Characterizing Soil Moisture Memory Timescale and the Xinanjiang Model Spin-up time by Basin Scale Hydro-climatic Data

Mohammad Mahfuzur Rahman

Nagaoka University of Technology

August, 2015

Characterizing Soil Moisture Memory Timescale and the Xinanjiang Model Spin-up time by Basin Scale Hydro-climatic Data

A thesis submitted to
Nagaoka University of Technology, Japan

for the Degree of
Doctor of Engineering

Presented by

Mohammad Mahfuzur Rahman (ID 12701982)

accepted on the recommendation of

Professor Dr. Minjia Lu, Academic supervisor & Chief examiner
Professor Dr. Tokuzo Hosoyamada, examiner
Associate Professor Dr. Toshiro Kumakura, examiner
Associate Professor Dr. Kazuyoshi Takahashi, examiner
Associate Professor Dr. Hiroshi Ishidaira, examiner

Nagaoka University of Technology

August, 2015

External publication:

1. Rahman, M. M.; Lu, M. (2015): Characterizing soil moisture memory by soil moisture autocorrelation. Journal of Water Resource and Hydraulic Engineering, 3, 85-94.

2. Rahman, M. M.; Lu, M.; Kyi, K. H. (2015): Variability of soil moisture memory for wet and dry basins. Journal of Hydrology, 523, 107-118, doi: 10.1016/j.jhydrol.2015.01.033.

3. Rahman, M. M.; Lu, M. (2015): Model spin-up behaviour for wet and dry basins: a case study using Xinanjiang model. Water, 7, 4256-4273; doi:10.3390/w7084256.

4. Rahman, M. M.; Lu, M.; Kyi, K. H. (2015): Seasonality of model spin-up time for wet and dry basins. Manuscript under preparation.

Oral presentation:

1. Rahman, M. M.; Lu, M. (2015): Characterizing soil moisture memory by soil moisture autocorrelation. Manuscript accepted for oral presentation at “2015 International Conference on Water Resource and Environment (WRE2015)”, July 25-28, 2015, Beijing, China.

Dedication

To

My father

Who departed untimely, leaving me a lame
A very costly exchange to acquire a fame.

Abstract

The persistence characteristics of soil moisture is known as soil moisture memory (SMM). Knowledge of SMM is important for both land surface and hydrological modelling exercises. SMM information can improve hydro-climatic prediction efficiency as well as provides useful insight about the speed of model spin-up process. Despite these advantages, SMM studies are restricted in certain regions due to the scarcity of observed soil moisture data. To overcome this limitation, this study explains the variability of SMM with respect to the dryness of a river basin and shows a way to predict basin scale SMM timescale using annual observed precipitation and potential evapotranspiration information only. Later, the linkage between the SMM and the model spin-up time has been investigated using the Xinanjiang (XAJ) model as a case study.

The study presents the basin average SMM timescale that indicates the duration of significant autocorrelations at 95% confidence intervals. The soil moisture autocorrelations were calculated using observed precipitation, potential evapotranspiration, streamflow and soil moisture data sets (soil moisture data was simulated using the XAJ model), for 26 river basins across the USA. The SMM timescale is highly influenced by precipitation variability and exhibits strong seasonality. Dry basins tend to show the highest memory during the winter months (December to February) and lowest in late spring (May). In contrast, wet basins have the lowest memory during winter and early spring (December to April) and highest in the late summer and early autumn (July to September). The analysis suggests that SMM timescale holds an exponential relationship with the basin aridity index.

The model spin-up behavior was evaluated by two separate approaches. Firstly, it records the model spin-up time by following a recursive simulation with single year climatology with three climatological input data sets and four initial conditions. The spin-up time was defined based on the percent cutoff-based time technique. Secondly, it performs a series of simulations using multi-year input data sets with varying simulation start time and two initial conditions ("saturated" and "dry"). The spin-up time was defined as the time required for the Mahalanobis Distance between the model soil moisture states of "saturated" and "dry" simulations becomes zero (0).

Both the analysis reveals that the model spin-up has a clear association with SMM timescale. Moreover, it reveals that model equilibrium is not only a function of initial conditions, but also is affected by input data sets. Model requires less time to be equilibrated under wetter initial conditions (lowest under saturated initial condition). Moreover, model spin-up time shows distinct variations with the dryness of the climate forcing. Analysis

suggests that wet basins require less time for model equilibrium in comparison to that of dry basins. In wet basins, model spin-up process is slower during the spring months and faster during winter months. In contrast, in dry basins, spin-up process is faster during the late spring and early autumn and slower during the winter. The spin-up time also displayed an exponential relationship with the basin aridity index. Both SMM timescale and the model spin-up time can be predicted from widely available observed data sets.

The highlighting points of this thesis:

- Computed and validated soil moisture autocorrelations using model simulated soil moisture data (validated against the observed seasonal cycles). This overcomes the soil moisture data limitation for SMM studies.
- This study presents SMM in the form of a timescale dealing with a single number that is more easier to understand. This improves the understanding and feeling of SMM variability.
- Analyses more than 25 river basins representing diverse climatic conditions across USA. This enables to show basin scale variations of SMM which is absent in the available literature.
- Explains seasonality and variability of SMM with respect to the dryness of the river basin. This establishes the linkage between the SMM and basin hydro-climatic properties.
- Analyses model spin-up time following recursive simulations using three different climatological forcing. This shows a way to improve the representativity of actual observations even with a recursive simulation-based spin-up studies.
- Discusses the variability of the model spin-up time with respect to the model initial conditions and basin dryness. This identifies the primary factors that affect the model spin-up process and provides important insights for model initialisation under data scarce situation.
- Explains the seasonality of model spin-up time based on simulations using multi-year climatologies. This provides a path to overcome the limitations with single year climatology-based recursive simulations for spin-up studies.
- Verifies the association between the SMM and the model spin-up time. This offers new knowledge to the modelling communities that are missing in the available literature.
- Finally, it shows the predictability of both SMM and model spin-up time from widely available precipitation and potential evapotranspiration data sets. This affords some knowledge of SMM and the model spin-up time, particularly for those areas where no knowledge is available.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Problem statement

The persistence characteristics of soil moisture is known as the soil moisture memory (SMM). The extent of this SMM has potential application in different research fields, primarily in weather and climate or hydrological predictions. Knowledge of SMM has proven to be an additional tool for improving traditional climate/hydrological forecast efficiency. Advantages of the SMM inclusive approach over traditional forecasts have been reported in several studies. Knowledge of SMM could improve soil moisture initialization for weather and climate prediction processes, and thus enhance prediction efficiency. Such knowledge might also improve the predictability of soil moisture and associated climate. Moreover, soil moisture, through its influence on land energy balance (partitioning sensible and latent heat flux), provides additional feedback on temperature and precipitation. Furthermore, a persistent soil moisture anomaly may prolong the effects of drought and influence the magnitude, occurrence, and receding of floods and streamflow dynamics. Finally, SMM is believed to be capable of propagating streamflow. Currently, streamflow forecasting efficiency mostly depends on the prediction accuracy of atmospheric forcing or snow accumulation. However, soil moisture conditions evidently affect streamflow forecast accuracy. A dry soil below the snow pack receives and stores infiltrated water from snow melt and later it evaporates rather than becoming runoff into streams. Conversely, wet soil beneath the melting snow pack can stimulate streamflow. Therefore, SMM potentially affects streamflow forecast skills.

In spite of its huge importance, SMM studies are restricted to certain areas of the world due to the scarcity of long-term soil moisture data. Presently, such kinds of data are available only for limited areas (i.e. Russia, China, Mongolia and Illinois, USA). Additionally, these observations were taken mostly on a weekly to half-monthly basis, depending on location and season. Therefore, to date SMM studies using observed soil moisture data are mostly limited to half-monthly or monthly timescales. Latest SMM studies are mainly done on a global scale with regional averages. However, SMM may vary from basin to basin, even those located within very close latitudes, due to their variations in hydro-climatological properties; hence the basin scale variations are unknown. Therefore, a clear relationship between the basin scale SMM and its hydro-climatic (i.e. aridity index) or soil property information could be another important alternative way to gain a rough estimation of SMM.

Latest techniques estimate SMM in the form of the strength of autocorrelation of soil moisture with different time lags. The variations of the strength of SMM through the changes of autocorrelation values over time lags. The higher the coefficients the higher the memory is. However, it is not always easy and straightforward to understand the change. A change of autocorrelation value from 0.9 to 0.2 only indicates a decrease in memory and does not specify the decrease of the timescale clearly, and vice versa. It also does not clarify whether this relationship is statistically significant. Even if the significance of the relationship is known, how long it will stay significant remains unknown. Therefore, it would be useful once the corresponding timescales of these autocorrelation values are known. Changes in SMM timescale from 60 days to 45 days would be easy and straightforward.

The strength of SMM also affect the spin-up process of hydrological or land surface models. When a model is calibrated with a different initial state compared to the long-term climatology, the model undertakes a period of spin-up during which its internal stores (i.e. soil moisture) adjust from the initial conditions to an equilibrium state. The duration of this adjustment process is mainly a function of soil moisture persistence (SMM). A low SMM indicates that the soil moisture anomalies are short-lived, dissipate quickly, enabling the model to recover relatively quickly from an undesirable initial state. On the other hand a high SMM that indicates the slowness of anomaly dissipation and would delay the process of model equilibrium. The model outputs during this spin-up process is often unrealistic or misleading, and thus affect the modelling accuracies. To avoid this, modellers usually exclude initial model outputs during the analysis. Due to the lack of information, the amount of output exclusion is mainly guided by a guess, experience or purpose. This practice bears certain risk of excluding essential information or including erroneous model outputs. Therefore, there is a need to understand the model spin-up process, its controls and the way to reduce the uncertainty associated with guessing spin-up time simply based on feeling or experience.

1.2 Research questions

Based on the problem statement, several research questions have been developed.

1. How the autocorrelation based SMM can be expressed as SMM timescale?
2. How the SMM timescale changes temporally and spatially?
3. What is the relationship between the basin scale SMM timescale and hydro-climatic properties?
4. How the model spin-up time varies with respect to initial conditions and input data climatology?
5. How the model spin-up time varies temporally and spatially?
6. What is the relationship between the SMM timescale and the model spin-up time?
7. How can we predict the model spin-up time using widely available observed dataset?

1.3 Objectives

The overall objective of this study was to characterize soil moisture memory timescale and the Xinanjiang model spin-up time by basin scale hydro-climatic data. This overall objective was segmented into several specific objectives to answer all the research questions mentioned in section 1.2.

Specific objectives:

1. To identify a way to convert the soil moisture autocorrelations into an easily understandable SMM timescale expressed as time duration. This specific objective answer the first research question.
2. To analyse the behaviour, seasonality and variability of SMM with respect to the dryness of the basin. This objective satisfies research question # 2.
3. To detect any relationship between the basin scale SMM timescale and its hydro-climatic properties that holds the potentials to be used as an alternative source of SMM prediction. This objective solves third research question.
4. To analyse the behaviour and seasonality of model spin-up time with respect to initial conditions and input data climatology. This objective resolves fourth and fifth research questions.
5. To understand the relationship between SMM and model spin-up time. This objective fulfills research question # 6.
6. To find a way to predict model spin-up time based on observed hydro-climatic dataset. This objective answers the last research question.

1.4 Structure of the thesis

Chapter 1: Summarizes the background, research questions and the objectives of this study.

Chapter 2: Provides literature review about the soil moisture memory.

Chapter 3: Discusses the conversion of soil moisture autocorrelations into SMM timescale linked with specific objective #1.

Chapter 4: Analyzes the seasonality and variability of SMM in wet and dry river basins. This chapter is connected with the second and third specific objectives.

Chapter 5: Examines the model spin-up time behaviour under different initial conditions and input data climatology using the Xinanjiang model. This chapter is associated with fifth, sixth and part of fourth specific objectives.

Chapter 6: Investigates how the model spin-up varies with simulations start time of year. This chapter is related with the fourth specific objective.

Chapter 2

Soil moisture memory : literature review

2.1 Soil moisture memory and its importance

Soil moisture shows special persistence characteristics (Delworth and Manabe, 1988; Entin et al., 2000; Koster and Suarez, 2001; Seneviratne and Koster, 2012). Soils can act as a temporary reservoir and has the ability to store deficits or surpluses of moisture. This anomalous conditions (deviation from the mean state) are usually created by the variability of atmospheric forcing (i.e. precipitation or radiation). For example, a period of heavy rainfall would create a positive anomaly in the soil moisture state. This soil moisture anomaly then dissipates through evapotranspiration or runoff. Similarly, a negative anomaly could also be created by a prolong dry spell and thereafter dissipated through rainfall, snow melt, or irrigation. Literature suggests that the complete dissipation process can take hours or months (Entin et al., 2000; Mahanama and Koster, 2003). Therefore, the soil can remember an anomalous condition long after it has occurred. This phenomenon of memorizing past anomalies is termed soil moisture persistence or soil moisture memory (hereinafter referred as SMM). However, this dissipation process and timescales differ from place to place due to variations in soil properties, scale of interest, and climate forcing. The extent of this SMM has potential application in different research fields, primarily in weather and climate or hydrological predictions. Knowledge of SMM has proven to be an additional tool for improving traditional climate/hydrological forecast efficiency (Koster et al., 2010a, 2004).

Advantages of the SMM inclusive approach over traditional forecasts have been reported in several studies. Koster et al. (2010a) and van den Hurk et al. (2012) documented the usefulness of SMM understanding to improve soil moisture initialization for weather and climate prediction processes. SMM knowledge can enhance prediction efficiency, mainly in reference to longer timescales and under high initial soil moisture anomaly conditions. Such knowledge might also improve the predictability of soil moisture and associated climate (Schlosser and Milly, 2002). Moreover, soil moisture, through its influence on land energy balance (partitioning sensible and latent heat flux), provides additional feedback on temperature (Seneviratne et al., 2006b; Mueller and Seneviratne, 2012) and precipitation (Koster et al., 2004; Dirmeyer et al., 2009). Furthermore, a persistent soil moisture anomaly may prolong the effects of drought (Hong and Kalnay, 2000; Nicholson, 2000) and influence the magnitude, occurrence, and receding of floods (Bonan and Stillwell-Soller, 1998) and

streamflow dynamics.

Low-frequency streamflow dynamics are believed to be controlled by the catchment wetness, SMM (Gudmundsson et al., 2011). Likewise, SMM is believed to be capable of propagating streamflow (Orth and Seneviratne, 2013). Currently, streamflow forecasting efficiency mostly depends on the prediction accuracy of atmospheric forcing or snow accumulation. However, soil moisture conditions evidently affect streamflow forecast accuracy. A dry soil below the snow pack receives and stores infiltrated water from snow melt and later it evaporates rather than becoming runoff into streams. Conversely, wet soil beneath the melting snow pack can stimulate streamflow (Koster et al., 2010b). Therefore, SMM potentially affects streamflow forecast skills (Koster et al., 2010b; Mahanama et al., 2012).

2.2 Review of SMM study approaches

2.2.1 Non-autocorrelation-based approach

Delworth and Manabe (1988) described this unique persistence characteristic of soil moisture through a first-order Markov process run by a random precipitation forcing, as shown in Eq. (2.1):

$$\frac{dw\left(t\right)}{dt} = -\frac{PE}{C_s} w\left(t\right) + P - Q \tag{2.1}$$

where random processes precipitation (including snow melt) P and runoff Q (runoff in this context and in the following is the sum of base flow and streamflow) force soil moisture, $w(t)$. PE is the potential evapotranspiration and C_s is the bucket model's (Manabe (1969)) water holding capacity. In Eq. (2.1), $\frac{PE}{C_s} w\left(t\right)$ term controls the soil moisture variability and generates soil moisture persistence. It computes the length of this persistence timescale by assuming an exponential decay function of autocorrelation with the lag time, usually known as the e-folding time (time required to reduce the calculated autocorrelation coefficient, ρ values to its 1/e value), as shown in Eq. (2.2):

$$\rho\left(t_{lag}\right) = e^{-\frac{PE}{C_s} t_{lag}} \tag{2.2}$$

where ρ is the autocorrelation at lag time t_{lag}.

This timescale ranges from no memory (zero) to infinite memory (infinity). This theory was validated and several observational SMM studies confirmed its effectiveness (Wu and Dickinson, 2004; Schlosser and Milly, 2002; Entin et al., 2000; Vinnikov et al., 1996; Vinnikov and Yeserkepova, 1991).

2.2.2 Autocorrelation-based approach

A. *Koster and Suarez*

Koster and Suarez (2001) proposed another framework that calculates the power of SMM in the form of autocorrelation of soil moisture. This approach estimates SMM for a specific time lag by calculating correlations between the soil moisture anomalies of a particular day (n) and soil moisture anomalies of day with a specific lag ($n+lag$)using multi-year data. As an example, Fig. 2.1 explains the calculation of SMM for 1 January with a specific lag.

According to this approach, calculation of autocorrelation using soil moisture or for its anomalies provides the same outcome, as using the anomalies would only include a constant offset that would not affect the resultant correlation coefficient (Orth, 2013). Thus, the autocorrelation equation for the soil moisture at day n and $n + lag$ can be written as Eq. (2.3).

$$\rho\left(w_n, w_{n+lag}\right) = \frac{COV\left(w_n, w_{n+lag}\right)}{\sigma_{w_n}\sigma_{w_{n+lag}}} \tag{2.3}$$

where $\rho\left(w_n, w_{n+lag}\right)$ is the autocorrelation between the degree of soil moisture saturation at the beginning of the time step w_n and the degree of soil moisture saturation at specific lag w_{n+lag}. $COV\left(w_n, w_{n+lag}\right)$ is the covariance between the soil moisture at day n and $n + lag$. w_n and w_{n+lag} are the standard deviation of soil moisture at day n and $n + lag$ respectively. Starting with a simple soil water balance equation (Eq. (2.4)), Koster and Suarez (2001) rewritten Eq. (2.3) and proposed the first autocorrelation based SMM calculation, Eq. (2.7).

$$C_s w_{n+lag,y} = C_s w_{n,y} + P_{n,y} - E_{n,y} - Q_{n,y} \tag{2.4}$$

where C_s, $w_{n,y}$ and $w_{n+lag,y}$ are the water holding capacity of the soil column, degree of soil moisture saturation at day n and day $n + lag$ of year y respectively. $P_{n,y}$, $E_{n,y}$ and $Q_{n,y}$ are the accumulated precipitation, evapotranspiration and streamflow during the time steps $(n, n + lag)$ of year y respectively.

