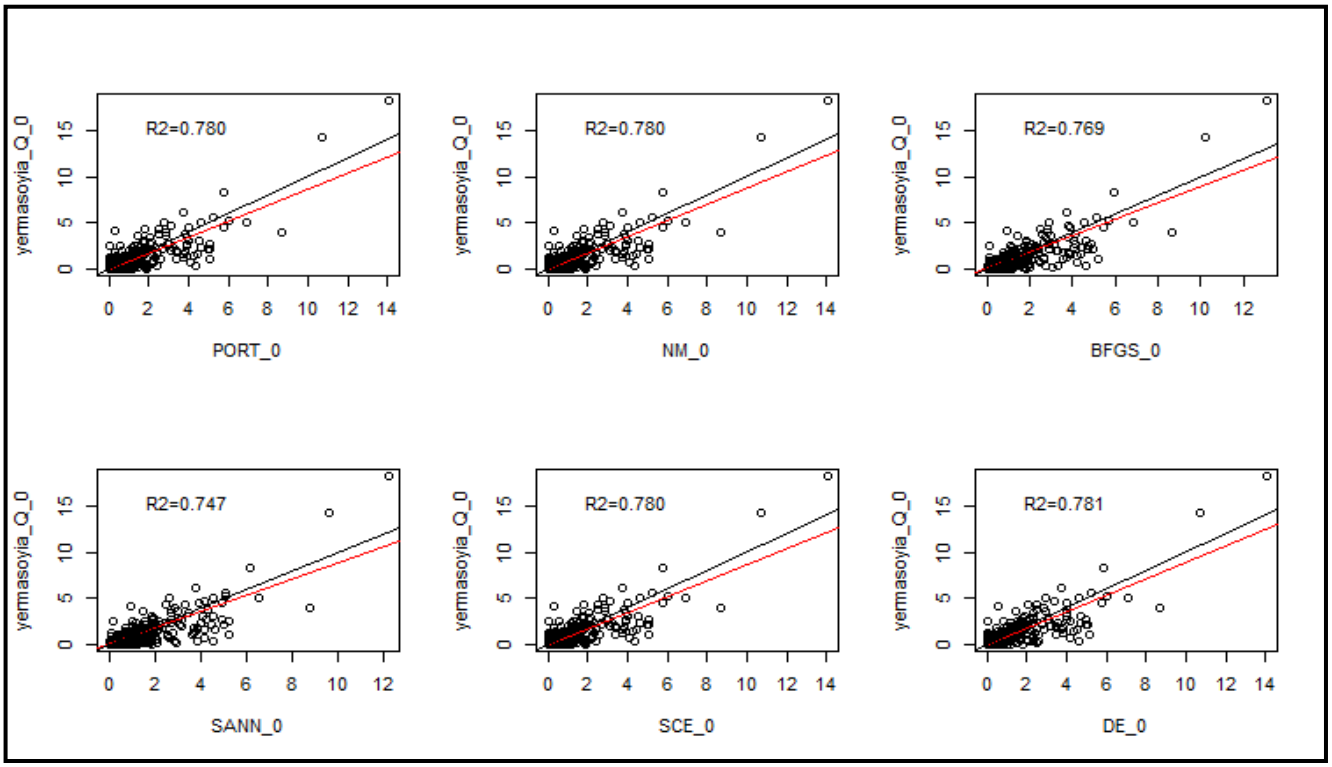


By BL: dry period

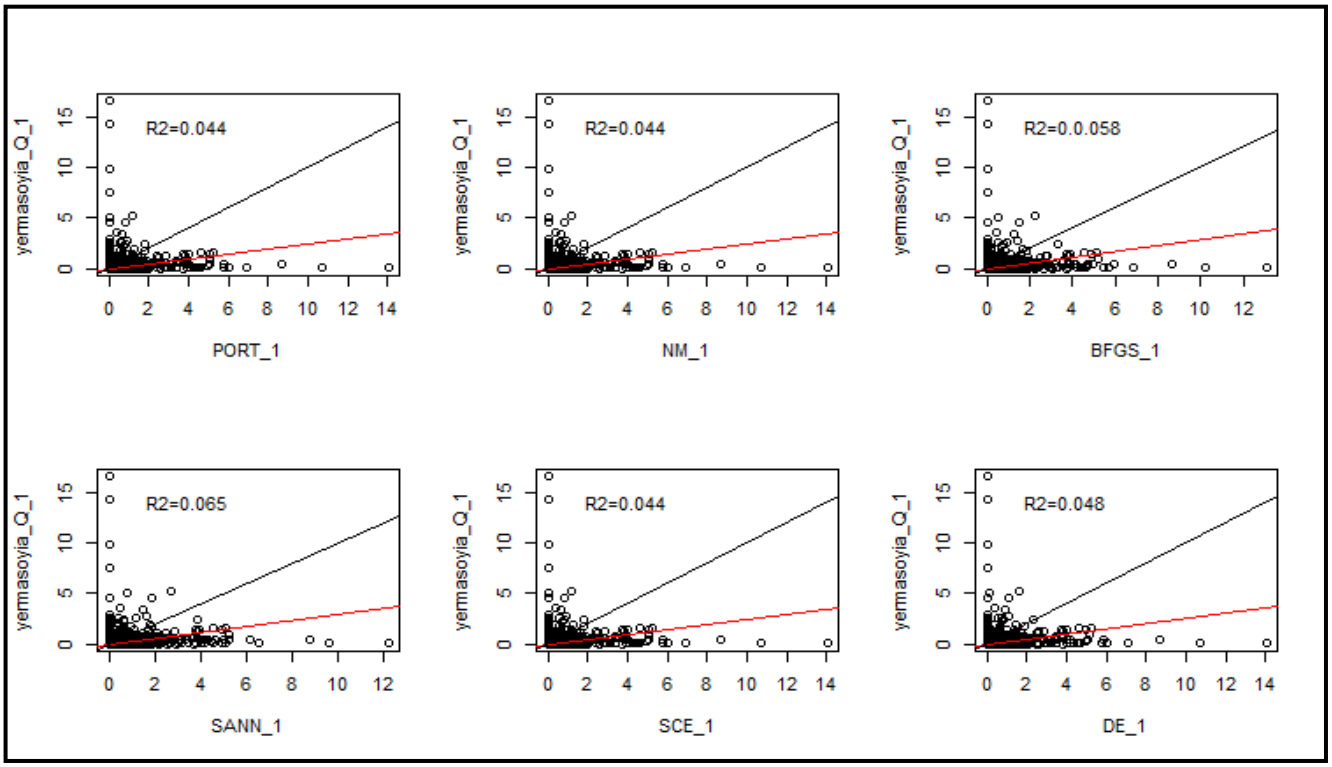


APPENDIX D

By NSE^3: wet period

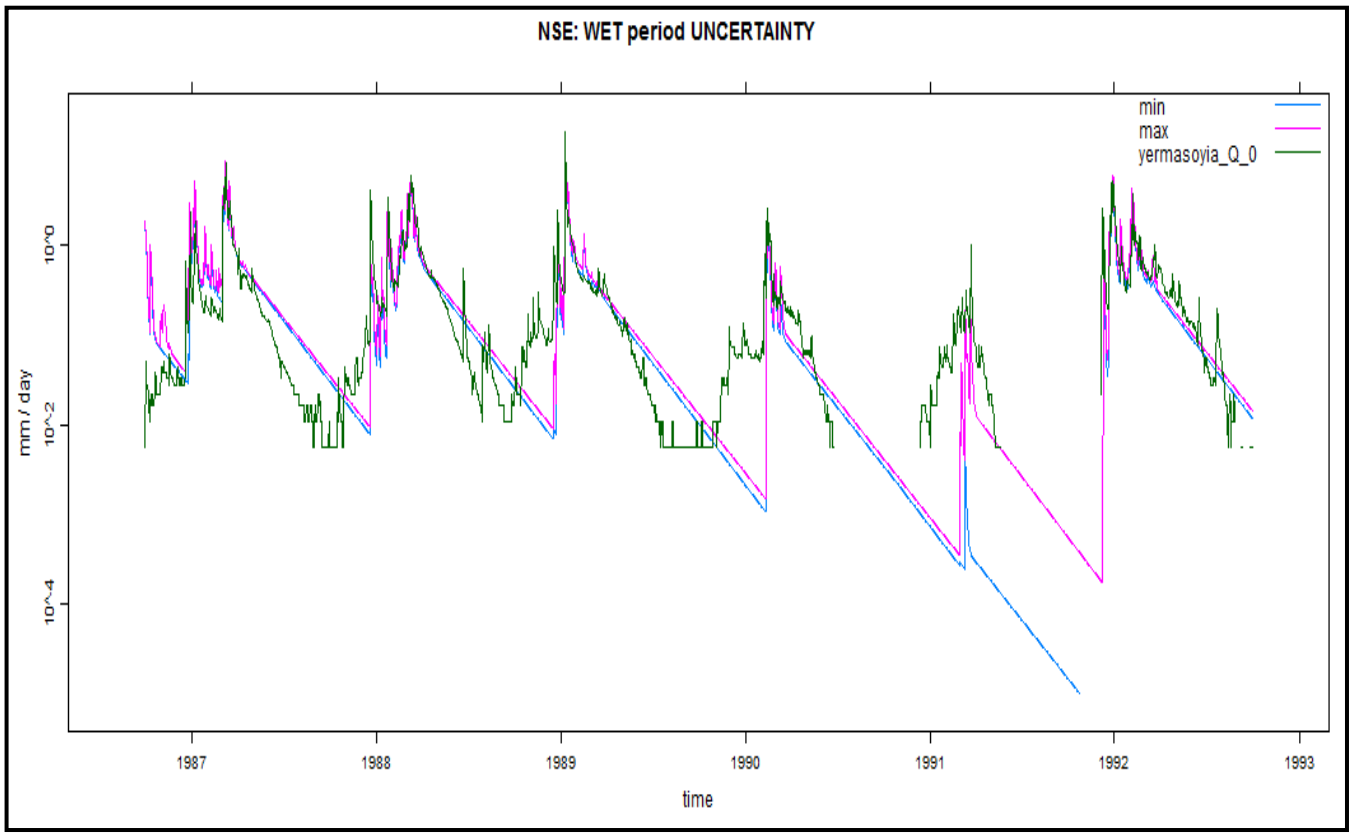


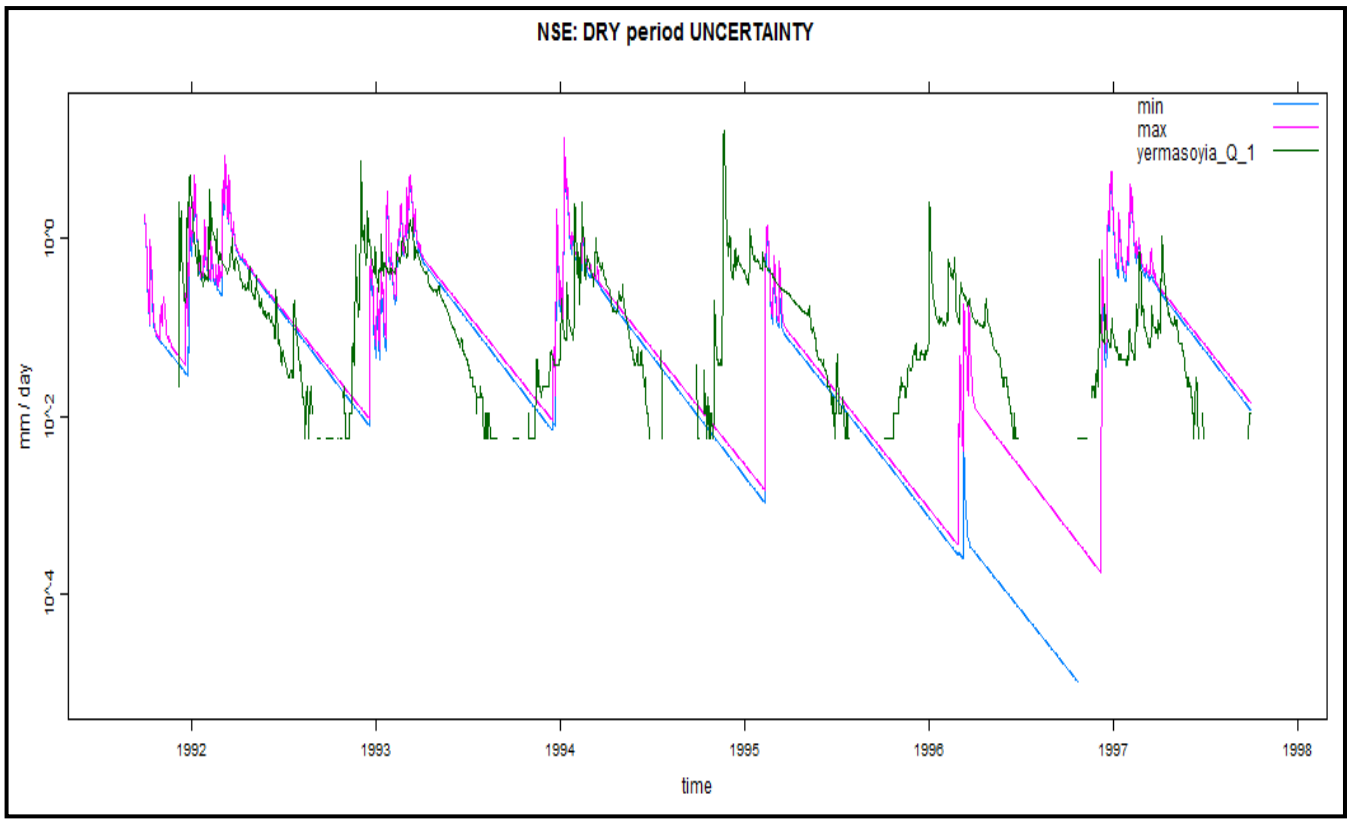
By NSE^3: dry period



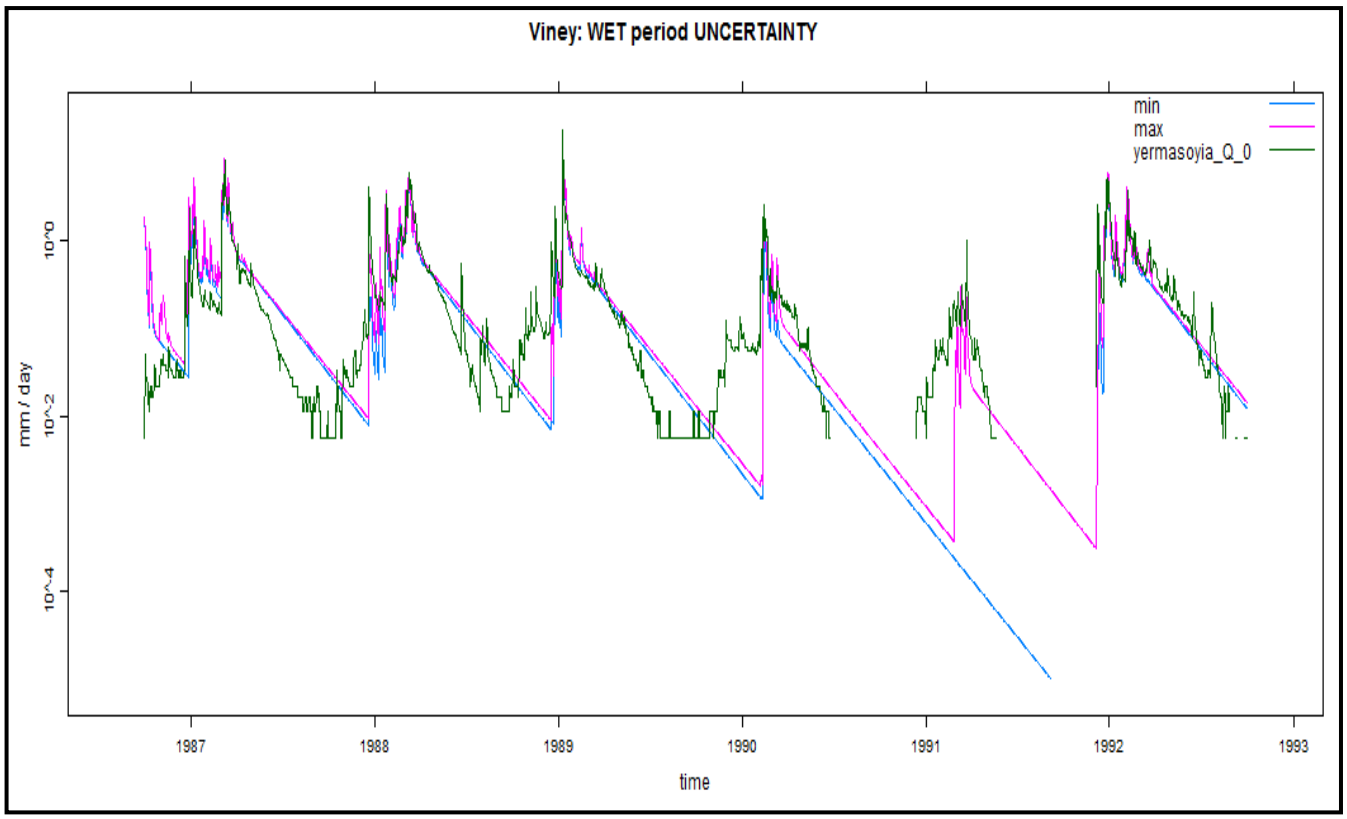
APPENDIX D

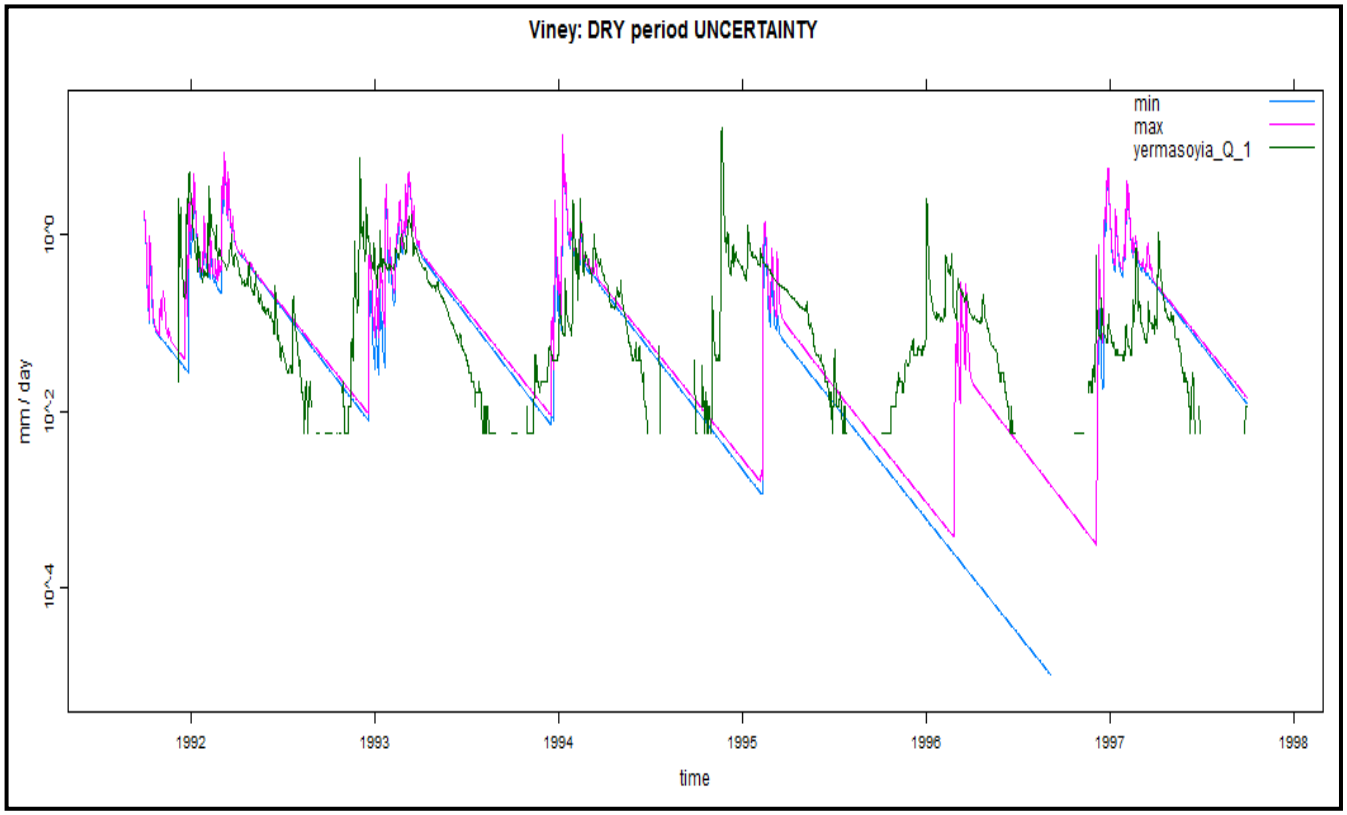
(D) Model structure uncertainty



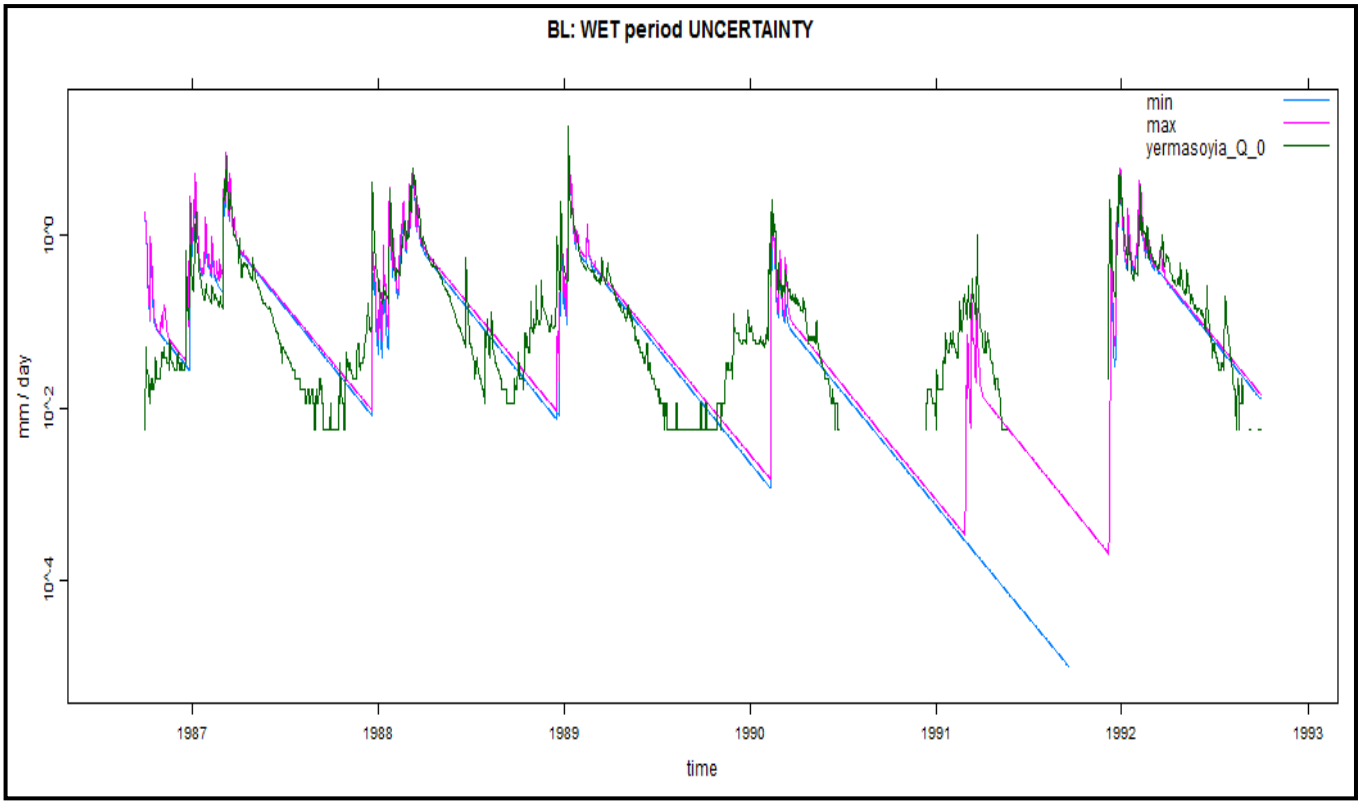


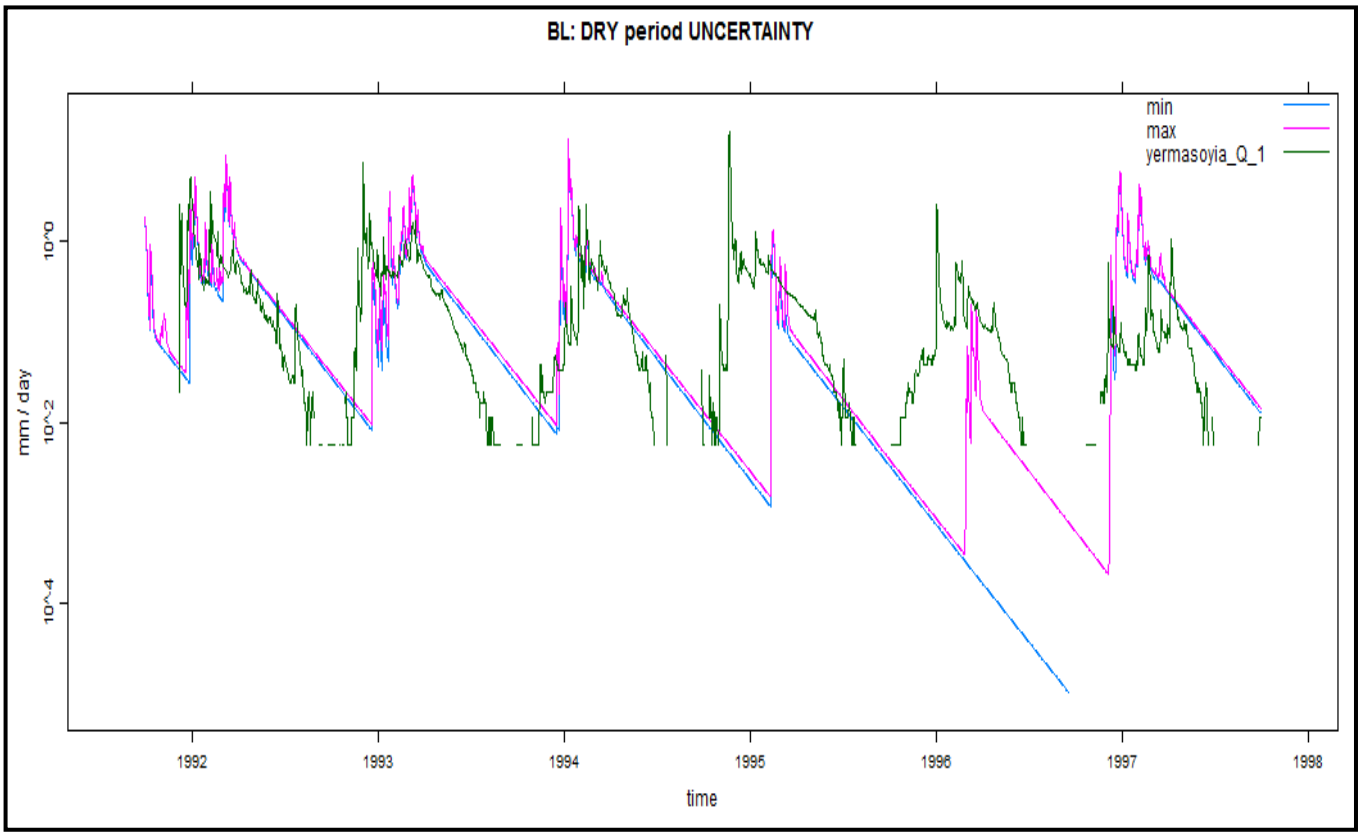
APPENDIX D



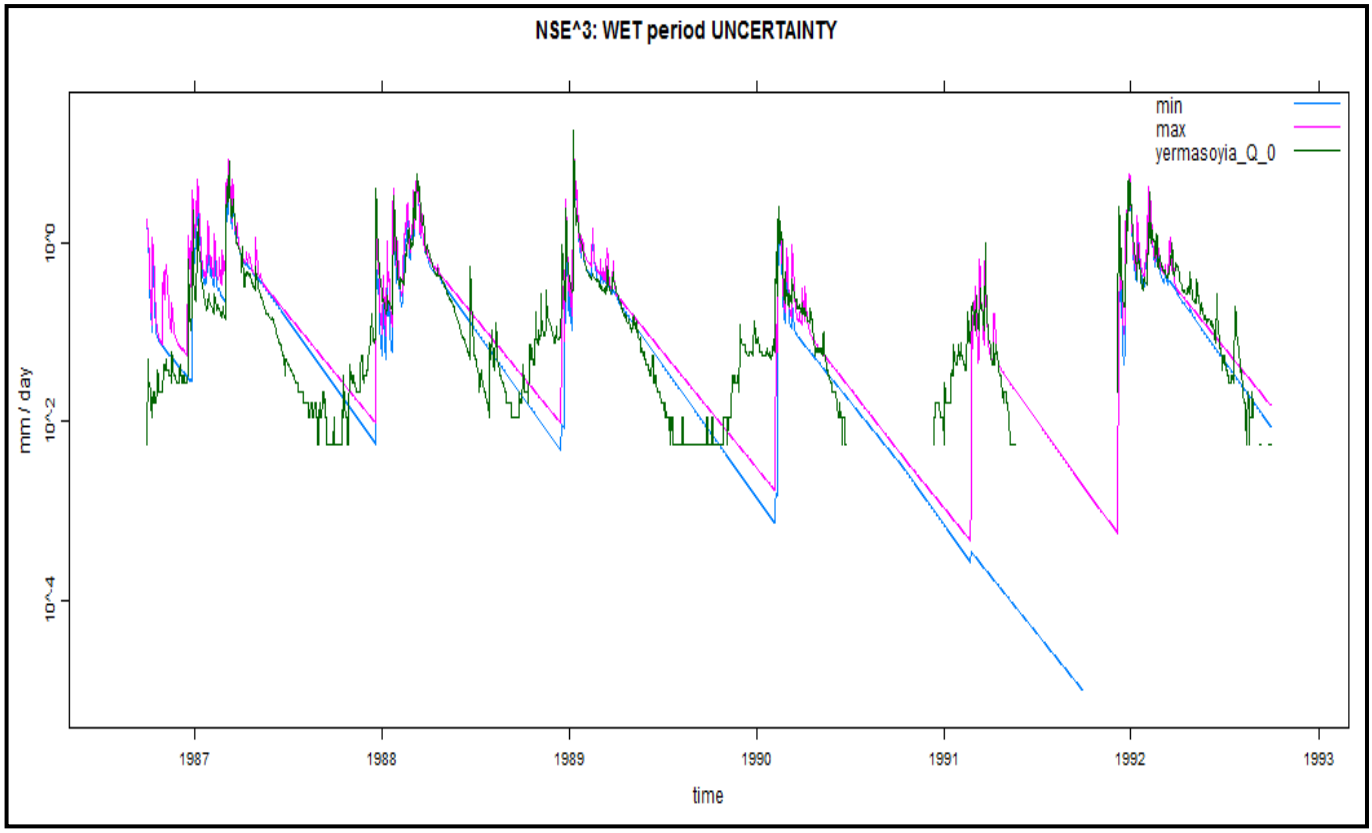


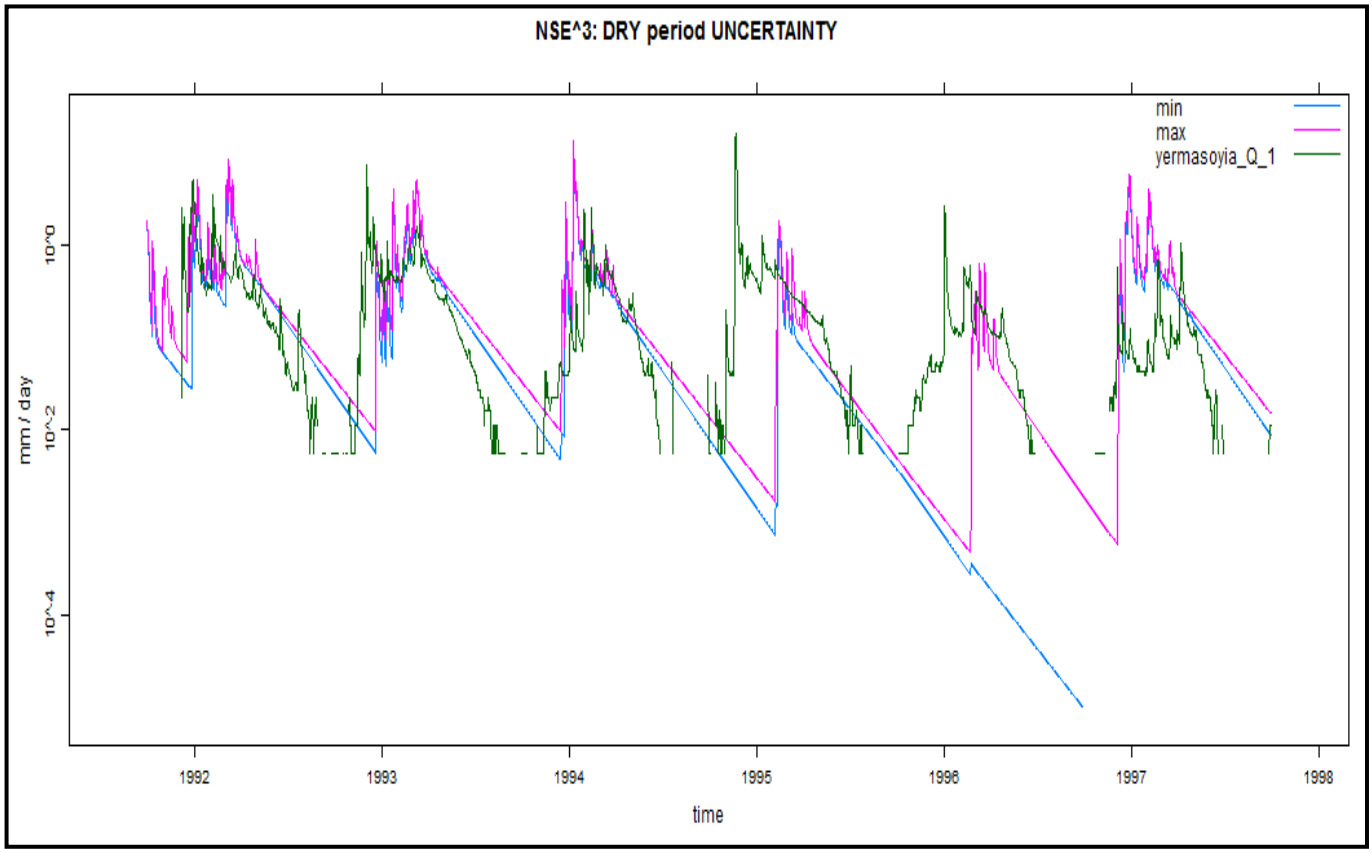
APPENDIX D





APPENDIX D



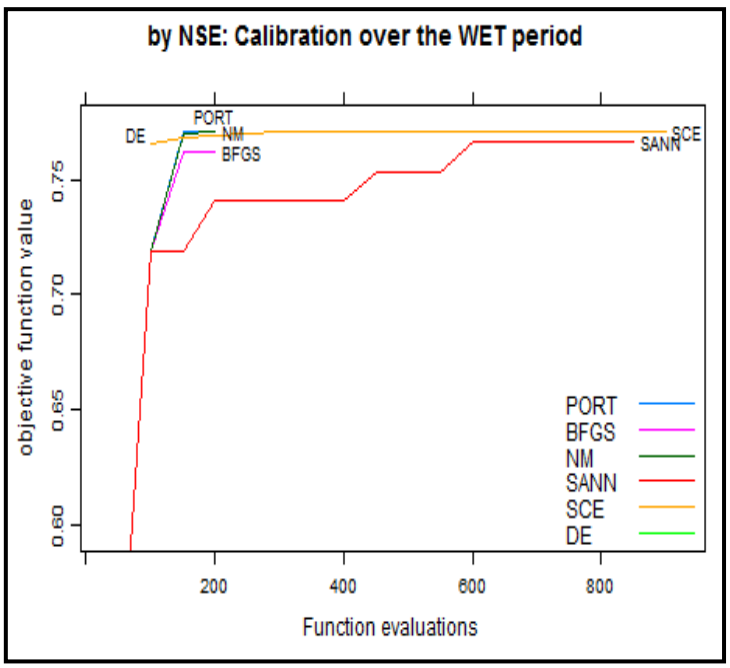