Supported by previous conclusions from Koster and Milly (1997), Koster and Suarez (2001) assumed that both streamflow (normalized by precipitation) and evapotranspiration (normalized by radiation) are linearly dependent on mean soil moisture over the considered time lag.

$$\frac{Q_{n,y}}{P_{n,y}} = a_n \frac{w_{n,y} + w_{n+lag,y}}{2} + b_n \tag{2.5}$$

$$\frac{E_{n,y}}{R_{n,y}} = c_n \frac{w_{n,y} + w_{n+lag,y}}{2} + d_n \tag{2.6}$$

where $w_{n,y}$, $w_{n+lag,y}$, $E_{n,y}$ and $P_{n,y}$ express the same meaning as mentioned above. $\bar{R}_n$ is the radiation. a_n, b_n, c_n and $d_n$are parameters after Koster and Milly (1997). By replacing evapotranspiration and streamflow in Eq. (2.4) by the Eqs. (2.5) and (2.6), Eq. (2.7) was

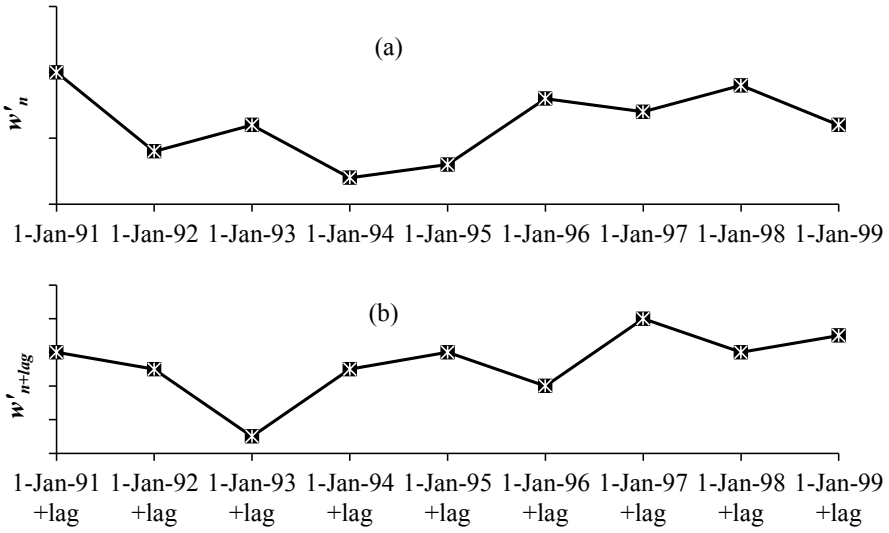

Figure 2.1: Illustration of SMM calculation for 1 January with a specific lag using hypothetical soil moisture time series data. The correlations between all soil moisture anomalies at day n (1 January here) of all years (plotted in panel-a) with the respective anomalies at day $n + lag$ (plotted in panel-b) is the measure of SMM.

developed. The detail step-by-step development of Eq. (2.7) is given in Appendix-A.

$$\rho\left(w_n, w_{n+lag}\right) = \frac{\sigma_{w_n}}{\sigma_{w_{n+lag}}}\left[A_n + \frac{COV\left(w_n, F_n\right)}{\sigma^2_{w_n}}\right] \tag{2.7}$$

with F_n , a forcing term defined by Eq. (2.8):

$$F_{n,y} = \frac{2 - 2\left(a_n \frac{w_{n,y}+w_{n+lag,y}}{2} + b_n\right)}{2C_s + a_n\bar{P}_n + c_n\bar{R}_n}P_{n,y} - \frac{2\left(c_n \frac{w_{n,y}+w_{n+lag,y}}{2} + d_n\right)}{2C_s + a_n\bar{P}_n + c_n\bar{R}_n}R_{n,y} \tag{2.8}$$

and A_n is defined by Eq. (2.9):

$$A_n = \frac{2 - \frac{a_n\bar{P}_n}{C_s} - \frac{c_n\bar{R}_n}{C_s}}{2 + \frac{a_n\bar{P}_n}{C_s} + \frac{c_n\bar{R}_n}{C_s}} \tag{2.9}$$

Koster and Suarez (2001) identified four main drivers of SMM:

1. nonstationarity in the statistics of the forcing, as induced by seasonality: $\frac{\sigma_{w_n}}{\sigma_{w_{n+lag}}}$;
2. reduction in anomaly differences through the functional dependence of evaporation on soil moisture (through the slope c): $\frac{c_n\bar{R}_n}{C_s}$;
3. reduction in anomaly differences through the functional dependence of runoff on soil moisture (through the slope a): $\frac{a_n\bar{P}_n}{C_s}$; and
4. the covariance between initial soil moisture and subsequent atmospheric forcing: $\frac{COV(w_n,F_n)}{\sigma^2_{w_n}}$.

However, Eq. (2.7) with its first controls of SMM holds an important conceptual weakness, as SMM over a given time lag is expressed as a function of soil moisture statistics at the following time step. Subsequent soil moisture may be viewed as a result rather than a control of soil moisture memory. Moreover, it is not monotonically linked with SMM and

can be both high (strong soil moisture control on evaporation or streamflow) or low (strong stochastic forcing) in situations of low soil moisture memory that is difficult to interpret.

B. *Seneviratne and Koster*

To overcome the limitation in Eq. (2.7), more than a decade later, Seneviratne and Koster (2012) revised the framework to form Eq. (2.10). Eq. (2.10) is an extension of the Eq. (2.7), based on the derivation of an expression for the variability of soil moisture at the following time step. See Appendix-A for more details about the development of Eq. (2.10).

$$\rho\left(w_n, w_{n+lag}\right) = \frac{\sigma_{w_n} A_n + \sigma_{F_n} \rho\left(w_n, F_n\right)}{\sqrt{\sigma_{w_n}^2 A_n^2 + 2\sigma_{w_n} A_n \sigma_{F_n} \rho\left(w_n, F_n\right) + \sigma_{F_n}^2}} \tag{2.10}$$

Through Eqs. (2.5) and (2.6), Eq. (2.10) is still dependent on the information of following time step $(n + lag)$. Seneviratne and Koster (2012) also proposed an explicit form of Eq. (2.10) that is fully independent of information from time step $n + lag$ based on the explicit form of Eqs. (2.5) and (2.6).

$$\frac{Q_{n,y}}{P_{n,y}} = a_n w_{n,y} + b_n \tag{2.11}$$

$$\frac{E_{n,y}}{R_{n,y}} = c_n w_{n,y} + d_n \tag{2.12}$$

Based on Eqs. (2.11) and (2.12), the final soil moisture equation was developed, Eq. (2.13).

$$\rho\left(w_n, w_{n+lag}\right) = \frac{\sigma_{w_n}\left(1-\alpha_n\right) + \sigma_{\Phi_n} \rho\left(w_n, \Phi_n\right)}{\sqrt{\left(\sigma_{w_n}\left(1-\alpha_n\right)\right)^2 + 2\sigma_{w_n}(1-\alpha_n)\sigma_{\Phi_n} \rho\left(w_n, \Phi_n\right) + \sigma_{\Phi_n}^2}} \tag{2.13}$$

with

$$\Phi_{n,y} = \frac{1}{C_s}\left[\left(\frac{\bar{P}_n - \bar{Q}_n}{\bar{P}_n}\right) P_{n,y} - \left(\frac{\bar{E}_n}{\bar{R}_n}\right) R_{n,y}\right] \tag{2.14}$$

and

$$\alpha_n = \frac{c_n \bar{R}_n}{C_s} + \frac{a_n \bar{P}_n}{C_s} \tag{2.15}$$

The final equation of SMM contains five controls of SMM:

1. the initial soil moisture variability σ_{w_w};
2. the forcing variability σ_{Φ_n};
3. the correlation between the initial soil moisture and the forcing $\rho\left(w_n, \Phi_n\right)$;
4. the sensitivity of evaporation to soil moisture $\frac{c_n \bar{R}_n}{C_s}$; and
5. the sensitivity of runoff to soil moisture $\frac{a_n \bar{P}_n}{C_s}$.

However, Eq. (2.13) is still model parameter dependent (i.e. a_n, b_n, c_n and d_n) and holds certain assumptions (Eqs. (2.11) and (2.12)).

C. *Orth and Seneviratne*

Based on Seneviratne and Koster (2012) and following the similar steps, an adjusted version of Eq. (2.13) was proposed by Orth and Seneviratne (2012a) to avoid the model parameters and underlying assumptions (streamflow and evapotranspiration dependencies on soil moisture; Eqs. (2.11) and (2.12)). This new framework computes memory using soil moisture, precipitation, streamflow, and evapotranspiration as the direct input, as shown in Eq. (2.16). Orth and Seneviratne (2012a) validated this new approach at five sites across Europe and computed acceptable memory outcomes.

$$\rho\left(w_n, w_{n+lag}\right) = \frac{C_s\sigma_{w_n} + \sigma_{G_n}\rho\left(w_n, G_n\right)}{\sqrt{\left(C_s\sigma_{w_n}\right)^2 + 2C_s w_{w_n}\sigma_{G_n}\rho\left(w_n, G_n\right) + \sigma_{G_n}^2}} \tag{2.16}$$

with

$$G_{n,y} = P_{n,y} - E_{n,y} - Q_{n,y} \tag{2.17}$$

where $\rho\left(w_n, w_{n+lag}\right)$, $C_s, P_{n,y}, E_{n,y}, Q_{n,y}$ and σ_{w_n} express the same meaning as mentioned previously. Here, $\sigma_{G_n}$and $\rho\left(w_n, G_n\right)$ represent the variability of the combined forcing term (based on Eq. (2.17)) and the correlation between the degree of soil moisture saturation at the beginning of the time step $w_n$and the accumulated flux of the combined forcing term between the time steps. The variability terms $(\sigma_{w_n}, \sigma_{G_n})$, and the correlation $\rho\left(w_n, G_n\right)$ are computed for all values at the time step n in all considered years.

In contrast to Delworth and Manabe (1988), autocorrelation-based approaches (Koster and Suarez, 2001; Seneviratne and Koster, 2012; Orth and Seneviratne, 2012a)measure SMM as a non-continuous discrete function. They do not assume an exponential decay of autocorrelations with the lag time. These approaches calculate the correlation between soil moisture values at time step n in all years with the corresponding values at time step $n + lag$. These inter-annual correlations range from zero (no memory) to one (infinite memory). In some regions, the soil moisture state remains static and hardly shows any seasonality, and thus prevail little or with no anomalous conditions. In such cases, the soil would have very little deviation from its mean condition on one day but may regain its normal state (no anomaly) the next day. Therefore, there would be little or no correlation ($\rho = 0$) between the anomalies of these two days. In other words, the anomaly would show no persistence. This condition is termed as no memory.

On the other hand, in some regions, the soil moisture state is very dynamic and displays high seasonality. A strong precipitation event or a prolonged dry spell could lead the soil moisture state to an unusual state (high anomalous condition). In such cases, the soil requires a relatively longer time to regain its normal condition. The anomalous conditions in one day would not change very much on the following day. Therefore, there would be high correlation ($\rho \approx 1$) between the anomalies of these two days. This implies soil moisture anomalies show

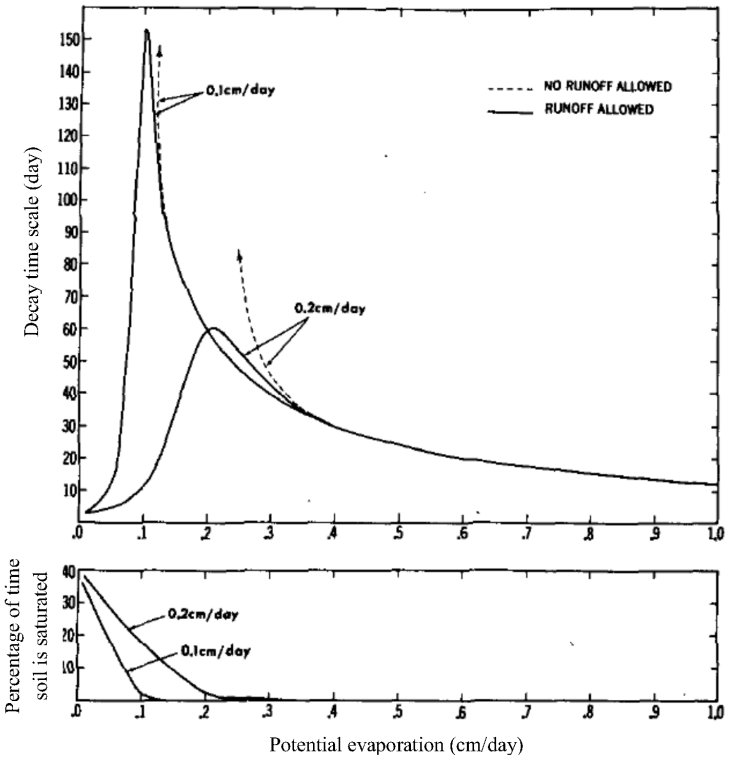

Figure 2.2: Dependency of SMM on potential evapotranspiration, precipitation (daily rates pointing at the graphs) and streamflow (solid vs. dashed lines) as explained by Delworth and Manabe (1988). Strength of SMM are expressed as decay timescale of the autocorrelation function. Adapted from Delworth and Manabe (1988).

strong persistence. These autocorrelation-based approaches allow computing seasonal cycles of SMM because it makes it possible to calculate for any interval $(n, n + lag)$ of the year. This approach is well suited for studying how SMM changes over different lag times.

2.3 Limitations of SMM study

2.3.1 Non-autocorrelation-based approach

Based on Eq. (1.2), the SMM timescale is mostly controlled by potential evapotranspiration and column water holding capacity as shown in Fig. 2.2. The water-holding capacity could be different for layered soils. High potential evapotranspiration results in faster soil moisture anomaly dissipation and thus diminishes soil memory. On the other hand, lower potential evapotranspiration generates strong memories. However, precipitation and runoff also affect the persistence of soil moisture (Orth, 2013). Moreover, this approach excludes the persistence of precipitation and radiation. A moist soil resulting from heavy rainfall could propel evapotranspiration, which, in turn, might provoke further precipitation. This process would therefore assist to maintain an analogous soil moisture anomaly (Koster and Suarez, 2001). Furthermore, this framework identifies meteorological forcing, soil and vegetation characteristics as the major controls of SMM. However, it does not separate the influences of these controls over SMM. It is also important to know which factor affects the SMM timescale

most, and how.

2.3.2 Autocorrelation-based approach

Although autocorrelation-based approaches have succeeded in overcoming the limitations involved with the non-autocorrelation-based approach, the resultant autocorrelation values are not easy to explain. These methods explain the variations of the strength of SMM through the changes of autocorrelation values over different time lags. A change of autocorrelation value from 0.9 to 0.2 only indicates a decrease in memory and does not specify the decrease of the timescale clearly, and vice versa. It also does not clarify whether this relationship is statistically significant. Even if the significance of the relationship is known, how long it will stay significant remains unknown. Therefore, it would be useful once the corresponding timescales of these autocorrelation values are known. Changes in the persistence timescale from 60 days to 45 days (dealing with a single number) would explain the SMM behaviour and seasonality.

2.3.3 Limitations regarding the scarcity of soil moisture data

SMM studies require long-term soil moisture data. Presently, such kinds of data are available only for limited areas (i.e. Russia, China, Mongolia and Illinois, USA). Additionally, the observations were taken mostly on a weekly to half-monthly basis, depending on location and season (Robock et al., 2000). Therefore, to date SMM studies using observed soil moisture data are mostly limited to half-monthly or monthly timescales (Vinnikov and Yeserkepova, 1991; Entin et al., 2000). These studies have assumed the average representation of monthly soil moisture data based on those few records. Moreover, SMM study outcomes based on these observed data sets cannot be compared among the regions due to their inconsistent sampling dates and frequency. Furthermore, soil moisture and other required hydro-climatic data (i.e. precipitation, evapotranspiration, streamflow) do not usually come from a single source, and thus, may have inconsistencies among the data sets. Therefore, any alternative measure of SMM estimation that does not require observed soil moisture data could potentially provides knowledge of SMM for most part of the world where soil moisture data is unavailable. Since it is impossible to create historical records, synthetic soil moisture data could be an alternative option to overcome these limitations (Georgakakos et al., 1995). However, the accuracy and credibility of such data are always a concern.

2.3.4 Limitations regarding the scale of SMM

Up to date SMM studies are mostly done on a global scale with regional averages. However, SMM may vary from basin to basin, even those located within very close latitudes, due to their variations in hydro-climatological properties; hence the basin scale variations are unknown. Several researchers (Liu and Avissar, 1999; Entin et al., 2000; Koster and Suarez,

2001; Wu and Dickinson, 2004; Seneviratne et al., 2006a) have discussed the variations of SMM with latitude, precipitation, evapotranspiration and soil wetness, with no clear relationship overall as basin scale SMM information is still unavailable. Therefore, the basin scale SMM information would assist to establish the relationship between the basin scale SMM and its hydro-climatic (i.e. aridity index) or soil property information, and thus provide important alternative way to gain a rough estimation of SMM.

Chapter 3

Autocorrelations to SMM timescale

3.1 Introduction

Considering the limitations of autocorrelation-based SMM study (refer to section 2.3.2), this chapter proposes a conversion of lagged autocorrelation into a single number SMM timescale considering the statistical significance at 95% confidence level. This approach was applied to compute the SMM timescale for the Spoon river basin in Illinois, USA and was validated against those calculated by Entin et al. (2000) for the Illinois area.

3.2 Materials and Methods

3.2.1 Study area

The Spoon river basin is located in Northeast Illinois with a drainage area of about 2776 km^2 (Fig. 3.1). The stream gauge location of the basin is located at latitude 40.71 and longitude -90.28 (U.S. Model Parameter Estimation Project (MOPEX) ID # 05569500). The annual precipitation and pan evaporation is approximately 896 mm and 1005 mm, respectively.

3.2.2 Data

The basin scale daily precipitation P (daily mean areal precipitation calculated from ground based gauge precipitation), potential evapotranspiration PE (developed from NOAA Evaporation Atlas), and streamflow Q data (developed from USGS hydro-climatic data) were obtained from the MOPEX dataset(Schaake et al., 2006). This study used streamflow and potential evapotranspiration data as an alternative to runoff and actual evapotranspiration data, respectively. A total of 54 years continuous data (1948-2001) was used for this analysis. To facilitate calculating lagged autocorrelations for any interval, this study used Xinanjiang model (Ren-Jun, 1992) (referred to as XAJ and discussed in Section 3.2.4) simulated soil moisture data (validated against the seasonal cycles of observed soil moisture data). The necessity to use the simulated soil moisture data is discussed in Section 3.2.3.