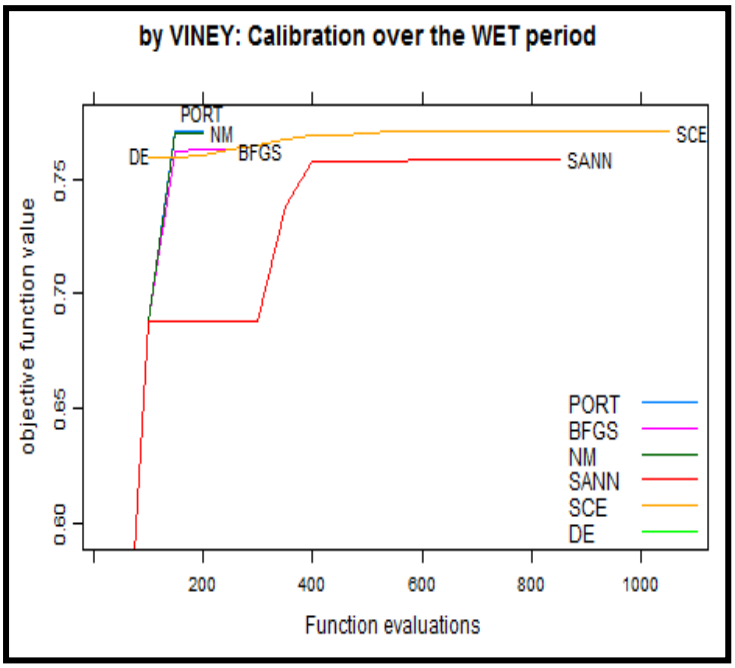
APPENDIX D

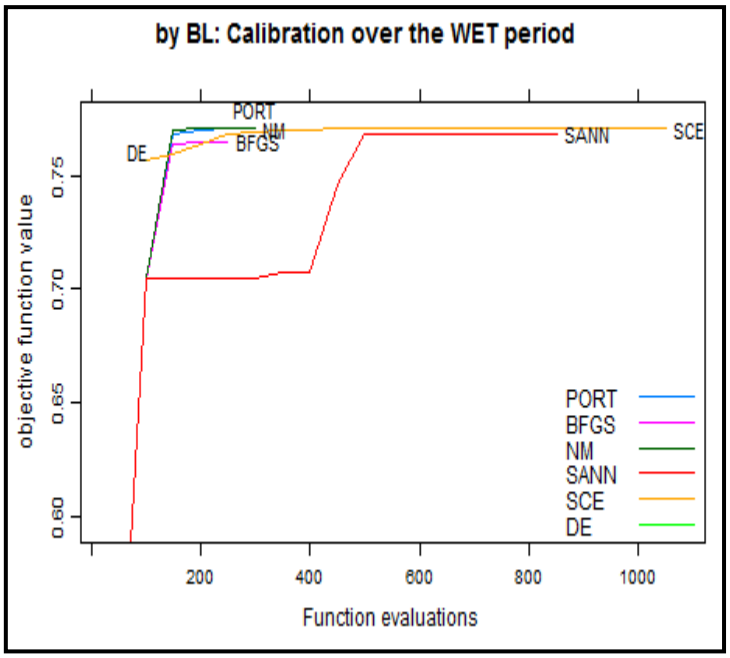
(E) Objective Function and Optimization Algorithm uncertainty

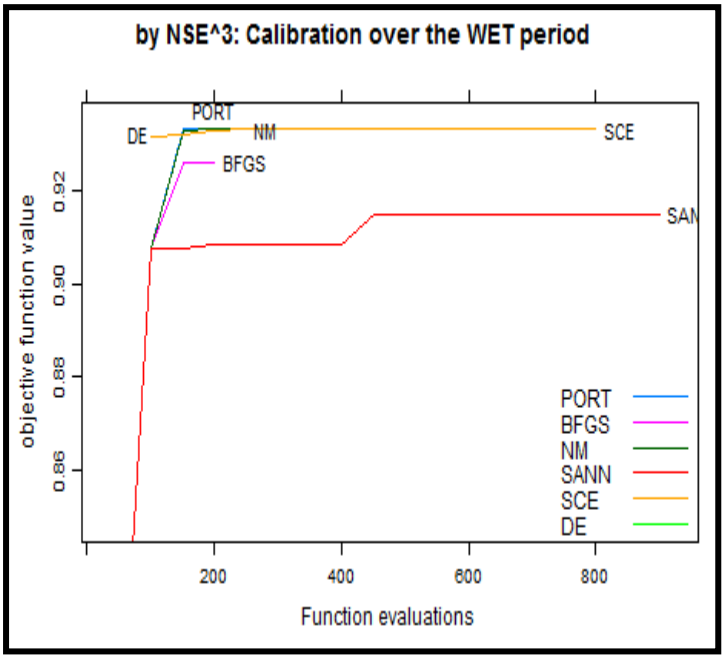
CALIBRATION: **WET** – VALIDATION: **DRY**

Optimization traces:









APPENDIX D

Validation FIT STATISTICS:

Objective Function = **NSE**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.155	0.595	0.640	0.582	0.625	0.845
BFGS	-0.147	0.653	0.593	0.543	0.578	0.792
NM	-0.130	0.618	0.646	0.610	0.633	0.850
SANN	-0.108	0.625	0.635	0.613	0.624	0.844
SCE	-0.155	0.595	0.640	0.582	0.625	0.845
DE	-0.176	0.591	0.620	0.538	0.602	0.826

Objective Function = **Viney**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.152	0.597	0.641	0.586	0.626	0.847
BFGS	-0.095	0.676	0.619	0.604	0.610	0.820
NM	-0.127	0.619	0.648	0.614	0.635	0.852
SANN	-0.131	0.632	0.586	0.549	0.572	0.795
SCE	-0.153	0.596	0.641	0.585	0.626	0.847
DE	-0.204	0.527	0.592	0.467	0.571	0.803

Objective Function = **BL**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.146	0.604	0.643	0.593	0.628	0.847
BFGS	-0.118	0.660	0.618	0.590	0.606	0.818
NM	-0.146	0.604	0.643	0.593	0.628	0.848
SANN	-0.172	0.559	0.635	0.558	0.618	0.843
SCE	-0.148	0.600	0.643	0.591	0.628	0.848
DE	-0.149	0.55	0.614	0.562	0.599	0.829