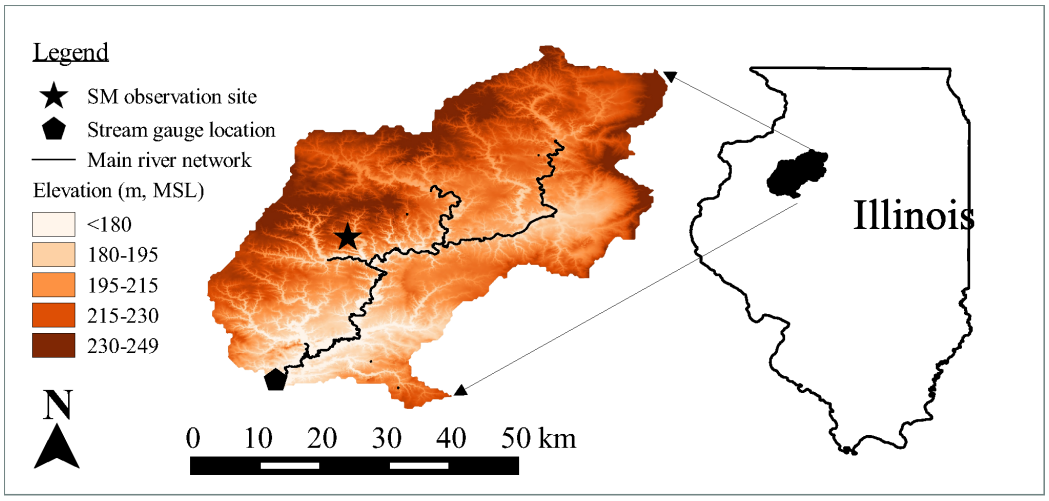

Figure 3.1: Location and digital elevation model map of the Spoon River basin, Illinois, USA (MOPEX ID # 05569500); scale applies for the basin map only.

3.2.3 Why simulated soil moisture data?

To test this new approach, long-term continuous soil moisture data was required. The currently available long-term soil moisture data is not only scarce (available only for limited areas, i.e. Russia, China, Mongolia, and Illinois), but is also non-continuous in nature (recorded mostly on a weekly to half-monthly basis depending on location and season (Robock et al., 2000). Georgakakos et al. (1995) proposed synthetic soil moisture data as an alternative option for overcoming these limitations. Therefore, considering the importance and verifiability of this new approach, this study used the XAJ model simulated soil moisture data. However, the accuracy and credibility of such data is always a concern. To testify the representativeness of the XAJ model simulated soil data against the observed one, the soil moisture data was validated against the observed data set.

3.2.4 Xinanjiang model and its calibration

The Xinanjiang model (Ren-Jun, 1992) is a conceptual hydrological model developed by the Flood Forecast Research Laboratory of the East China Technical University of Water Resources (currently, Hohai University). This model is widely used in China to simulate streamflow within a catchment, particularly for humid and semi-arid regions (Lu and Li, 2014). Runoff in the XAJ model is based on the repletion of storage concept. The runoff does not start to generate until the soil moisture content of the aeration zone reaches its field capacity. Once it reaches field capacity the rainfall excess equals the subsequent runoff without further loss (Ren-Jun, 1992). The XAJ model accepts areal mean precipitation and pan evapotranspiration as the input and provides streamflow records at the outlet of the basin as output. It also generates soil moisture data as an internal model state. The XAJ model was calibrated with the aid of a web-based application available at http://lmj.nagaokaut.ac.jp/~khin/(last accessed on 29 October, 2014) (Khin et al., 2015). This web platform not only allows the

Table 3.1: Parameters in the Xinanjiang model and calibrated value.

Parameter	Physical meaning	Parameter value
C_p	Ratio of measured precipitation to actual precipitation	1.05
C_{ep}	Ratio of potential evapotranspiration to pan evaporation	0.9015
b	Exponent of the tension water capacity curve	0.3
imp	Ratio of the impervious to the total area of the basin	0
WUM	Water capacity in the upper soil layer (mm)	20
WLM	Water capacity in the lower soil layer (mm)	70
WDM	Water capacity in the deeper soil layer (mm)	50
C	Coefficient of deep evapotranspiration	0.3
SM	Areal mean free water capacity of the surface soil layer (mm)	50
EX	Exponent of the free water capacity curve	0.5
KI	Outflow coefficient of the free water storage to inter-flow	0.65
KG	Outflow coefficient of the free water storage to groundwater	0.05
c_s	Recession constant for channel routing	0.5
c_i	Recession constant for the lower inter-flow storage	0.7
c_g	Daily recession constant of groundwater storage	0.985

XAJ model calibration in a user-friendly environment but also provides handy calibration support by suggesting parameter settings after Li and Lu (2014), hydrograph visualization, and calculating Nash-Sutcliffe (NASH) efficiency (Nash and Sutcliffe, 1970). The model was calibrated and validated with a 54-year data set (1948-2001). There were 15 parameters in the XAJ model and a list of them and their calibrated values for the studied basin is presented in Table 3.1. The simulated streamflow was validated against the daily-observed streamflow, and NASH efficiency was recorded. NASH efficiency was calculated based on Eq. (3.1).

$$NASH = 1 - \frac{\sum_{t=1}^{n} \left[Q_o\left(t\right) - Q_s\left(t\right)\right]^2}{\sum_{t=1}^{n} \left[Q_o\left(t\right) - \bar{Q}_o\right]^2} \tag{3.1}$$

where Q_o, Q_s and $\bar{Q}_o$ are the observed annual streamflow, simulated annual streamflow, and average observed annual streamflow, respectively.

3.2.5 Validation of simulated soil moisture data

After calibration and validation of the XAJ model's simulated streamflow, the internal model state of the soil moisture data was selected for validation. The XAJ model produces three layers of soil moisture data. Total soil moisture data for all three layers (XAJ model simulated) were compared to those of the observed total soil moisture data for the top 30cm and top 90cm of the soil layers, recorded at Oak Run station situated near the centre of the studied catchment (station ID #17; longitude -90.15, latitude 40.97; see Fig. 3.1). Illinois soil moisture climatology was obtained from Hollinger and Isard (1994). The observations were taken at roughly half-monthly intervals. XAJ soil moisture values were chosen for the corresponding soil moisture observation date and validated thereafter. A total of 321 soil moisture observations between 1st June, 1981 and 15th June, 1998 were used for this vali-

dation purpose. Simulated soil moisture data was further verified by calculating one month (30 days) of lagged autocorrelations for the 28$^{\text{th}}$ day of every month between March and October using both observed (top 30cm and top 90cm) and simulated soil moisture data, and compared thereafter.

3.2.6 Calculation of soil moisture autocorrelation and SMM timescale

This study calculated soil moisture autocorrelation for several lags using Eq. (2.16) as proposed by Orth and Seneviratne (2012a). Thereafter, the calculated lagged autocorrelation values were converted into SMM timescale (τ_{SMM}) considering statistical significance at 95% confidence level. The autocorrelations for the 28$^{\text{th}}$ day (the 28$^{\text{th}}$ day is the most frequent soil moisture observation day) of every month between March and October were calculated for all lag days (starting from a minimum 5-day lag) until the autocorrelation value crossed a minimum threshold. This minimum threshold value was set to be equal to the critical ρ value at 95% confidence level for one-sided test (alternate hypothesis, $H_a : \rho > 0$. The critical ρ value was calculated after Mitchell et al. (1966), Eq. (3.2). The SMM timescale τ_{SMM} was assumed to be equal to the highest number of lag days that produced a significant correlation at 95% confidence level.

$$\rho_{0.95} = \frac{-1 + 1.645\sqrt{N-2}}{N-1} \tag{3.2}$$

where $\rho_{0.95}$ is the critical ρ value at 95% confidence level, and N is the number of pairs of data (length of data record). For example, the studied basin with 45-year data has a threshold ρ value of 0.22243. The study accepted the alternate hypothesis until the lag day that satisfied the criteria of $\rho \geq \rho_{0.95}$. Once it produced $\rho \geq \rho_{0.95}$, it stopped the memory calculation whether it became significant on the next day or not. The last lag-day at which it satisfied the criterion was counted as the length of the SMM timescale for that particular day (i.e., the 28$^{\text{th}}$ day of any month in this study). Figure 3.2 shows that the memory for one particular day March 28$^{\text{th}}$ was counted as 31 days because the 32$^{\text{nd}}$ day-lagged autocorrelations crossed the threshold value. The study ignored the lagged autocorrelation for less than five days. In this case, the 5$^{\text{th}}$ day ρ value was $< \rho_{0.95}$, and zero memory was recorded for that particular day. Similarly, τ_{SMM} for the 28$^{\text{th}}$ day of every month from March to October (winter months were excluded to avoid impacts of snow) was calculated. Later, basin average τ_{SMM} was calculated by averaging these eight months' (March to October) SMM timescales.

3.3 Results and Discussions

3.3.1 Hydrograph

The model calibration attained good agreement between the observed streamflow and simulated streamflow (Fig. 3.3). The annual (between the annual observed streamflow and

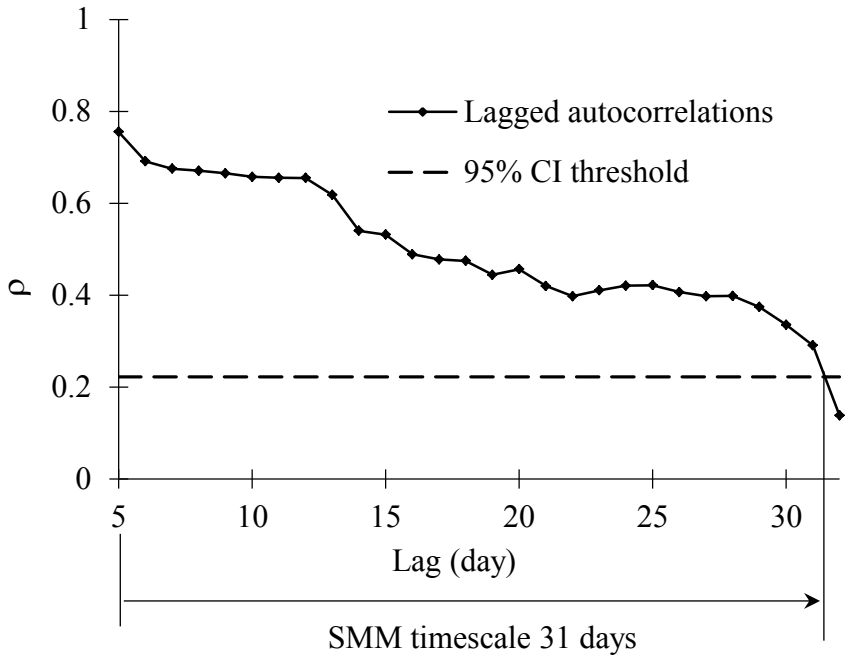

Figure 3.2: The methodology of SMM timescale estimation explained.

annual simulated streamflow) and daily (between the daily observed streamflow and daily simulated streamflow) $NASH$ efficiencies were 0.91 and 0.59, respectively.

3.3.2 Validation of simulated soil moisture data

The objective of this study was to compute the autocorrelation values using degree of soil moisture saturation (refer to Eq. (2.16)). Autocorrelation accounts for the dissipation of soil moisture anomalies. Therefore, the question is how well the XAJ model represented the temporal anomalies of soil moisture. Validation of the XAJ simulated soil moisture data suggested that the model could fairly well reproduce the absolute soil moisture value for the top 30cm soil layer (daily $NASH$ efficiency 0.57). However, it underestimated the top 90cm by a nearly constant offset (Fig. 3.4a). On the other hand, the model represented the temporal anomalies (calculated from the mean value of 321 soil moisture observations between June 1st, 1981 and June 15th, 1998) of soil moisture for all the soil layers up to 90cm in depth (see Fig. 3.4b). The correlation coefficients between the simulated soil moisture and observed soil moisture are 0.83 and 0.85 for top 30cm and top 90cm, respectively. Hollinger and Isard (1994) suggested that the study area is mainly dominated by agriculture (i.e. mainly corn fields). Due to the shallow rooting effect, deeper soil layers beyond 30cm may not be hydrologically active throughout the year. Hence, the actual soil moisture could be higher in the deeper layers. Similarly, Ren et al. (2006) argued that the XAJ model could not only simulate the streamflow but also represent soil moisture data. Entin (1999) noted that most land surface models are incapable of simulating absolute soil moisture values but they are quite able to capture the seasonal cycles of soil moisture. A few studies (Georgakakos et al., 1995; Wu et al., 2007) have been conducted using simulated soil moisture data for

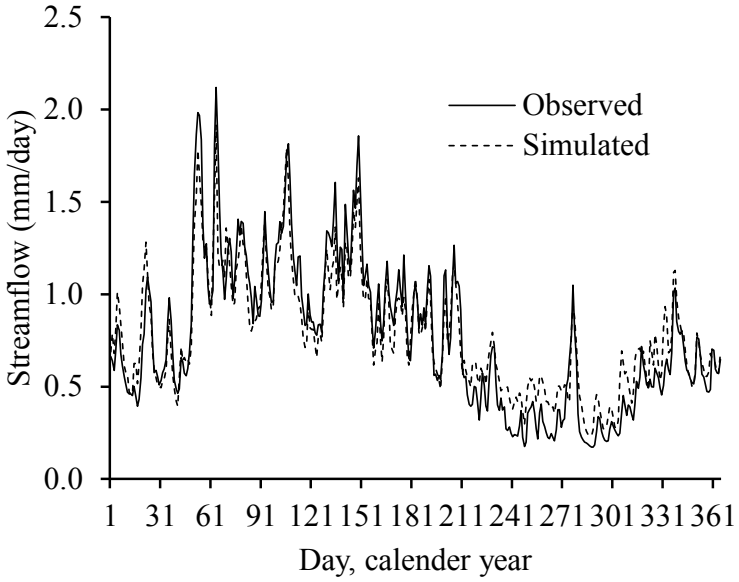

Figure 3.3: Validated daily hydrograph of the Spoon river basin, Illinois, USA (MOPEX ID # 05569500).

similar studies.

3.3.3 Autocorrelations and timescale

Calculated one-month (30-day) lagged autocorrelation (based on Eq. (2.16)) using observed soil moisture (top 30cm and top 90cm) and simulated soil moisture (total soil moisture) confirmed that simulated soil moisture corresponded fairly well with those of observed top 30 cm ($R^2 = 0.88$), Fig. 3.5a. In contrast, it underestimated the autocorrelations for the top 90cm soil layers ($R^2 = 0.77$). However, the overall autocorrelation values were higher for top 90cm soil layers compared to those of the top 30cm. The higher the values of autocorrelations, the deeper the soil layers seemed to be natural, and several studies reported the same. Entin et al. (2000); Wu and Dickinson (2004) and Vinnikov et al. (1996) documented that soil moisture persistence increased with the soil depth.

The basin average τ_{SMM}, calculated based on the new approach, was counted as 56.13 days. The calculated SMM timescale was highly consistent with the previous estimation for this region. Entin et al. (2000) reported the average memory for 18 stations in Illinois as 1.8-2.1 months (calculated based on Delworth and Manabe (1988) approach, Eq. (2.2)). The basin indicated the highest memory in April and August and lowest memory in March. Based on the aridity (ratio of annual potential evapotranspiration over actual precipitation, PE/P =1.16), the Spoon river basin can be considered a dry basin. Entin et al. (2000); Orth and Seneviratne (2012a) and Rahman et al. (2015) argued that dry basins demonstrate higher memory in the winter months. This study excluded the winter months' memory to avoid the impacts of snow. Therefore, the actual memory would be slightly higher than this estimation.

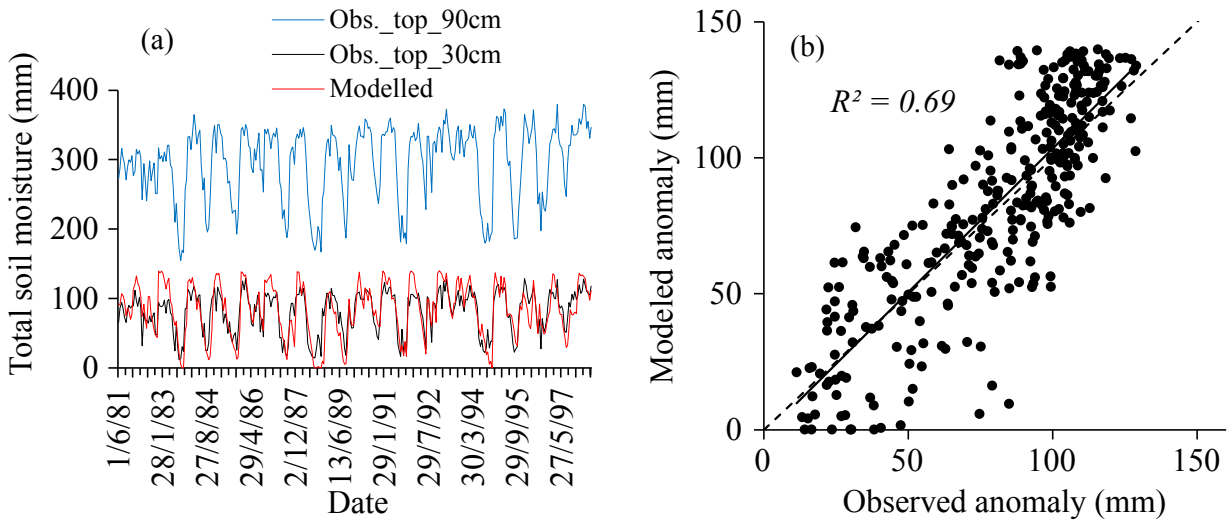

Figure 3.4: Validation of XAJ model simulated soil moisture against observed soil moisture for the Spoon river basin, Illinois, USA (MOPEX ID: 05569500):(a) Absolute soil moisture; (b) Soil moisture anomalies (top 90cm); solid and dashed lines are regression fit and 1 by 1 lines, respectively.

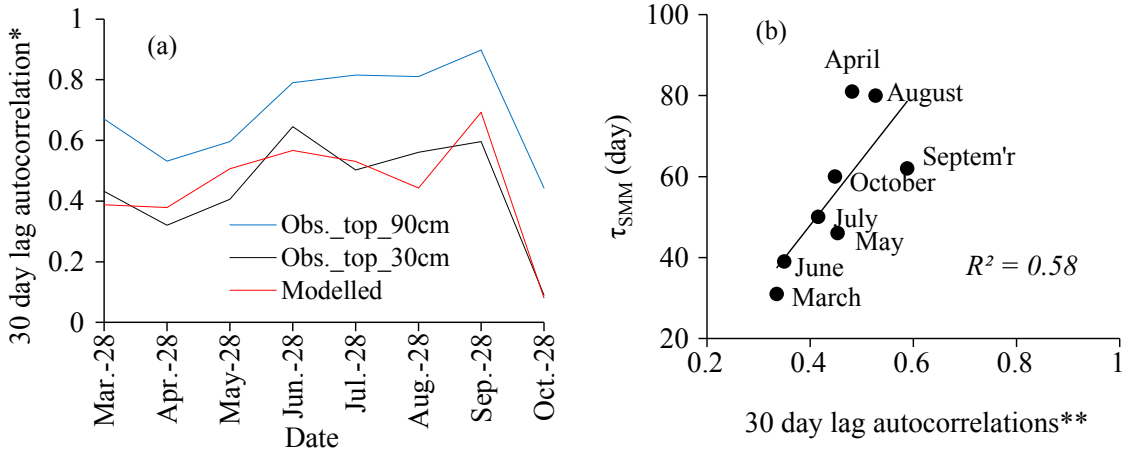

Figure 3.5: Soil moisture autocorrelations and memory timescales: (a) 30-day-lagged autocorrelations using simulated soil moisture, total soil moisture for top 30cm and top 90cm; (b) Estimated memory timescales and 30-day-lagged autocorrelations. *Autocorrelations calculated from simulated soil moisture data limited to soil moisture observation duration only (11-17 years); ** Autocorrelation calculated from 45 years simulated soil moisture data (1956-2000).

The calculated monthly τ_{SMM} was plotted against the one-month (30-day) lagged autocorrelation coefficients (for the 28^{th} day of every month between March and October) in Fig. 3.5b. The high R^2 (0.58) value indicates the consistency in representation of the strength of autocorrelations through the newly converted SMM timescales. Figure 3.5b also reveals the persistence of soil moisture at different angles. The 30-day-lagged autocorrelation values for May ($\rho = 0.45$) and October ($\rho = 0.45$) were almost the same. However, this study suggested that the strength of autocorrelation in October would remain significant for 60 days. On the other hand, the autocorrelation strength of May seemed to decline quickly and remained significant for only 46 days. Similarly, in September the 30-day-lagged autocorrelation ($\rho = 0.59$) was higher compared to that for August ($\rho = 0.53$). However, the relatively lower-strength autocorrelation in August seemed to remain significant for longer (80 days) when compared to the autocorrelation for September (60 days). Consequently, comparing only two autocorrelation values may not necessarily explain the actual strength of soil moisture memory. This new SMM timescale adds more information compared to simple autocorrelations. This new SMM timescale enables better understanding of the behaviour of soil moisture persistence and its seasonality.

3.4 Conclusions

Soil moisture is an important component in climate and weather predictions due to its special persistence characteristics. Any anomalous condition in the soil moisture state tends to persist long after the event that caused the anomaly. This behaviour is commonly termed as soil moisture memory. There are two main approaches to computing the strength of this memory. The latest approach measures this memory in the form of autocorrelations and ranges between 0 and 1. The strength of this estimated memory was judged by the autocorrelation score. The higher the score, the stronger the memory was. To compare the memory of two regions or basins or months, autocorrelation scores were used. However, this autocorrelation score can be confusing when attempting to represent the general nature of the basin or area. Moreover, it does not necessarily explain whether it is statistically significant or not. Furthermore, a significant autocorrelation score for a particular lag-day reveals the strength of that particular lag and does not clarify how long (until how many lag-days) this relation would remain significant.

To improve the understanding of the effects of changes in autocorrelation values, this study proposed converting autocorrelation values into SMM timescales considering the statistical significance. This study estimated the SMM timescale of the Spoon river basin in Illinois using observed precipitation, potential evaporation, and streamflow and simulated soil moisture data. The estimated SMM timescale (56.13 days) was highly consistent with those of regional estimation. The consistency of converting the autocorrelation values into SMM timescale was confirmed with high R^2(0.58) values between 30-day-lagged autocorrelations and SMM timescales. This SMM timescale is easy to comprehend and conveys more information than simple autocorrelation values. This study suggested that similar scores in

autocorrelation values calculated for two different months or two different basins could lose their strength of autocorrelation with a different pace. Despite having similar scores, one relationship may remain significant for longer than the other. This new SMM timescale not only corresponds to the autocorrelation scores, but also suggests the duration of its robustness of statistical significance. Finally, it offers easy comparison between two regions or river basins with just a single number.

Chapter 4

Variability of SMM for wet and dry basins

4.1 Introduction

The absence of basin scale SMM information and its necessities are discussed in section 2.3.4. The basin scale SMM information would open the window to find its relationship with basin's hydro-climatic or physical properties. Any relationship between SMM and basin properties could potentially overcome the soil moisture data limitation discussed in section 2.3.3. Aiming to find and alternative way to solve this data limitation, this chapter analyses basin scale SMM of 26 river basins over United States. It computes basin scale SMM timescale using a well-known hydrological model's simulated soil moisture data (Xinanjiang model; discussed in section 3.2.4). The model was run with observed precipitation and potential evapotranspiration data to simulate streamflow. The simulated streamflow was then validated against daily observed streamflow. Finally, the internal model state of soil moisture data was picked for SMM studies. The applicability of the model simulated soil moisture data was verified and discussed in section 3.3.2. Based on the result of 26 river basins, this chapter describes the seasonality of the SMM timescale and how it is affected by the basin's dryness. Moreover, it investigates and analyses the relationship between the SMM timescale and the aridity index (ratio of annual potential evaporation over annual precipitation).

4.2 Materials and method 8

A summary and flow chart of the overall methodology of SMM analysis is shown in Fig. 4.1. The details are described in the followings.

4.2.1 Calculation of soil moisture autocorrelation and SMM timescale

This study calculated soil moisture autocorrelations based on the equation (Eq. (2.16)) proposed by Orth and Seneviratne (2012). The SMM timescale for the 1st day of every month was calculated based on statistical significance of autocorrelations at 95% confidence

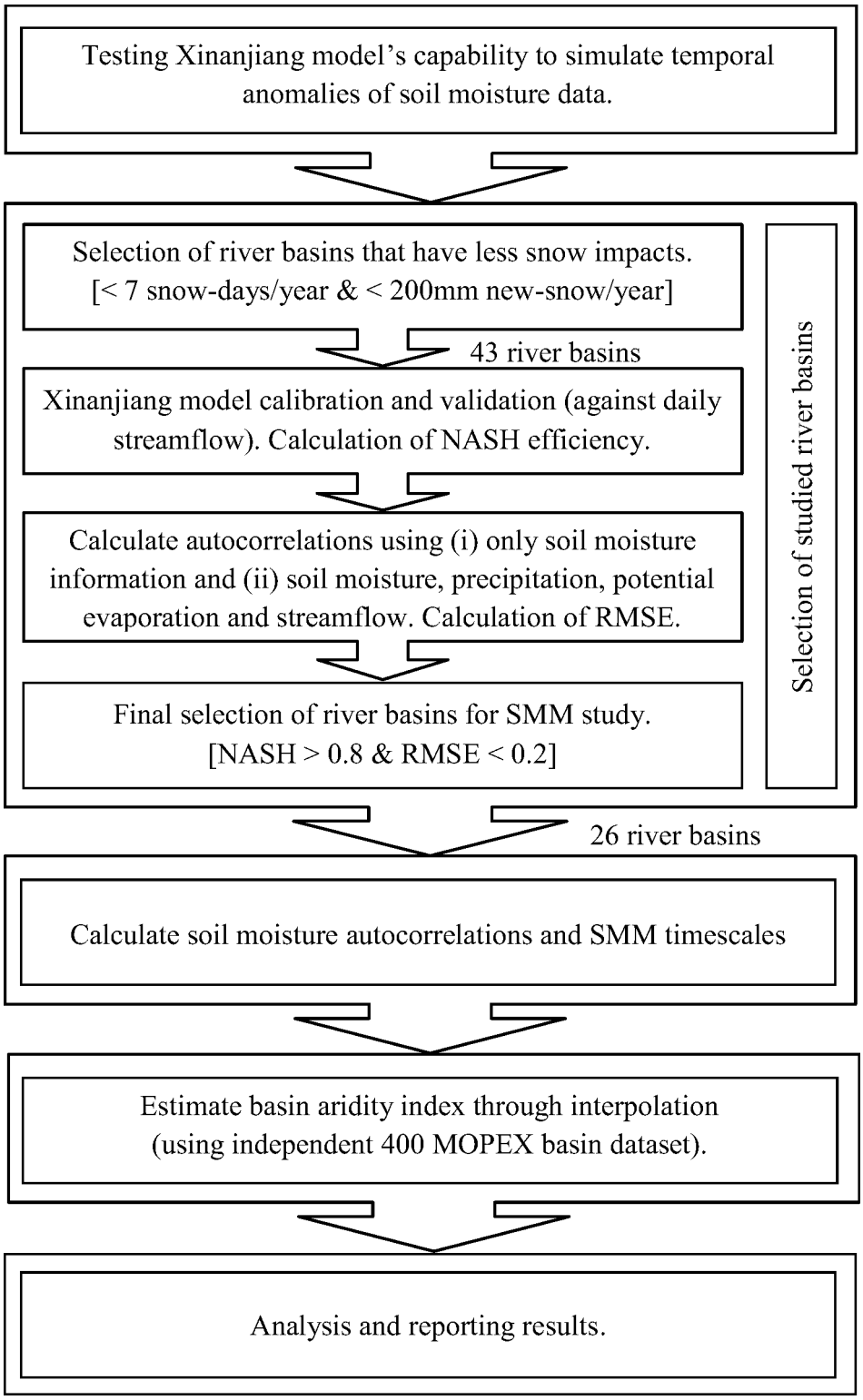

Figure 4.1: Flow chart showing overall methodology.

level similarly those of discussed in Chapter 3. The resultant twelve SMM timescale values (January to December) representing every month were then averaged to represent the SMM timescale of a particular basin.

4.2.2 Data

The basin scale daily precipitation, P (daily mean areal precipitation calculated from ground based gauge precipitation), potential evapotranspiration, PE (developed from NOAA Evaporation Atlas), and streamflow, Q data (developed from USGS hydro-climatic data) were used in this study. These data were obtained from the U.S. Model Parameter Estimation Project (MOPEX) data set freely accessible at ftp://hydrology.nws.noaa.gov/ (accessed on 19 October 2013) (Schaake et al., 2006). It is to be noted that this study used streamflow and potential evapotranspiration data instead of runoff and actual evapotranspiration data respectively. MOPEX provides 54 years continuous data (1948-2001) for most of the analysed basins; the few exceptions contain missing records and hold fewer continuous records. This study simulated the streamflow for every basin with the XAJ model (Ren-Jun, 1992) (discussed in section 3.2.4) and validated the results against the daily observed streamflow. For the purpose of the XAJ model run and subsequent SMM computation, only continuous data was used.

4.2.3 Selection of studied basins

The selection of river basins consists of four steps.

Step-1

Since the autocorrelation equation (Eq. 2.16) does not consider snow, the studied river basins were selected carefully to avoid snow impacts. To avoid snow impacts, a total of 15 USA states were selected on the basis of annual average total snow-days (a snow-day is a day that receives at least 2.5mm snow/day) and annual average total new snow-depths. The selection was based on 30 year climate normal (1981-2010) released by NOA's National Climatic Data Centre (NCDC), accessible at http://www.ncdc.noaa.gov/oa/climate/normals/usnormals.html (accessed on 13 November 2013). Fourteen states out of these 15 have <7 snow-days and receive <200mm of total new snow per year and one state has 10 snow-days and receives 373mm total new snow per year.

Step-2

A total of 43 MOPEX river basins located within those 15 states were then selected for XAJ model simulation (details are in section 4.2.4), considering data availability and prior calibration experience by Khin et al. (2015). The XAJ model was calibrated separately for every basin and simulated streamflow for all 43 river basins. The simulated streamflow was validated against daily observed streamflow. Nash-Sutcliffe (NASH) efficiency (Nash and

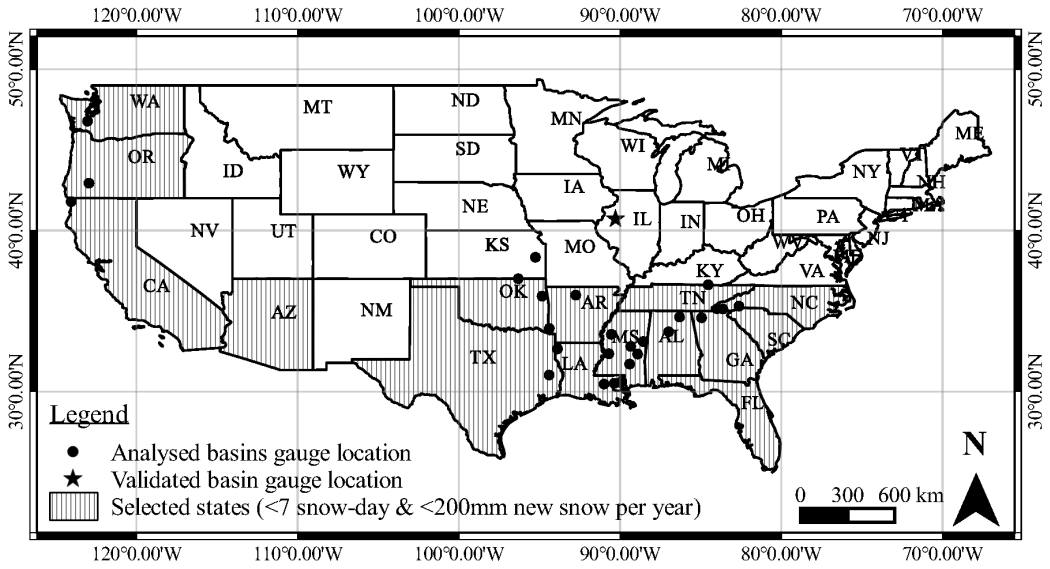

Figure 4.2: Stream gauge location map of analyzed basins over USA mainland.

Sutcliffe, 1970) for annual streamflow was recorded. NASH efficiency was calculated based on Eq. (3.1).

Step-3

To verify the applicability of Eq. (2.16) and XAJ model simulated soil moisture data, calculated autocorrelation coefficient values (based on Eq. (2.16) using soil moisture, P, PE and Q data as input) were validated against directly calculated correlation coefficient values (based on Eq. (2.3) using only soil moisture data as input). It is to be noted that Eq. (2.3) was only used to verify the applicability of Eq. (2.16) for SMM calculation. All the SMM timescales discussed in this chapter are based on Eq. (2.16) only.

The Root Mean Square Error (RMSE) for 5, 10, 15, 20, 25 and 30 day-lagged soil moisture autocorrelations (Eq. (2.3) and Eq. (2.16)) was calculated based on Eq. (4.1).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(\rho_{obs,i} - \rho_{cal,i}\right)^2}{n}} \tag{4.1}$$

where $RMSE$ is root mean square error, ρ_{obs} is observed autocorrelation coefficient (Eq. (2.3)), ρ_{cal} is calculated autocorrelation coefficient (Eq. (2.16)) values at time and n is the number of observations.

Step-4

Finally, those basins that attained annual $NASH$ efficiency over 0.8 (step-2) and $RMSE$ less than 0.2 (step-3) were picked for SMM analysis. Figure 4.2 shows the stream gauge locations of 27 river basins (26 analysed and 1 validated) over the USA mainland map. A list of studied MOPEX basins, locations and basic characteristics are presented in Table 1.

Table 4.1: Studied MOPEX basins, locations and basic characteristics. Dry basins (aridity index >0.9) are marked with bold face, Italic font style.

MOPEX ID	Location			Area (sq.km)	Ave. P (mm/year)	Ave. PE (mm/year)	Ave. snow-days (day/year)	Ave. total new snow (mm/year)
	Long.	Lat.	State					
11532500	-124.05	41.79	CA	1577	2687	740	0.00	0
12027500	-123.03	46.78	WA	2318	1599	579	3.00	127
03550000	-83.98	35.14	NC	269	1846	771	3.90	193
03504000	-83.62	35.13	NC	135	1893	762	3.90	193
03443000	-82.62	35.30	NC	740	2156	817	3.90	193
03410500	-84.53	36.63	TN	2471	1389	817	6.20	160
02387500	-84.94	34.58	GA	4144	1480	901	0.70	18
03574500	-86.31	34.62	AL	829	1467	941	0.80	41
14308000	-122.95	42.93	OR	1163	1347	805	2.20	76
07378500	-90.99	30.46	LA-MS	3315	1594	1077	0.60	23
07375500	-90.36	30.51	LA-MS	1673	1633	1074	0.60	23
02492000	-89.90	30.63	LA-MS	3142	1583	1071	0.60	23
02456500	-86.98	33.71	AL	2292	1425	982	0.80	41
02472000	-89.41	31.71	MS	1924	1492	1060	0.60	23
02475500	-88.91	32.33	MS	956	1447	1056	0.60	23
02482000	-89.34	32.80	MS	2341	1447	1056	0.60	23
02448000	-88.56	33.10	MS	1989	1421	1057	0.60	23
07290000	-90.70	32.35	MS	7283	1435	1073	0.60	23
07056000	-92.75	35.98	AR	2147	1180	916	3.80	132
07288500	-90.54	33.55	MS	1987	1381	1112	0.60	23
07340000	-94.39	33.92	OK	6895	1329	1156	5.60	198
07197000	***-94.84***	***35.92***	***OK***	***795***	***1162***	***1113***	***5.60***	***198***
07348000	***-93.88***	***32.65***	***LA***	***8125***	***1173***	***1223***	***0.10***	***0***
05569500*	***-90.28***	***40.71***	***IL***	***2776***	***896***	***1005***	-	-
08033500	***-94.40***	***31.02***	***TX***	***9417***	***1100***	***1308***	***1.30***	***4***
06914000	***-95.25***	***38.33***	***KS***	***865***	***957***	***1206***	***10.0***	***373***
07172000	***-96.32***	***37.00***	***OK-KS***	***1153***	***898***	***1301***	***6.90***	***198***

*Indicates the validated river basin at Illinois State, USA.

The *NASH* efficiencies and *RMSE* are reported in Table 3.

4.2.4 Xinanjiang model and its calibration

The Xinanjiang model(Ren-Jun, 1992) (discussed in section 3.2.4) was calibrated with the aid of a web-based application, accessible at http://lmj.nagaokaut.ac.jp/~khin/ (last accessed on 20 October 2014) (Khin et al., 2015). This web platform not only allows the user to run the XAJ model in a user friendly environment, but also provides handy calibration support by suggesting parameter settings after Li and Lu (2014), hydrograph visualization and calculating *NASH* efficiency. The model was calibrated with a 54 year data set (1948-2001 for most basins) and validated against the last 20 years. A list of XAJ model parameters and their ranges are presented in Table 2.