Objective Function = **NSE^3**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.151	0.577	0.649	0.596	0.634	0.856
BFGS	0.009	0.713	0.659	0.659	0.659	0.864
NM	-0.149	0.578	0.650	0.597	0.635	0.857
SANN	0.014	0.712	0.598	0.598	0.597	0.814
SCE	-0.151	0.577	0.649	0.595	0.634	0.856
DE	-0.078	0.637	0.671	0.661	0.663	0.875

APPENDIX D

Parameter Stability:

Objective Function = **NSE**

	f	e	d	shape	tau_s	tau_q	v_s	v_q	delay
PORT	0.895	1.500	178.366	0	60.999	2.121	0.484	0.516	0
BFGS	0.757	1.500	260.252	0	58.713	2.245	0.448	0.552	0
NM	0.861	1.500	188.704	0	61.055	2.150	0.479	0.521	0
SANN	0.850	1.458	196.701	0	60.753	2.178	0.470	0.530	0
SCE	0.895	1.500	178.349	0	60.999	2.121	0.484	0.516	0
DE	0.880	1.495	189.025	0	60.882	2.125	0.480	0.520	0

Objective Function = **Viney**

	f	e	d	shape	tau_s	tau_q	v_s	v_q	delay
PORT	0.892	1.500	178.886	0	61.012	2.123	0.483	0.517	0
BFGS	0.745	1.499	258.479	0	58.613	2.271	0.446	0.554	0
NM	0.861	1.500	188.292	0	61.068	2.151	0.479	0.521	0
SANN	0.799	1.419	238.781	0	59.217	2.232	0.449	0.551	0
SCE	0.894	1.500	178.281	0	61.009	2.122	0.484	0.516	0
DE	0.955	1.451	171.543	0	60.721	2.085	0.484	0.516	0

Objective Function = **BL**

	f	e	d	shape	tau_s	tau_q	v_s	v_q	delay
PORT	0.882	1.500	181.951	0	61.031	2.131	0.482	0.518	0
BFGS	0.769	1.499	243.365	0	59.346	2.238	0.453	0.547	0
NM	0.883	1.500	181.640	0	61.031	2.131	0.482	0.518	0
SANN	0.943	1.492	165.007	0	60.826	2.092	0.489	0.511	0
SCE	0.890	1.500	179.075	0	61.030	2.126	0.483	0.517	0
DE	0.955	1.422	167.231	0	60.852	2.114	0.481	0.519	0

Objective Function = **NSE^3**

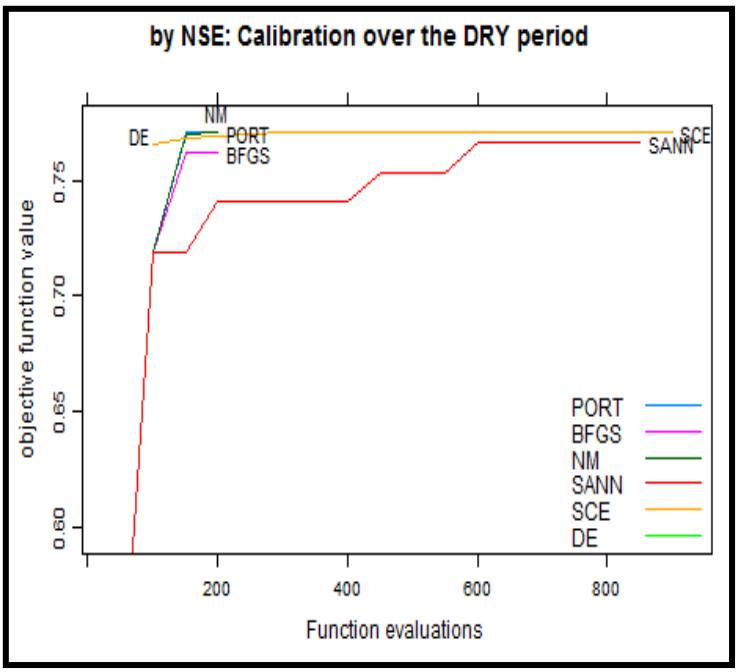
	f	e	d	shape	tau_s	tau_q	v_s	v_q	delay
PORT	0.930	1.500	164.990	0	60.979	2.103	0.489	0.511	0
BFGS	0.722	1.499	253.271	0	58.585	2.310	0.444	0.556	0
NM	0.930	1.500	164.925	0	60.987	2.104	0.489	0.511	0
SANN	0.667	1.432	341.338	0	53.590	2.416	0.405	0.595	0
SCE	0.931	1.500	164.758	0	60.974	2.102	0.489	0.511	0
DE	0.856	1.498	182.361	0	61.333	2.168	0.479	0.521	0

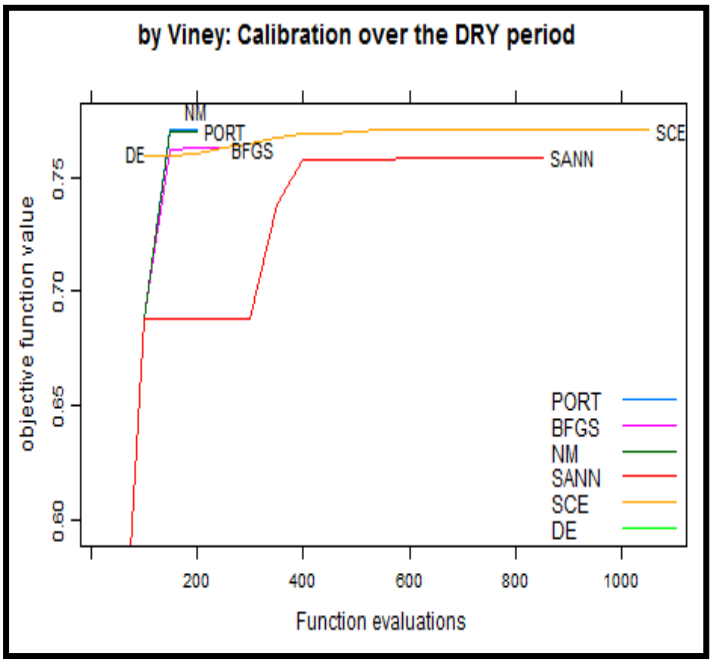
APPENDIX D

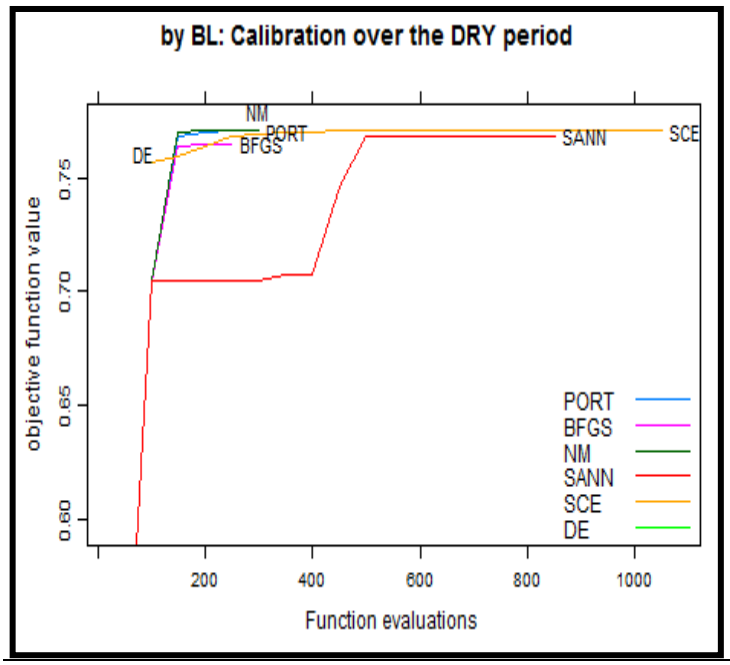
(F) Objective Function and Optimization Algorithm uncertainty