Table 4.2: Range of calibrated parameters in the Xinanjiang model.

Parameter	Physical meaning	Parameter value
C_p	Ratio of measured precipitation to actual precipitation	0.8-1.2
C_{ep}	Ratio of potential evapotranspiration to pan evaporation	0-2.0
b	Exponent of the tension water capacity curve	0.1-0.3
imp	Ratio of the impervious to the total area of the basin	0-0.005
WUM	Water capacity in the upper soil layer (mm)	5-20
WLM	Water capacity in the lower soil layer (mm)	60-90
WDM	Water capacity in the deeper soil layer (mm)	10-100
C	Coefficient of deep evapotranspiration	0.1-0.3
SM	Areal mean free water capacity of the surface soil layer (mm)	1-50
EX	Exponent of the free water capacity curve	0.5-2.5
KI	Outflow coefficient of the free water storage to inter-flow	0-0.7; KI+KG=0.7
KG	Outflow coefficient of the free water storage to groundwater	0-0.7; KI+KG=0.7
c_s	Recession constant for channel routing	0.5-0.9
c_i	Recession constant for the lower inter-flow storage	0.5-0.9
c_g	Daily recession constant of groundwater storage	0.9835-0.998

4.2.5 Calculation of basin aridity index

The aridity index value, ζ was calculated from the independent set of precipitation and potential evapotranspiration data. The aridity index was estimated by interpolating aridity index values of 400 MOPEX river basins (excluding the basins analysed in this article). The aridity index was calculated after Li and Lu (2014), Eq. (4.2).

$$\zeta = \frac{PE}{P} \tag{4.2}$$

where, ζ, PE and P are the aridity index, mean annual potential evaporation and ground based mean annual areal precipitation respectively.

The interpolation was done through the Kriging method with the aid of the Spatial Analyst tool of the ArcGIS 10.0 version. The interpolated aridity index values showed a high agreement with those calculated (using the same precipitation and potential evapotranspiration data used for SMM analysis), with R^2 value of 0.99. Analysis suggests that the basins can be categorized into two groups based on SMM validation, the pattern of SMM seasonality, and SMM timescale. To simplify the analysis and discussion, this study defines dry basins and wet basins based on their aridity scores. Basins having an aridity index of less than 0.9 are called wet basins, while the rest are referred to as dry basins in the following sections.

4.2.6 Optimization of regression equation

Theoretically, we assume the SMM timescale would approach zero and lose all the memories when a basin aridity index approaches zero. To understand the behaviour of the SMM timescale beyond the aridity ranges of the analysed basins, the regression equation between

Table 4.3: Basin-wise summary of soil moisture memory analysis. Dry basins (aridity index >0.9) are marked with bold face, Italic font style.

MOPEX ID	Data length of SMM calculation (year)	Area (sq.km)	NASH	RMSE	Aridity index (ζ)	τ_{SMM} (day)		
						Min. (month)	Max. (month)	Mean
11532500	45	1577	0.89	0.05	0.29	0 (December)	56 (May)	19.67
12027500	45	2318	0.88	0.04	0.39	0 (December)	53 (July)	21.83
03550000	45	269	0.85	0.05	0.40	0 (January)	64 (October)	21.25
03504000	45	135	0.91	0.05	0.40	0 (January)	28 (August)	11.08
03443000	45	740	0.89	0.04	0.46	0 (January)	38 (August)	15.25
03410500	45	2471	0.84	0.04	0.58	0 (January)	93 (September)	32.08
02387500	45	4144	0.87	0.05	0.61	5 (January)	78 (October)	24.00
03574500	45	829	0.80	0.16	0.64	0 (January)	58 (October)	23.83
14308000	45	1163	0.83	0.03	0.68	0 (January)	53 (August)	20.58
07378500	44	3315	0.92	0.02	0.70	0 (February)	79 (October)	36.67
07375500	44	1673	0.91	0.03	0.71	0 (March)	64 (October)	29.17
02492000	44	3142	0.87	0.04	0.71	0 (March)	65 (October)	31.08
02456500	45	2292	0.90	0.02	0.72	0 (February)	82 (October)	30.83
02472000	45	1924	0.89	0.02	0.76	0 (February)	153 (Septemb.)	46.17
02475500	45	956	0.87	0.01	0.77	0 (February)	200 (August)	64.33
02482000	45	2341	0.84	0.03	0.79	0 (February)	197 (August)	55.50
02448000	34	1989	0.89	0.02	0.80	0 (February)	85 (August)	36.67
07290000	32	7283	0.87	0.04	0.80	0 (February)	118 (August)	40.42
07056000	45	2147	0.93	0.04	0.81	5 (February)	70 (October)	36.50
07288500	35	1987	0.90	0.12	0.86	0 (March)	49 (August)	26.00
07340000	38	6895	0.92	0.07	0.88	5 (February)	53 (September)	27.25
07197000	***45***	***795***	***0.80***	***0.01***	***0.94***	***28 (August)***	***118 (Novem.)***	***68.67***
07348000	***36***	***8125***	***0.81***	***0.04***	***1.09***	***13 (May)***	***140 (July)***	***74.83***
05569500*	***45***	***2776***	***0.91***	***0.01***	***1.16***	-	-	-
08033500	***45***	***9417***	***0.80***	***0.02***	***1.19***	***31 (May)***	***172 (June)***	***101.50***
06914000	***45***	***865***	***0.88***	***0.01***	***1.34***	***28 (May)***	***217 (August)***	***124.08***
07172000	***45***	***1153***	***0.88***	***0.00***	***1.54***	***56 (May)***	***209 (Decem.)***	***132.92***

*Indicates the validated river basin at Illinois State, USA.

the aridity index and SMM timescale was optimized using the solver function of Microsoft Excel (version 10.0).

4.3 Results and discussions

4.3.1 Hydrograph

Based on the basins selection criteria mentioned in section 4.2.3, all the analysed basins attained a $NASH$ efficiency of no less than 0.8 in the annual scale. Moreover, the validated daily hydrograph between the observed and simulated streamflow shows a good agreement (for example, see Fig. 4.2 and Table 3). The annual $NASH$ efficiencies of analysed basins range from 0.80 to 0.93.

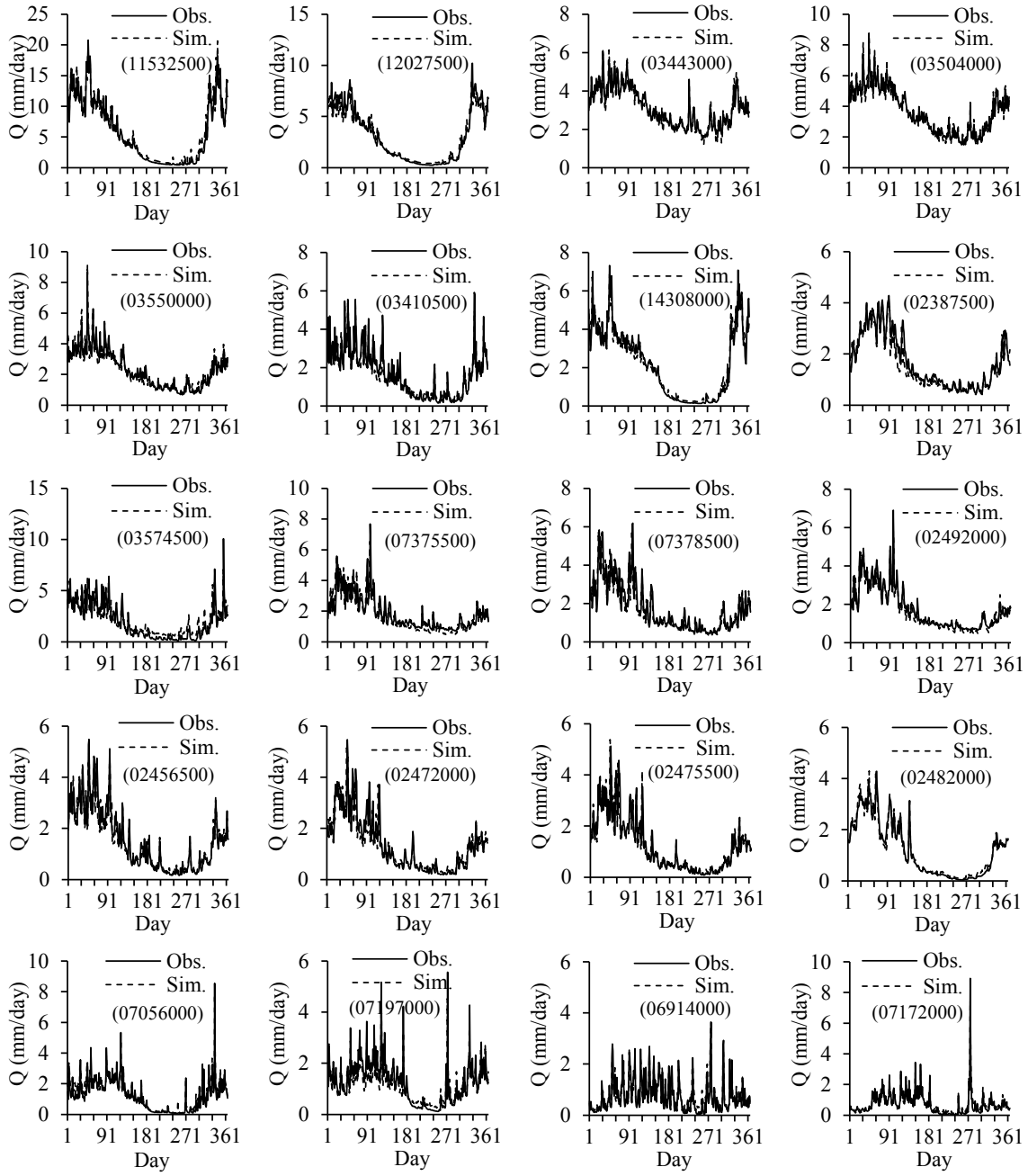

Figure 4.3: Validated daily hydrograph of 20 analyzed river basins. MOPEX ID is presented in parenthesis.

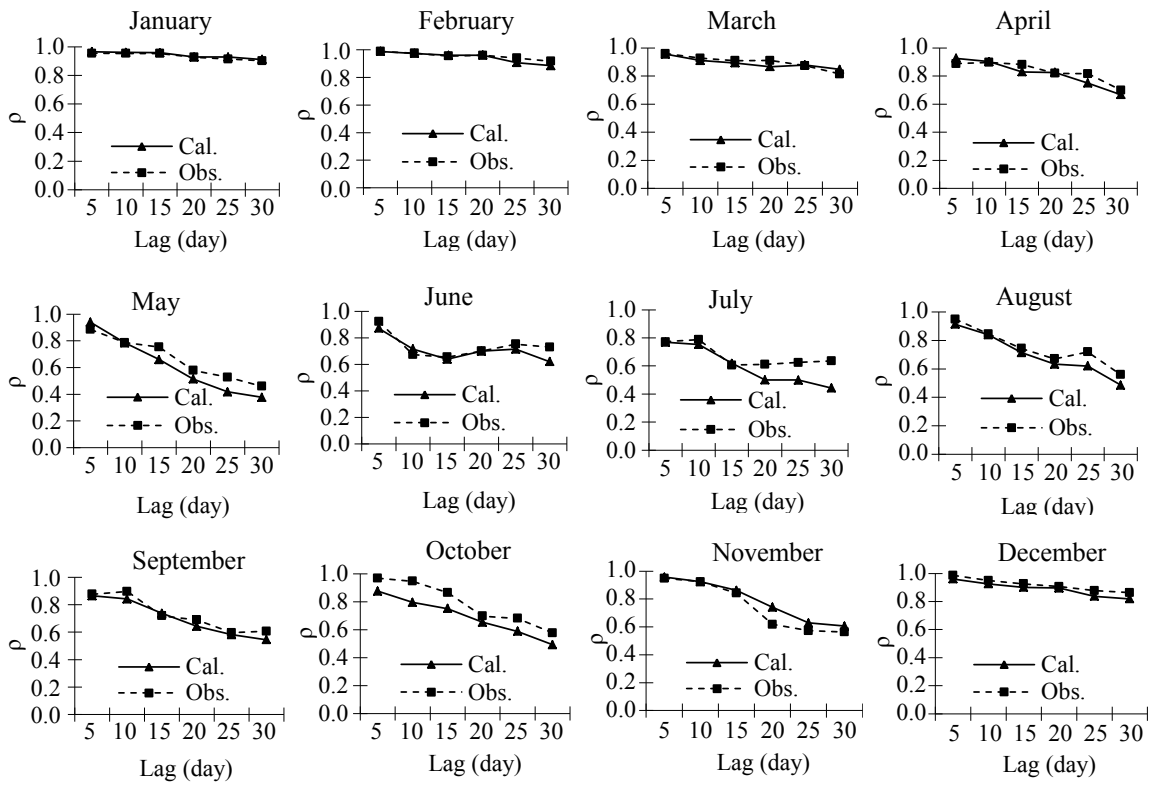

Figure 4.4: Validation of monthly memory (autocorrelation calculated by Eq. (2.16) against observed autocorrelations (by Eq. (2.3)). The Caney River basin, Kansas, USA, MOPEX ID: 07172000.

4.3.2 Validation of autocorrelation equation

The computed SMM of the Caney River basin, Kansas, USA for all the months for time lags between 5 and 30 days is presented in Fig. 4.4. The SMM exhibits a declining trend with the increase of time lag, except for a few minor fluctuations. The extents of these declines are generally not the same for all the months and in dry or wet basins. The SMM normally diminishes at a faster rate with the increase of lag time in the wet basins, while it weakens slowly in dry basins. Similarly, a longer persistence of dry soil moisture anomaly was reported by Orth and Seneviratne (2012b).

The faster rate of SMM decay in the wet basins could be a function of degrees of anomaly conditions, precipitation frequency and the amount of precipitation per event. Usually, the wet basins have smaller soil moisture anomalies, since soil remains saturated or near saturation during most of the time of year (cannot get any wetter). On the contrary, the dry basin anomalies could be very high due to any unusual high precipitation event or prolonged dry spell. The rains are more frequent and heavier in the wet basins, and thus dissipate the anomaly conditions in a shorter time. Inversely, it might take a long time to dissipate similar anomalies in drier basins, as they receive few rainfall events with lesser quantities per event. Consequently, the anomaly dissipates slowly in dry basins. Similar to $NASH$ efficiencies, analysed basins exhibit better agreement between the autocorrelations calculated by Eq. (2.16) and Eq. (2.3). The average $RMSE$ is 0.041 (minimum 0.004 and

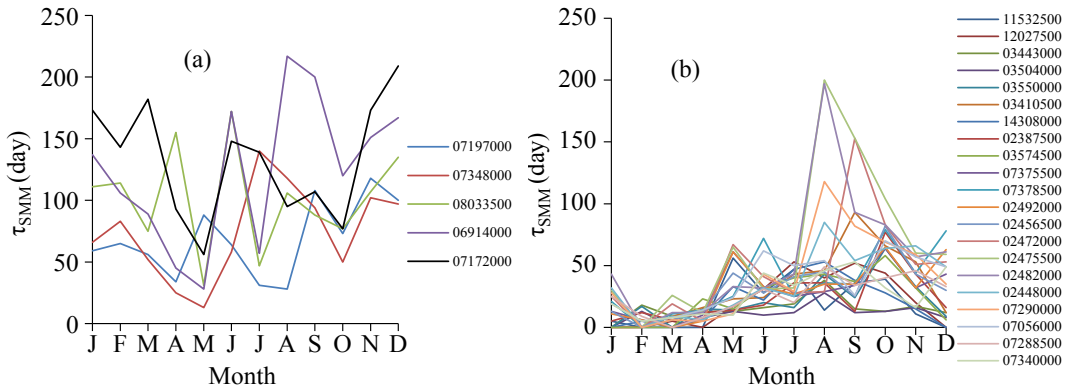

Figure 4.5: Memory timescales and seasonal cycles; (a) dry basins ($\zeta > 0.9$) (b) wet basins ($\zeta < 0.9$).

maximum 0.156). The validation of autocorrelation coefficients confirms the applicability of the equation to calculation of SMM, using precipitation, potential evapotranspiration and streamflow as input. The equation skilfully captures the SMM and its decay over increase of time lag (see Fig. 4.4).

4.3.3 SMM timescale and seasonality

Basin average SMM timescales ranges were calculated from 11 days to 133 days. The SMM timescale varies in wet ($\zeta < 0.9$) and dry ($\zeta > 0.9$) basins. The mean memory of wet and dry basins is about 31 and 100 days respectively. A summary of the basin-wise SMM timescale is given in Table 3. The SMM timescale also showed strong seasonality. However, the seasonal cycles are not the same in dry or wet basins (Fig. 4.5). Generally, in wet basins ($\zeta < 0.9$), SMM tends to rise with the onset of summer and attains its peak during late summer (July-August) or early autumn (September-October). The data showed minimum memory during winter and early spring (December-April), while dry basins ($\zeta > 0.9$) displayed the smallest memory in late spring (May) and fluctuated during the summer months before starting to rise in autumn and peaking in winter. In contrasting to wet basins, dry basins exhibit the highest memory in winter months. The observed seasonal cycles of memory are consistent with Orth and Seneviratne (2012a). Orth and Seneviratne (2012a) analysed soil moisture memory of four basins from Germany and Switzerland (humid) and one basin from Italy (Mediterranean climatic region). Those humid basins also showed maximum memory in late summer and minimum in spring. Opposite cycles were reported for the Italian basin.

Koster and Suarez (2001), Seneviratne and Koster (2012) and Orth and Seneviratne (2013) mentioned at least six different controls of SMM. However, this study only analysed the influence of precipitation variance on the SMM timescale. Monthly mean standard deviation of precipitation (calculated from 45 years climatology, 1956-2000) and mean SMM timescale (calculated from 21 wet basins and 5 dry basins) are presented in Fig. 4.6. Figure 4.6(a) displays a specific pattern of SMM timescales in wet or dry basins compared with two adjacent months' precipitation variability.

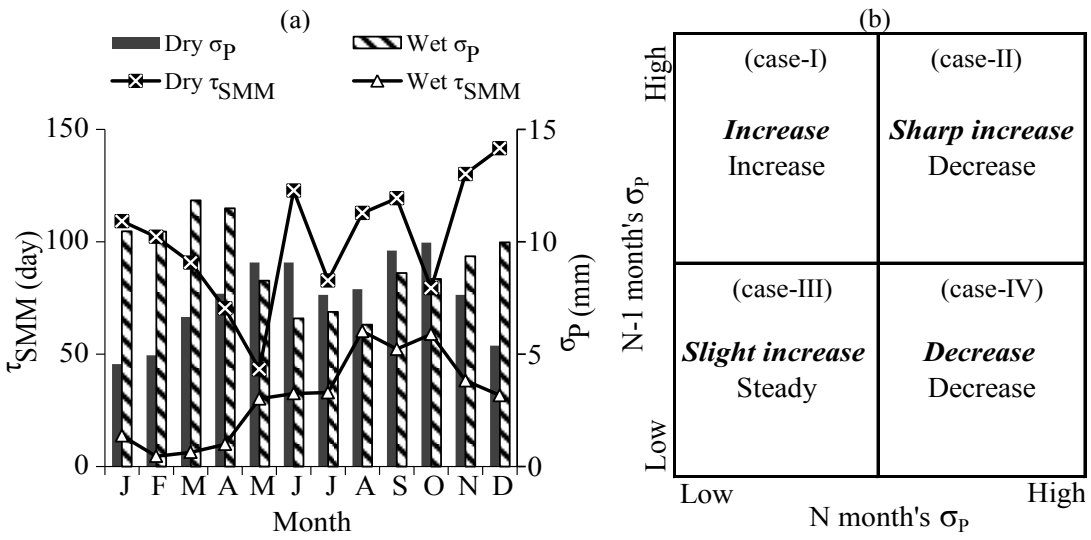

Figure 4.6: (a) Monthly mean variability of precipitation and SMM timescales; (b) conceptual sketch of the effects of two adjacent months ($N-1$ and N month) precipitation variability on the later months (N month) SMM timescales; bold Italic face font indicates the effects in dry basins ($\zeta > 0.9$); regular font represents the effects in wet basins ($\zeta < 0.9$).

Figure 4.6(b) shows a conceptual sketch of the effects of two adjacent months ($N-1$ and N month) precipitation variability on the later month's (N month) SMM timescale. The SMM timescale of month N tends to increase when high precipitation variability in month $N-1$ is followed by a low precipitation variability in month N (case-I, Fig. 4.6(b)). The effects are similar for both wet and dry basins. Figure 4.6(a) shows the increase of memory due to the case-I scenario in the month of May, August and October for wet basins and November, December for dry basins. A decrease of July month's memory in dry basins is the only exception where other controls of SMM could be superior. In contrast, low precipitation variability in month $N-1$ followed by a high precipitation variability in month N results in the decrease of memory in the month N (case-IV, Fig. 4.6(b)). The effects are visible in the memory of September, November and December for wet basins and February, March, April, May and October for dry basins (Fig. 4.6(a)). Memory responds differently in wet or dry basins when high precipitation variability in month $N-1$ is followed by high precipitation variability in month N (case-II, Fig. 4.6(b)) or low precipitation variability in month $N-1$ is followed by low precipitation variability in month N (case-III, Fig. 4.6(b)). Under the former condition (case-II), memory increases sharply in dry basins during June, and decreases in wet basins during February. The latter condition (case-III) causes a slight increase in memory in dry basins (August) but the memory remains steady in wet basins (July).

4.3.4 SMM timescale and aridity index

The analysed basins' aridity index ranges from 0.29 to 1.54. Figure 4.7 shows the annual aridity index (interpolated from 400 MOPEX basins information) map over USA. Analysis

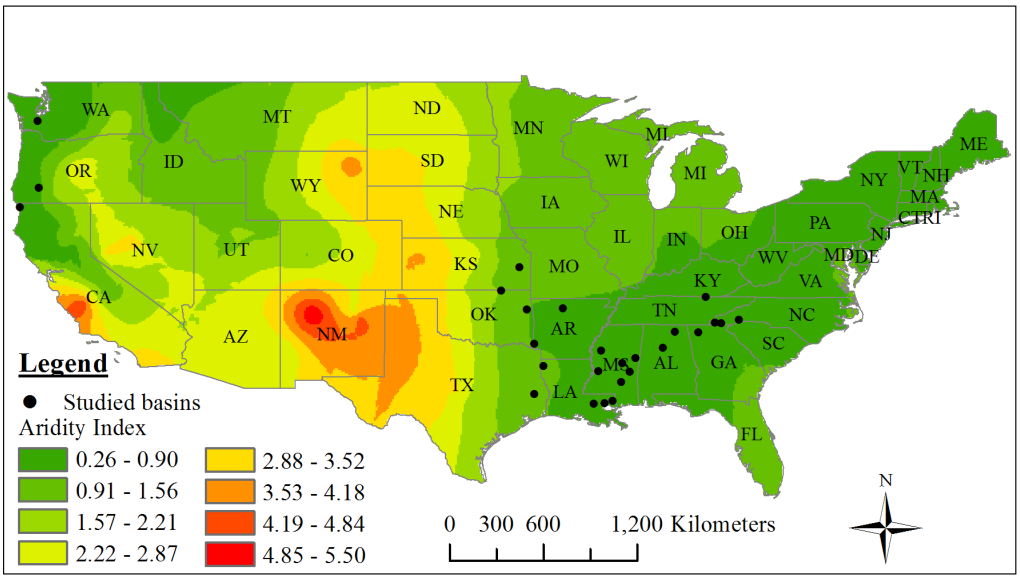

Figure 4.7: Annual aridity index map over USA.

reveals that the dryer basins appear to have larger memories compared with the wetter basins.

The overall relationship between the SMM timescale and aridity index is exponential (Fig. 4.8) with an R^2 value of 0.9. The relationship between the basin aridity index and SMM timescale can be expressed by Eq. (4.3).

$$\tau_{SMM} = 24.76\left(e^{1.25\zeta} - 1\right) \tag{4.3}$$

where τ_{SMM} is the SMM timescale in days and ζ is the basin aridity index.

Despite the dissimilarity in approaches to memory calculation, this result is consistent with the available literature (Liu and Avissar, 1999; Entin et al., 2000; Wu and Dickinson, 2004; Seneviratne et al., 2006a). These literatures suggest that stronger memories exist in the dryer regions. Orth (2013) in another study revealed that there was a tendency to memory increment to a lesser extent compared to the increase of catchment dryness index from the data from 13 catchments of Switzerland. The findings of the present study partly agree with the conclusions reported by Koster and Suarez (2001), Seneviratne and Koster (2012) and Orth and Seneviratne (2012a). These studies argued that both extremes of dry and wet conditions create larger memories, and offered two separate reasons for this phenomenon. Koster and Suarez (2001) reported high values of 31-day-lagged autocorrelation for the first day of July in desert areas and parts of the wet tropics. In contrast, this study only supports the existence of extended memories in the drier regions. However, prolongation of memory under both the two extreme conditions might be true when two adjacent months forcing variability remain close. Orth and Seneviratne (2012a) pointed out that the influence of evapotranspiration and runoff is limited in humid regions. Therefore, the influence of precipitation on combined forcing variability is prominent. The wet basins might have relatively longer memories in a few months but that may be offset by the other months, due

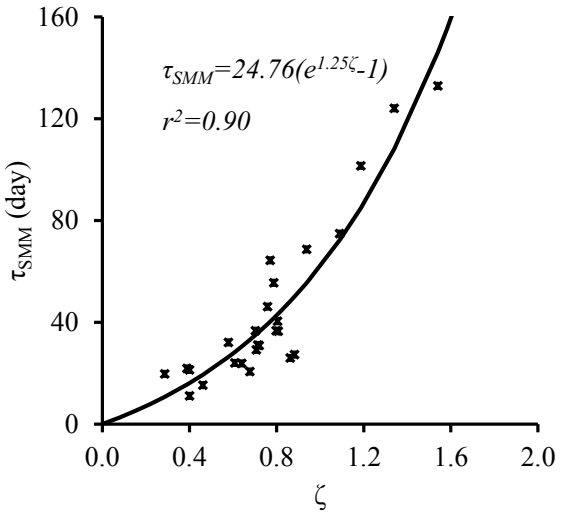

Figure 4.8: Relationship between SMM timescale and aridity index.

to the influence of higher precipitation variability over a combined forcing term, and thus, the overall memory remains short.

4.4 Conclusions

This study investigates the soil moisture memory behaviour of 26 river basins across USA using XAJ model simulated soil moisture data. Since the simulated soil moisture data could not be validated against the observed ones due to the scarcity of such data in the studied river basins, the XAJ model's capability to simulate soil moisture data was validated against observed soil moisture data in one basin in Illinois, USA. Moreover, one month-lagged memory calculated from the XAJ model simulated soil moisture was validated against those calculated from the observed data set. Soil moisture validation suggests that the XAJ model can represent soil moisture anomalies (seasonal cycles) well. Therefore, the XAJ model simulated soil moisture could be used for SMM studies as it analyses the anomaly dissipation timescale. The applicability of the approach using SMM timescale estimation and the XAJ model simulated soil moisture data is further justified with a high agreement between the calculated memories from the observed and simulated data sets. Results from 26 river basins suggest that the SMM timescale increases exponentially with the aridity index of the basin. A basin aridity index range of 0.29 to 1.54 corresponds to the memory timescale ranges from 11 to 133 days. The wet basins' ($\zeta < 0.9$) showed an average memory of 31 days, whereas the dry basins' ($\zeta > 0.9$) average memory is about 100 days. The memories showed distinct seasonality in dry and wet basins. In dry basins, memory is maximum in winter and minimum in late spring. In contrary, in wet basins, memory is maximum in late summer or early autumn and minimum in winter or early spring. The SMM timescale estimation based on the significance of autocorrelations is proved to be consistent and useful to understand the pattern and seasonality over seasons or dryness of the basin. It is to be noted that the relationship between SMM timescale and aridity would be robust, even if a

different threshold value of autocorrelation coefficient is set to find the significance of autocorrelations. This study did not consider the altitude or land cover of the analysed basins. Therefore, it is important to know how it is affected by altitude or land cover. Accepting the limitations (i.e. using model simulated soil moisture data, uncertainty associated with observed datasets) and assumptions (i.e. ignoring snow impacts while selecting basins and considering statistical significance to measure SMM timescale), this result is useful to the basin-scale hydrological perspectives and even for global scale atmospheric weather prediction, particularly in the case of drought and flood prediction. Moreover, the calculated SMM timescale is easily understandable, comparable and provides important possibilities to roughly predict the SMM of a particular region from widely available observed precipitation and potential evapotranspiration dataset. This could produce some knowledge of SMM under no knowledge conditions.

Chapter 5

Variability of XAJ model spin-up time against initial conditions

5.1 Introduction

Hydrological models constitute an important tool for managing water resources. They can be used as a support instrument for understanding physical processes or prediction purposes. A hydrological model can serve to predict a risk of flooding, indicate the susceptible areas and timing of inundation, and be useful in preparing for evacuation in advance. Likewise, a prediction of future floods and their magnitudes could assist with planning of protective measures. A hydrological model could also be used to assess climate change impacts on water resources. However, sound hydrological prediction requires both access to quality hydrological data and the application of suitable modelling techniques.

Hydrological models are unique and their accuracy could differ greatly from model to model due to differences in model structure (i.e. different field capacities), input data sets and parameterizations. Even a single model could produce diverse outputs and achieve dissimilar accuracies due to variations in calibrations. A number of literatures discuss the effect of model initial condition to its outputs (Berthet et al., 2009; Castillo et al., 2003; Goodrich et al., 1994; Minet et al., 2011; Nikolopoulos et al., 2011; Senarath et al., 2000; Zhang et al., 2011). These studies highlighted the complex interaction among soil moisture initial conditions, climatic factors and soil properties.

When a model is calibrated with a different initial state compared to the target basin's long-term climatology, the model undertakes a period of spin-up during which its internal stores (i.e. soil moisture) adjust from the initial conditions to an equilibrium state (Yang et al., 1995; de Goncalves et al., 2006). The model output during this adjustment period is highly impacted on by the initial condition, and consequently may show huge drift and not be usable. Literature suggests that the typical spin-up time of the land surface model (LSM) could range from one to several years (de Goncalves et al., 2006; Yang et al., 1995; Chen and Mitchell, 1999; Cosgrove et al., 2003; Rodell et al., 2005). Once the model achieves its equilibrium state, the simulated output usually agrees better with the observations and responds realistically to the inputs (Yang et al., 1995; Cosgrove et al., 2003; Seck et al., 2015). Hence, special attention is required for specifying the model initial conditions. However,

due to the scarcity of long-term records or spatially distributed information specifying the catchment states, the model's initial conditions are usually inferred from limited observations or an initial guess (Ajami et al., 2014b). Rodell et al. (2005) suggests using climatological average states from the same model for the purpose of initialisation in the absence of long-term forcing data.

Several researchers claimed that spin-up time is not only associated with the water holding capacity and its initial values, but also atmospheric forcing and surface conditions (de Goncalves et al., 2006; Yang et al., 1995; Chen and Mitchell, 1999; Cosgrove et al., 2003; Rodell et al., 2005). In a LSM model study Cosgrove et al. (2003) demonstrated that spin-up time varies spatially and is highly correlated with precipitation and temperature. Moreover, they noted that spin-up time is highly influenced by the soil moisture persistence or soil moisture memory (SMM). A low SMM indicates that the soil moisture anomalies are short-lived, dissipate quickly, enabling the model to recover relatively quickly from an undesirable initial state. On the other hand a high SMM that indicates the slowness of anomaly dissipation and would delay the process of model equilibrium. Seck et al. (2015) also documented the link between initial conditions and meteorological conditions. They mentioned the slowness of model equilibrium under dry initial condition due to the longer system memory. Rahman et al. (2015) proposed an easy way to estimate basin scale SMM using aridity index (ratio of annual evaporation over annual precipitation) information only. Since SMM and model spin-up time are interlinked, it is intuitive to have a relationship between aridity index and model spin-up time too.

To minimize the uncertainty associated with the model spin-up process, modellers often implement two main techniques. Firstly, the model is often run repeatedly using a single or multiple years of forcing data until it reaches an equilibrium state and thereafter initialises the model according to this equilibrium state (Ajami et al., 2014a; Wood et al., 1998). However, this repeated model run with single year forcing data might not be sufficient to train the model given the extremes of climatological phenomenon. Moreover, it demands computation time and energy (Ajami et al., 2014a). Secondly, modellers often perform the analysis task by excluding the first few months' (years') model outputs (Lim et al., 2012). The length of this data exclusion (spin-up time) is mostly defined by a guess. However, guessing a spin-up time does have its limitations. Excluding initial model outputs could be a very costly task in developing countries where hydro-climatic data is very scarce. Over-estimating the spin-up period will lead to a loss of important information. Likewise, an underestimation would affect the conclusion by incorporating erroneous initial model outputs. Moreover, guessing spin-up time (if any) for a shorter period, particularly for seasonal or monthly simulation would be very problematic. Therefore, understanding the spin-up behaviour of a model is essential for better calibration and simulation experience.

Despite its importance only a very few studies have examined the spin-up behaviour of land surface (Yang et al., 1995; de Goncalves et al., 2006; Cosgrove et al., 2003; Rodell et al., 2005; Lim et al., 2012) or hydrological models (Seck et al., 2015; Ajami et al., 2014b,a). These studies have been done to examine the model spin-up behaviour under diverse conditions of

climate, vegetation and soil types. Although the conclusions have often been model-specific, they delivered essential guidelines on model initial condition setting, and thus reduced modelling errors. However, all most all these studies have been conducted on the basis of multiple years (mostly 10-year) recursive simulations using only a particular year's input data sets. Recursive runs with a single year input data sets would not be sufficient to train the model with climatological extremes. Moreover, conclusions of these studies have been mainly based on the results of one basin or study site. The present study attempted to overcome these limitations by employing a 10-year recursive runs using three different climatological input data sets under four different initial conditions. This study has been done using the XAJ model (Ren-Jun, 1992) over 22 river basins throughout United States.

The XAJ model is a conceptual hydrological model discussed in section 3.2.4. The XAJ model has been widely employed to simulate runoff generation within a catchment in China's humid and semi-arid regions and other parts of the world (Lu and Li, 2014). Researcher considers spin-up time based on their personal feeling, experience and purpose. Lin et al. (2006) considered a spin-up period of 19 days during a four-month streamflow simulation for the Shiguanhe River basin, China. In another study, Lu et al. (2008) considered only 12 hours of spin-up time while forecasting floods at the Huaihe River basin's Wangjiaba sub-basin. It is very difficult to comment on the acceptable duration of the XAJ model spin-up time, as it is mainly controlled by the purpose, scope and scale of interest. However, it could be useful to know the spin-up behaviour of the XAJ model under different conditions to judge and decide the spin-up time for improved simulation exercise. Considering this objective, this study investigates the spin-up behaviour of the XAJ model for 22 river basins across the USA.

Firstly, this study examines the model spin-up times for three different climatological input data sets (precipitation and evaporation). Secondly, it analyses the model spin-up times under four initial conditions for each of the input data sets. Thirdly, it assesses the link between the model spin-up time and soil moisture memory. Fourthly, it explores the relationship between the model spin-up time and the basin's aridity index (ratio of annual evaporation over annual precipitation). Finally, it shows an easy way to predict the maximum model spin-up time using the aridity index information only.

5.2 Materials and Methods

5.2.1 Study area

This chapter analyses 22 river basins across the USA. Stream gauge locations of the analysed basins are shown in Fig. 5.1.