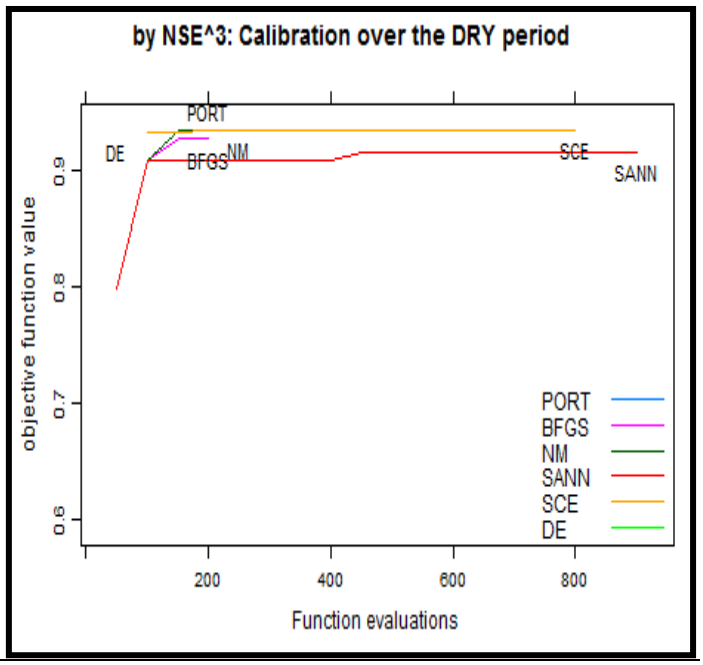
CALIBRATION: **DRY** – VALIDATION: **WET**

Optimization traces:









APPENDIX D

Validation FIT STATISTICS:

Objective Function = **NSE**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.006	0.776	0.771	0.771	0.770	0.933
BFGS	-0.029	0.791	0.762	0.761	0.759	0.921
NM	0.010	0.782	0.771	0.771	0.770	0.933
SANN	0.027	0.782	0.766	0.766	0.763	0.931
SCE	-0.006	0.776	0.771	0.771	0.770	0.933
DE	-0.032	0.773	0.770	0.769	0.767	0.931

Objective Function = **Viney**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	-0.003	0.776	0.771	0.771	0.771	0.933
BFGS	0.018	0.799	0.763	0.763	0.761	0.924
NM	0.012	0.782	0.771	0.770	0.769	0.933
SANN	-0.005	0.783	0.759	0.759	0.758	0.923
SCE	-0.004	0.776	0.771	0.771	0.771	0.933
DE	-0.047	0.757	0.764	0.762	0.760	0.930

Objective Function = **BL**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.000	0.778	0.771	0.771	0.771	0.933
BFGS	0.000	0.795	0.765	0.765	0.765	0.926
NM	0.000	0.778	0.771	0.771	0.771	0.933
SANN	-0.012	0.770	0.770	0.769	0.768	0.933
SCE	0.000	0.777	0.771	0.771	0.771	0.933
DE	0.010	0.763	0.762	0.762	0.761	0.930

Objective Function = **NSE^3**

	rel.bias	r.sq.sqrt	r.squared	viney	BL	nse^3
PORT	0.008	0.774	0.770	0.770	0.769	0.933
BFGS	0.111	0.805	0.758	0.740	0.747	0.926
NM	0.009	0.774	0.770	0.770	0.769	0.933
SANN	0.106	0.797	0.742	0.726	0.731	0.915
SCE	0.008	0.774	0.770	0.770	0.769	0.933
DE	0.061	0.789	0.768	0.764	0.762	0.933

APPENDIX D

Parameter Stability:

Objective Function = **NSE**

	f	e	d	shape	tau_s	tau_q	v_s	v_q	delay
PORT	0.895	1.500	178.366	0	60.999	2.121	0.484	0.516	0
BFGS	0.757	1.500	260.252	0	58.713	2.245	0.448	0.552	0
NM	0.861	1.500	188.704	0	61.055	2.150	0.479	0.521	0
SANN	0.850	1.458	196.701	0	60.753	2.178	0.470	0.530	0
SCE	0.895	1.500	178.349	0	60.999	2.121	0.484	0.516	0
DE	0.880	1.495	189.025	0	60.882	2.125	0.480	0.520	0

Objective Function = **Viney**

	f	e	d	shape	tau_s	tau_q	v_s	v_q	delay
PORT	0.882	1.500	181.951	0	61.031	2.131	0.482	0.518	0
BFGS	0.769	1.499	243.365	0	59.346	2.238	0.453	0.547	0
NM	0.883	1.500	181.640	0	61.031	2.131	0.482	0.518	0
SANN	0.943	1.492	165.007	0	60.826	2.092	0.489	0.511	0
SCE	0.890	1.500	179.075	0	61.030	2.126	0.483	0.517	0
DE	0.955	1.422	167.231	0	60.852	2.114	0.481	0.519	0

Objective Function = **BL**

	f	e	d	shape	tau_s	tau_q	v_s	v_q	delay
PORT	0.882	1.500	181.951	0	61.031	2.131	0.482	0.518	0
BFGS	0.769	1.499	243.365	0	59.346	2.238	0.453	0.547	0
NM	0.883	1.500	181.640	0	61.031	2.131	0.482	0.518	0
SANN	0.943	1.492	165.007	0	60.826	2.092	0.489	0.511	0
SCE	0.890	1.500	179.075	0	61.030	2.126	0.483	0.517	0
DE	0.955	1.422	167.231	0	60.852	2.114	0.481	0.519	0

Objective Function = **NSE^3**

	f	e	d	shape	tau_s	tau_q	v_s	v_q	delay
PORT	0.930	1.500	164.990	0	60.979	2.103	0.489	0.511	0
BFGS	0.722	1.499	253.271	0	58.585	2.310	0.444	0.556	0
NM	0.930	1.500	164.925	0	60.987	2.104	0.489	0.511	0
SANN	0.667	1.432	341.338	0	53.590	2.416	0.405	0.595	0
SCE	0.931	1.500	164.758	0	60.974	2.102	0.489	0.511	0
DE	0.856	1.498	182.361	0	61.333	2.168	0.479	0.521	0

OSUG

Observatoire des Sciences de l'Univers de Grenoble

Université Joseph Fourier GRENOBLE

GRENOBLE UFR PhITEM

Master Sciences de la Terre et de l'Environnement

Non-plagiarism certificate

I, the undersigned (First name, FAMILY NAME)

Eleftherios GKilimanakis

Author of the report entitled (Title)

"Identification of Uncertainty in Hydrological Modeling using several model structures, optimization algorithms and objective functions".

Declare that the above-cited report results from my personal work and that I have neither forged, falsified nor copied all or part of another persons work to present it as mine.

All sources of information used and all author citations have been included following standard usage.

I am aware of the fact that failing to cite a source or failing to cite it fully and properly constitutes plagiarism, and that plagiarism is considered a serious offence within the university that can be sanctioned severely by law.

At (place) Volos,

the (date) 07/01/2013

Student's signature