The river basins were selected based on prior calibration experience of the XAJ model by Rahman et al. (2015). Rahman et al. (2015) selected these river basins to avoid snow impacts on SMM calculation and mentioned the XAJ models capability to simulate river discharge with good accuracies. This study selected the same river basins or basins located within the

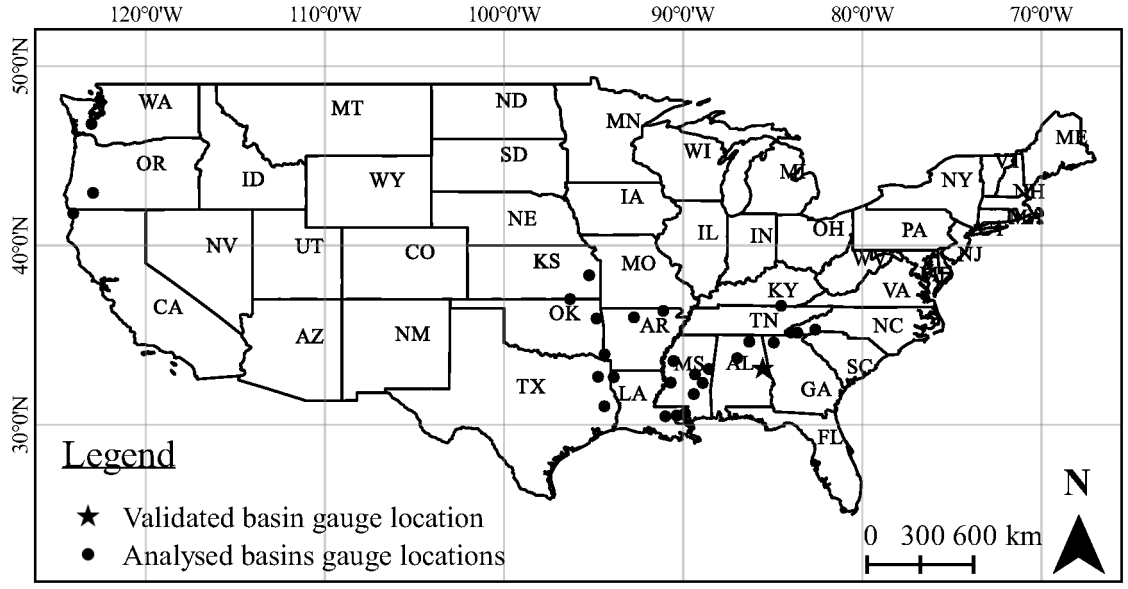

Figure 5.1: Stream gauge location map of studied river basins over USA mainland.

area analysed by Rahman et al. (2015) intending to reduce calibration efforts and to enable linking between SMM and spin-up time. Moreover, snow process would introduce additional system memory and affect its spin-up behaviour. Analysed basins are located in nearly snow-free areas. Based on 30 year climate normal (1981-2010) released by NOAA's National Climatic Data Centre (NCDC) (available at http://www.ncdc.noaa.gov/oa/climate/normals/usnormals.html; accessed on November 13, 2013), the basins have less than 7 snow-days (a snow-day is a day that records at least 2.5mm snow/day) and receive less than 200mm of total new snow per year. A summary of the analysed basins' physical and hydro-climatic characteristics is presented in Table 5.1.

5.2.2 Data

The basin scale daily precipitation, P (daily mean areal precipitation calculated from ground based gauge precipitation), potential evaporation, PE (developed from NOAA Evaporation Atlas), and streamflow, Q data (developed from USGS hydro-climatic data) were obtained from the U.S. Model Parameter Estimation Project (MOPEX) data sets (Schaake et al., 2006). These are freely available at ftp://hydrology.nws.noaa.gov/ (accessed on October 19, 2013).

5.2.3 XAJ model parameters, calibration and validation

The XAJ model calibration for this study has been carried out with the aid of a web-based application (available at http://lmj.nagaokaut.ac.jp/~khin/; last accessed on January 13, 2015) (Khin et al., 2015). This web platform not only allows the user to calibrate the XAJ

Table 5.1: Studied MOPEX basins, locations and basic characteristics. Dry basins (aridity index >0.9) are marked with bold face, Italic font style.

MOPEX ID	Location			Ave. P (mm/year)	Ave. PE (mm/year)	Ave. snow-days (day/year)	Ave. total new snow (mm/year)	Ave. soil moisture saturation (%)
	Long.	Lat.	State					
11532500	-124.05	41.79	CA	2687	740	0.00	0	82
12027500	-123.03	46.78	WA	1599	579	3.00	127	75
03550000	-83.98	35.14	NC	1846	771	3.90	193	75
03504000	-83.62	35.13	NC	1893	762	3.90	193	90
03410500	-84.53	36.63	TN	1389	817	6.20	160	74
02387500	-84.94	34.58	GA	1480	901	0.70	18	73
03574500	-86.31	34.62	AL	1467	941	0.80	41	74
14308000	-122.95	42.93	OR	1347	805	2.20	76	62
07378500	-90.99	30.46	LA-MS	1594	1077	0.60	23	63
07375500	-90.36	30.51	LA-MS	1633	1074	0.60	23	64
02492000	-89.90	30.63	LA-MS	1583	1071	0.60	23	47
02456500	-86.98	33.71	AL	1425	982	0.80	41	66
02414500*	-85.56	33.12	AL	1370	975	0.80	41	65
02472000	-89.41	31.71	MS	1492	1060	0.60	23	64
02448000	-88.56	33.10	MS	1421	1057	0.60	23	72
07290000	-90.70	32.35	MS	1435	1073	0.60	23	57
07056000	-92.75	35.98	AR	1180	916	3.80	132	68
07288500	-90.54	33.55	MS	1381	1112	0.60	23	62
07340000	-94.39	33.92	OK	1329	1156	5.60	198	70
07072000	***-91.11***	***36.35***	***AR***	***1114***	***964***	***3.80***	***132***	***62***
07348000	***-93.88***	***32.65***	***LA***	***1173***	***1223***	***0.10***	***0***	***47***
07346050	***-94.75***	***32.67***	***TX***	***1128***	***1246***	***1.3***	***4***	***53***
06914000	***-95.25***	***38.33***	***KS***	***957***	***1206***	***10.0***	***373***	***61***

*Indicates the validated river basin at Illinois State, USA.

model in a user friendly environment, but also provides: firstly, helpful calibration support by suggesting parameter settings developed by Li and Lu (2014); and secondly, hydrograph visualisation and calculating Nash-Sutcliffe ($NASH$) efficiency (Nash and Sutcliffe, 1970). $NASH$ efficiency was calculated based on Eq. 3.1. Inputs to the XAJ model are areal mean precipitation and potential evaporation. Input data sets throughout this chapter indicate time series of daily precipitation and potential evaporation. A list of XAJ model parameters and their calibrated ranges are presented in Table 5.2.

5.2.4 Recursive simulation design

To detect the spin-up trends, one year input data from 1st January to 31st December was repeated in a yearly cycle for 10-year. Similar recursive experiment was done in several model spin-up studies (Yang et al., 1995; Cosgrove et al., 2003; Seck et al., 2015; Ajami et al., 2014b). This yearly recursive simulation removes inter-annual climate variability and links any model adjustment processes to the equilibrium state of its internal stores (i.e. soil moisture) from an initial anomaly directly to the spin-up processes. However, this single-year recursive simulation may not be able to represent an accurate climatology, and may or may

Table 5.2: Range of calibrated parameters in the Xinanjiang model.

Parameter	Physical meaning	Parameter value
C_p	Ratio of measured precipitation to actual precipitation	0.8-1.2
C_{ep}	Ratio of potential evapotranspiration to pan evaporation	0-2.0
b	Exponent of the tension water capacity curve	0.1-0.3
imp	Ratio of the impervious to the total area of the basin	0-0.005
WUM	Water capacity in the upper soil layer (mm)	5-20
WLM	Water capacity in the lower soil layer (mm)	60-90
WDM	Water capacity in the deeper soil layer (mm)	10-100
C	Coefficient of deep evapotranspiration	0.1-0.3
SM	Areal mean free water capacity of the surface soil layer (mm)	1-50
EX	Exponent of the free water capacity curve	0.5-2.5
KI	Outflow coefficient of the free water storage to inter-flow	0-0.7; KI+KG=0.7
KG	Outflow coefficient of the free water storage to groundwater	0-0.7; KI+KG=0.7
c_s	Recession constant for channel routing	0.5-0.9
c_i	Recession constant for the lower inter-flow storage	0.5-0.9
c_g	Daily recession constant of groundwater storage	0.9835-0.998

not achieve an unnatural equilibrium (Schlosser et al., 2000). To overcome this limitation, recursive simulations were done with three separate input data sets representing mean, 5^{th} and 95^{th} percentile climatology.

Preparation of input files

Streamflow for each basin was simulated with three separate input files, created with a single year data that is close to: i) 5^{th} percentile, ii) mean, and iii) 95^{th} percentile climatology. To maintain consistency among the simulations, a single parameter set was used to simulate all three input files. However, practically it is very unlikely to achieve good calibration accuracy for different climatology using same parameter set due to the difference in water balance and high parameter sensitivity to precipitation (even some basins would produce negative $NASH$ efficiency). As a solution, we tried to manipulate the input data sets in such a way that can represent different climatology by keeping the same distribution pattern throughout the year. The input to the XAJ model is precipitation and potential evaporation. The potential evaporation climatology does not vary between a 'dry year' and 'wet year'. Therefore, we opted to generate hypothetical precipitation (also streamflow for validation purpose) climatology by manipulating that of mean year. This modification steered to gain relatively good calibration accuracy while still capturing the climatology. The objective of this study is to present the spin-up behaviour under different climatology. The modified intra-year precipitation distribution definitely differs from the actual one. However, we believe that this is still sufficient to fulfill our objective. This experimental design may not be the perfect one, but it is an improvement from the earlier approaches.

Firstly, the mean, 5^{th} percentile and 95^{th} percentile precipitation and streamflow climatology were computed from 52-year observed data sets (1948-1999). Secondly, the year that closely represents the mean year climatology was selected to prepare mean year input file by repeating 1^{st} January to 31^{st} December for 10 years. Thirdly, 5^{th} percentile and 95^{th} per-

Table 5.3: Xinanjiang model soil moisture initial conditions.

Initial condition	Physical meaning
Saturated	100% of the field capacity
Intermediate	50% of the field capacity
Dry	Zero soil moisture
Climatology	Mean climatology initial condition

centile input files were created by manipulating the mean year precipitation and streamflow data based on Eqs. (5.1) and (5.2).

$$P_{x,i} = P_{mean,i} * \frac{P_x}{P_{mean}} \tag{5.1}$$

$$Q_{x,i} = Q_{mean,i} * \frac{Q_x}{Q_{mean}} \tag{5.2}$$

where $P_{mean,i}$ and $P_{x,i}$ are the i^{th} day precipitation for the mean and 5^{th} or 95^{th} percentile year, respectively; $Q_{mean,,i}$ and $Q_{x,i}$ are the i^{th} day streamflow for the mean and 5^{th} or 95^{th} percentile year, respectively; P_x, Q_x, P_{mean} and Q_{mean} are annual precipitation for the 5^{th} or 95^{th} percentile year, annual streamflow for the 5^{th} or 95^{th} percentile year, annual precipitation for the mean year and annual streamflow for the mean year, respectively.

Initial conditions

The XAJ model was run with four soil moisture initial conditions for each of the input climatology. The details of initial conditions are given in Table 5.3.

Model calibration

The XAJ model was firstly calibrated with the mean year input file declaring an initial condition as intermediate. Once it achieves a good agreement between the daily observed and simulated discharge, the same parameter sets were used for the remaining simulations. A total of twelve simulations (4 initial conditions X 3 climatologies for each basin) were conducted for each basin. Average layered soil moisture values, obtained from the output of first simulation (mean year input file with an intermediate initial condition) were considered to be the average climatology of the basin.

5.2.5 Definition of model spin-up time

There are several accepted definitions of model equilibrium or spin-up. Yang et al. (1995) define a complete model equilibrium state as the state at which the "model's state at year $n+1$ is identical to that at year n". However, in practice, it is very difficult to achieve identical sates between two recursive simulations, thus quite a few approaches have been proposed

(Cosgrove et al., 2003). Spin-up can be defined based on the e-folding time (time required to reduce the yearly differences in daily/monthly model output to its 1/e value) (Delworth and Manabe, 1988), halving time (time required to reduce the yearly differences in daily/monthly model output to its half) (Simmonds and Lynch, 1992) or percent cut off-based (PC) time (time required for yearly changes in daily/monthly model output to decrease to a certain threshold; see Cosgrove et al. (2003); de Goncalves et al. (2006)). Of these, PC time has been widely used for detecting the model equilibrium (Yang et al., 1995; de Goncalves et al., 2006; Cosgrove et al., 2003; Ajami et al., 2014b; Lim et al., 2012; Henderson-Sellers et al., 1996; Chen et al., 1997).

In this study, the model equilibrium state has been defined on the basis of PC time. PC time defines the extent of time required for yearly changes in daily model output to decrease to a certain threshold. Generally, the threshold value for the model equilibrium varies from 1 to 0.01% depending on the purpose and scope (Yang et al., 1995; de Goncalves et al., 2006; Cosgrove et al., 2003; Ajami et al., 2014b; Henderson-Sellers et al., 1996). This study detects the equilibrium at 0.01% threshold. The percentage change of daily values of total soil moisture was calculated by Eq. (5.3).

$$PC = \left| \frac{D_{i,n} - D_{i,n+1}}{D_{i,n+1}} \right| * 100 \tag{5.3}$$

where PC, $D_{i,n}$ and $D_{i,n+1}$ are the percentage change, the total soil moisture at day i of year n and $n+1$, respectively.

5.2.6 Reporting of model spin-up time

Every basin produces twelve different spin-up times (4 initial conditions X 3 climatologies for each basin). The analysis relating to the basin aridity index considers the highest spin-up time produced by the initial condition that is closest to the average climatology. The average saturation of the river basins is shown in Table 5.1. The average saturation of 18 out of 22 river basins is close to 50% of their respective field capacity. Therefore, the spin-up time produced with an intermediate initial condition was reported for those basins. The remaining four river basins seem to have average saturation close to their full capacity, and thus spin-up times produced with a saturated initial condition was reported for those basins (1^{st} four basins of Table 5.1).

As discussed earlier, model achieves equilibrium state quickly under less SMM condition. Rahman et al. (2015) argued that soil moisture state loses all the memory once it becomes saturated. In harmony, this study also assumed that the XAJ model will take little or no time to achieve an equilibrium state under highly wet conditions (aridity index approaches zero). Thus, the regression equation presented in this paper which shows the relationship between the spin-up time and aridity index was optimised so that the model's behaviour in arid conditions beyond the examined basins could be better understood.

5.2.7 Calculation of basin aridity index and SMM timescale

The independent aridity index values interpolated from the aridity index values of 400 MOPEX river basins across the USA (discussed in section 4.3.4) were used to discuss the relationship between the model spin-up time and the aridity index. Consistent with Chapter 4, the river basins are grouped as wet ($\zeta < 0.9$) and dry ($\zeta > 0.9$) basins for the simplicity of analysis. The basin average soil moisture memory (SMM) timescale in days was estimated based on Eq. (4.3).

5.3 Results and Discussion

5.3.1 Hydrograph, SMM timescale and aridity index

The daily $NASH$ efficiencies of the analysed basins range from 0.43 to 0.87. The validated daily hydrograph for the mean year simulation is presented in Fig. 5.2.

The validation result suggests that the simulated daily streamflow agrees very well with those of daily observed streamflow. The basin-wise range of $NASH$ efficiency, SMM timescale and aridity index are included in Table 5.4.

5.3.2 XAJ model spin-up time and SMM timescale

Analysed basins' spin-up time ranged from 2 to 655 days. The wet basins ($\zeta < 0.9$) require less time (mean spin-up time 55 days) to be equilibrated compared to the dry basins ($\zeta > 0.9$; mean spin-up time 298 days). Basin-wise model spin-up times produced with an initial condition that is close to the average climatology (intermediate for 18 basins and saturated for 4 basins) are given in Table 5.4.

Average spin-up times of the XAJ model in wet and dry basins for all three input data sets with four initial conditions are shown in Fig. 5.3.

Spin-up time tends to increase with the dryness of initial conditions in all basins for both mean and 95^{th} percentile input data sets. In contrast, wet and dry basins respond differently when the XAJ model is run with the 5^{th} percentile input data sets. XAJ model spin-up time increases with dryness of the initial conditions for wet basins while calibrated with 5^{th} percentile input data sets. In contrast, the XAJ model takes less time to achieve equilibrium for 5^{th} percentile input data sets with dry initial condition. In wet basins, saturated initial condition requires less time to reach equilibrium. Similarly, Seck et al. (2015) also suggests that spin-up for dry initial condition is slower than that of wet initial conditions. However, available literature does not clarify which initial condition should facilitate equilibrium condition in least time.

Theoretically we believe that any initial condition that is close to the average climatology should lead equilibrium quickly. In wet basins (average climatology around 70% of the field capacity), dry initial condition creates maximum anomalies, and thus would take longest time to reach equilibrium. Similarly, in dry basins (average climatology slightly over 50% of

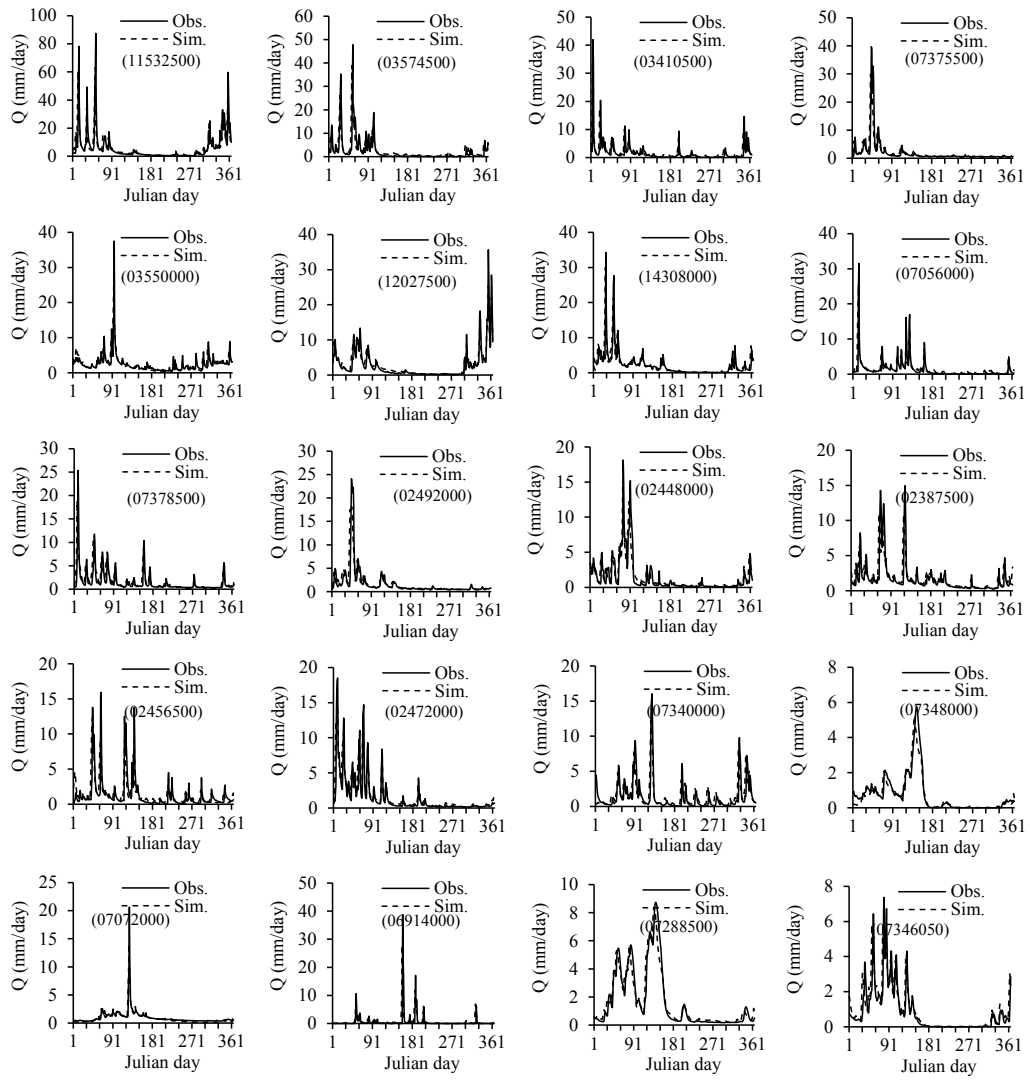

Figure 5.2: Validated daily hydrograph of 20 analyzed river basins (calibrated with mean year input data sets). MOPEX ID is presented in parenthesis.

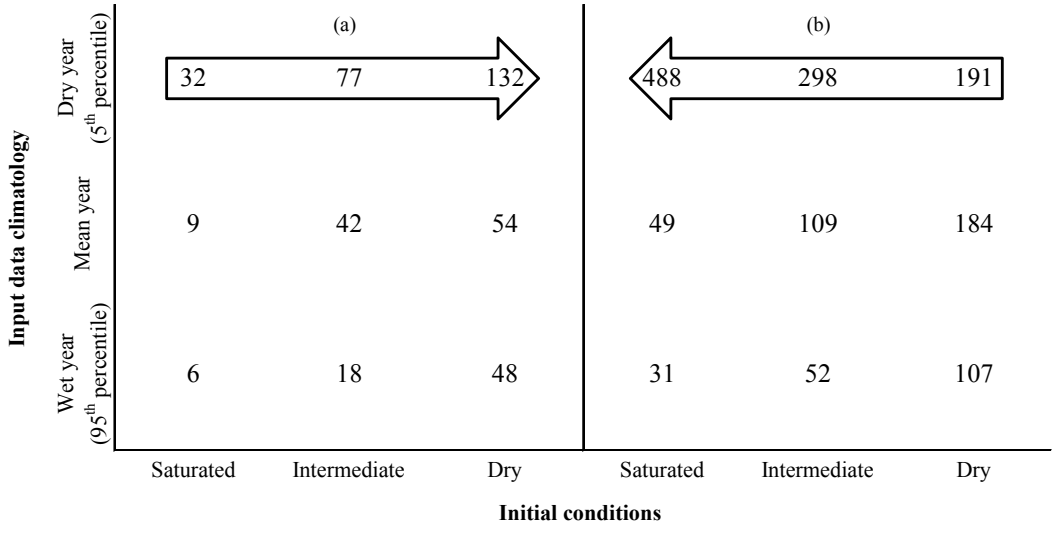

Figure 5.3: Average XAJ model spin-up time (in days) produced with different initial conditions and input data sets; (A) wet basins ($\zeta < 0.9$) (B) dry basins ($\zeta > 0.9$).

Table 5.4: Summary of the XAJ model spin-up time analysis.

MOPEX ID	Area (sq.km)	Daily NASH	Aridity index (ζ)	τ_{SMM} (day)	τ_{Xsp}(day)
11532500	1577	0.70-0.75	0.29	11	7
12027500	2318	0.71-0.73	0.39	16	3
03550000	269	0.58-0.72	0.40	16	2
03504000	135	0.70-0.77	0.40	16	9
03410500	2471	0.59-0.70	0.58	26	14
02387500	4144	0.67-0.72	0.61	28	18
03574500	829	0.72-0.82	0.64	30	23
14308000	1163	0.77-0.84	0.68	33	27
07378500	3315	0.43-0.61	0.70	35	43
07375500	1673	0.75-0.85	0.71	35	40
02492000	3142	0.77-0.82	0.71	36	24
02456500	2292	0.65-0.70	0.72	36	43
02414500*	2696	0.79	0.73	37	55
02472000	1924	0.48-0.79	0.76	39	40
02448000	1989	0.43-0.83	0.80	42	73
07290000	7283	0.54-0.61	0.80	43	131
07056000	2147	0.64-0.81	0.81	43	65
07288500	1987	0.75-0.84	0.86	48	68
07340000	6895	0.58-0.61	0.88	50	342
07072000	***1134***	***0.61-0.87***	***0.90***	***52***	***192***
07348000	***8125***	***0.46-0.71***	***1.09***	***72***	***134***
07346050	***383***	***0.55-0.74***	***1.15***	***79***	***211***
06914000	***865***	***0.43-0.69***	***1.34***	***108***	***655***

*Indicates the validated river basin at Illinois State, USA.

the field capacity), dry initial condition creates the maximum anomalies too, and thus might equilibrate lately. In both wet and dry basins, we expected that intermediate initial condition (50% of the field capacity) would achieve equilibrium quickly. However, present study reveals that the XAJ model consistently tends to achieve equilibrium shortly under saturated initial condition for all the basins (except for dry basins simulated with 5^{th} percentile input data sets) irrespective to their average climatology. This might be a XAJ model dependent phenomena and the XAJ model would always approach equilibrium quickly under a saturated initial condition. Moreover, the XAJ model seems to behave differently under dry-dry (dry basins simulation with dry climatologies) conditions. The exceptional behaviour of the XAJ model spin-up time while simulating with 5^{th} percentile climatology could be better understood by analysing more dry basins. Unfortunately, the XAJ model is reported to work better under humid and semi-humid areas (Ren-Jun, 1992; Lu and Li, 2014; Li and Lu, 2014), thus such investigation under dry-dry conditions would be challenging. Nevertheless, the outcomes of this present study would be very essential for the XAJ model applications for most of the areas.

Among the input data sets, 95^{th} percentile exhibits the least spin-up time requirement for any initial condition. Moreover, saturated initial condition with 95^{th} percentile input data

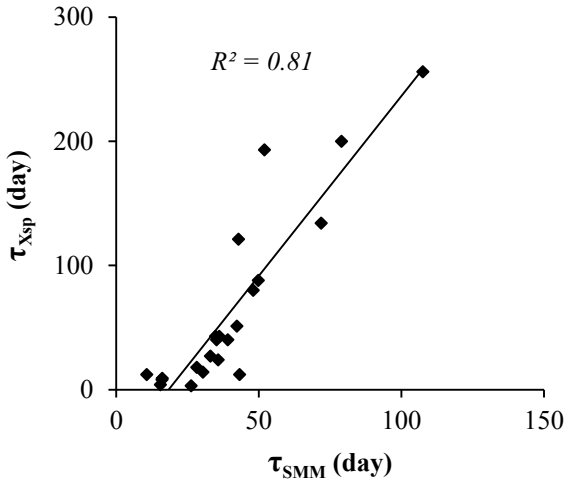

Figure 5.4: Relationship between basin-wise soil moisture memory and the XAJ model spin-up time.

sets displayed the minimum time requirement to be equilibrated. Fig. 5.3 reveals that the XAJ model spin-up time tends to increase with both the dryness of initial condition and the climatology of input data sets. Therefore, the findings of the present study indicate that for wet basins, a saturated initial condition could save the XAJ model spin-up time regardless the climatology of input data sets. While, for dry basins, drier initial condition could be wise in case the input data sets represent the drier climatology. Model simulation with climatology initial condition also disclosed substantial time requirement for the XAJ model equilibrium. This implies that model initialisation based on observed or model derived climatological mean may not always be sufficient to avoid the spin-up error. A precise model initialisation might also require spin-up time to be considered for the subsequent analysis. Estimated model spin-up time (with climatology initial condition) exhibits a strong agreement with the basin average SMM timescales (calculated from independent data sets) with an R^2 of 0.81 (Fig. 5.4). This is consistent with Cosgrove et al.'s (2003) argument about the association between model equilibrium and soil moisture persistence. This high R^2 value not only indicates the influence of soil moisture anomaly dissipation speed on the model spin-up time, but also justifies the use of recursive simulation to detect the XAJ model's spin-up behaviour. Even though a single year forcing data was used to run the model in a recursive way, it can still sufficiently capture the basin's characteristics.

5.3.3 Predictability of XAJ model spin-up time from basin aridity index

Literatures (Seck et al., 2015; Ajami et al., 2014b) suggest that spin-up time for integrated hydrological model is much longer than that typically reported for LSMs. Comparing spin-up time of models from different types would be a very tricky task. Soil moisture persistence is stronger as compared with those of meteorological fluxes (Koster and Suarez, 2001). Similarly, soil moisture persistence in the deeper layer is much stronger than that of surface layer

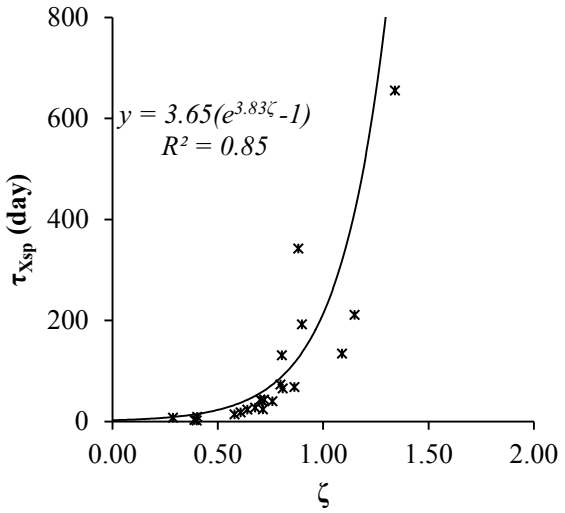

Figure 5.5: Relationship between basin aridity index and the XAJ model spin-up time.

(Vinnikov et al., 1996; Entin et al., 2000; Wu and Dickinson, 2004). Therefore, the model spin-up study considering equilibrium for different state variable (sensible/latent heat flux, total soil moisture, root zone soil moisture, depth of water table, discharge, ground water storage etc.) would provide different results. However, model spin-up behaviour (how it approaches towards equilibrium under different circumstances) could be compared quite easily. The XAJ model's spin-up behaviour seems to be consistent with those of LSMs. Noah LSM spin-up study (Lim et al., 2012) on Korean Data Assimilation System argued that dry land areas takes more than 40 months for spin-up compared to the wet areas. Similarly, Rodell et al. (2005) claimed that Mosaic (Koster and Suarez, 1992) LSM shows less spin-up time in humid regions compared to arid regions. Moreover, Cosgrove et al. (2003) demonstrated a strong spatial variation and correlation of spin-up time with precipitation and temperature.

Computed basin-wise XAJ model spin-up time (mostly with an intermediate initial condition) reveals an exponential relationship with basin aridity index (calculated from independent data sets) with an R^2 value of 0.85 (Fig. 5.5).

The relationship between the basin aridity index and model spin-up time can be expressed by Eq. (5.4).

$$\tau_{Xsp} = 3.65\left(e^{3.83\zeta} - 1\right) \tag{5.4}$$

where τ_{Xsp} is the XAJ model spin-up time in days and ζ is the basin aridity index.

This relationship could be useful for roughly estimating the XAJ model spin-up time when no information about the soil moisture climatology is available. Declaring an intermediate initial condition is easy and straightforward compared to setting a climatology initial condition. However, it should be noted that this relationship is based on the daily scale model simulation only, thus the XAJ model spin-up time for shorter or longer scale might be different.

The equation was validated against the actual spin-up behaviour of the XAJ model for

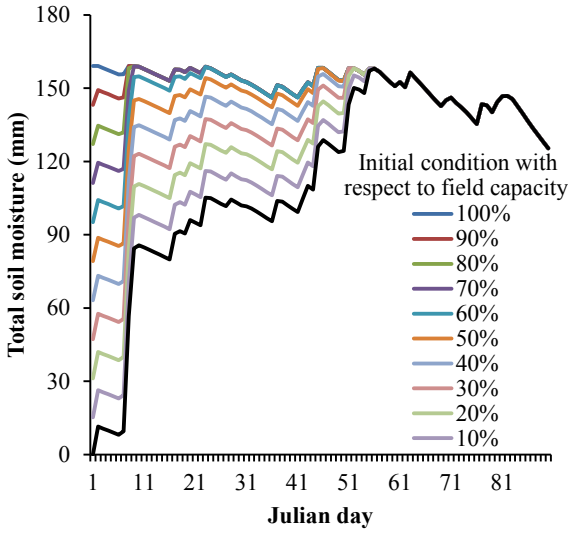

Figure 5.6: Time series plot of total column soil moisture (mm) over the 11 year simulation for validated river basin at Alabama, USA (the Tallapoosa River basin, MOPEX ID # 02414500).

the Tallapoosa River basin, Alabama, USA (MOPEX ID # 02414500; $\zeta = 0.73$, gauge location shown in Fig. 5.1). The equation suggests a maximum spin-up time of 56 days for a basin with an aridity index of 0.73. A recursive simulation with several initial conditions (mean year input data) indicates that the model takes a maximum of 55 days to reach an equilibrium soil moisture state (dry initial condition, see Fig. 5.6).

5.4 Conclusions

Spin-up is the process during which a model adjusts its internal stores to an equilibrium state from an unusual initial state. Model outputs during this adjustment process are highly affected by the initial conditions, and consequently could be unrealistic and misleading. To avoid this problem, modellers often prefer to set the model initial condition as close to the reality and/or exclude the model outputs for the first few months. However, studies suggest that perfect initialisation may not be sufficient for eliminating the risk of erroneous model output. The model adjustment process is not only affected by the initial condition but also by the characteristics of input data sets. Similarly, exclusion of the first few months' model outputs is not an ideal solution. Exclusion of model output guided by a feeling could lead to underestimating or overestimating spin-up time. Therefore, prior information about the model's behaviour under different conditions or preferable initial conditions will improve the detection of spin-up time or reducing spin-up time, respectively. This study investigates the XAJ model's spin-up behaviour using different initial conditions and input data sets (representing separate climatology) for 22 river basins across the USA. The XAJ model shows an increasing trend of spin-up times against both the dryness of input data sets and initial conditions. The responses are identical in wet and dry basins for the mean and 95^{th} percentile input data sets. In contrast, it behaves differently in wet and dry basins

for the 5^{th} percentile input data sets. In wet basins, spin-up times tend to increase with the dryness of initial condition, while dryer initial condition produces less spin-up time in dry basins. Among the input data sets, 95^{th} percentile exhibited least spin-up time requirement regardless of the basin dryness. For all the basins, a 95^{th} percentile input data sets with saturated initial condition showed the minimum time to be equilibrated. Analysis suggests that a saturated initial condition is preferable for mean year or 95^{th} percentile data sets for all the basins. However, it would be wise to utilize saturated and dry initial condition for the dryer input data sets (5^{th} percentile) for wet and dry basins, respectively. Finally, the wet basins require less time for model equilibrium compared to those of dry basins. The spin-up time displays a high correlation with the basin soil moisture memory timescale. Moreover, the XAJ model spin-up timescale exhibits an exponential relationship with basin aridity index. This relationship allows estimating the XAJ model spin-up time using only precipitation and evaporation information only. Estimation of the XAJ model spin-up time could be valuable to reduce uncertainty associated with guessing spin-up time, simply based on feeling or experience. Prior information about model spin-up time would allow us to fully use the information included in short data records under data scarce situation.

Chapter 6

Seasonality of model spin-up time

6.1 Introduction

When a model is calibrated with an unusual initial condition, the model undergoes some adjustment process to reach the normal equilibrium state (Yang et al., 1995; Cosgrove et al., 2003; de Goncalves et al., 2006). Time requires completing this model adjustments or reach its equilibrium condition in its internal stores (i.e. soil moisture) is called as the model spin-up time. The length and behaviour of this spin-up process is a function of chosen initial conditions, parameters that describes the model domain and the model forcing (Seck et al., 2015). The model output during this spin-up time is hugely impacted on by the initial condition, and often unrealistic or misleading. The model outputs after its initial adjustments normally correspond better with the observations and reacts realistically to the inputs (Yang et al., 1995; Cosgrove et al., 2003; Seck et al., 2015). Consequently, it is important to pay particular attention to the model spin-up process, its length and behaviour for the modellers. However, clear information about the length of spin-up time is often missing or model specific and cannot be applicable to all. In practice, modellers tend to reduce this spin-up period or exclude the initial model outputs for improved modelling accuracies mostly guided by a guess. These techniques of reducing spin-up errors hold certain limitations. Therefore, understanding the factors affecting the spin-up process and its behaviour is highly important for modelling communities.

The influence of initial conditions in hydrological models have been studied by several researchers (Goodrich et al., 1994; Senarath et al., 2000; Castillo et al., 2003; Zehe et al., 2005; Berthet et al., 2009; Nikolopoulos et al., 2011; Zhang et al., 2011; Minet et al., 2011). However, these studies were done focusing mainly on event-scale or short-term response. Moreover, these literatures not necessarily quantify the spin-up time or discussed the criteria to specify the equilibrium condition of model state once it finishes the spin-up time (Seck et al., 2015). Recently, few studies have discussed about the spin-up time and behaviour of integrated hydrological model (Ajami et al., 2014b; Seck et al., 2015; Rahman and Lu, 2015). Rahman and Lu (2015) (also chapter 5) suggested an easy way to estimate the maximum spin-up period of the XAJ model under extreme climatology using only basin aridity index information. Estimating maximum model spin-up time could reduce uncertainty under extreme conditions. In contrary, spin-up time of land surface models (LSMs) is well docu-

mented (Yang et al., 1995; Robock et al., 1998; Schlosser et al., 2000; Cosgrove et al., 2003; Rodell et al., 2005; Lim et al., 2012). These studies have been done to examine the model spin-up behaviour under diverse conditions of climate, vegetation and soil types. Reported spin-up time of LSMs varies from models to models and range from one to several years (de Goncalves et al., 2006; Yang et al., 1995; Chen and Mitchell, 1999; Cosgrove et al., 2003; Rodell et al., 2005). Despite the conclusions of these literatures are often model-specific, they provide important insights and guidelines for all modelling communities about the spin-up analysis.

Up-to-date spin-up studies are mostly done on the basis of a recursive model runs through a specific period (typically a single year) where the outputs at the end of one simulation become the initial conditions for the next simulation (Yang et al., 1995). This single year recursive simulations are claimed to be eliminating the inter-annual climatic variability and links any model adjustments from year to year exclusively to the spin-up process (Cosgrove et al., 2003). By repeating the same annual forcing, it applies the same temporal dynamics to the system and enables to distinguish between the effects of persistence in initial conditions and high and low-frequency dynamics that would be added while using multi-year climatology (Seck et al., 2015). This recursive simulation assumes that single year forcing is representative of the observed climatology (Cosgrove et al., 2003). Rahman and Lu (2015) tried to improve the representativity of this single year simulation by analysing the model spin-up behaviour based on simulation results using three different climatological input data sets (mean, 5th and 95th percentile). However, this repeated model runs using a single year climatology may not be sufficient to train the model with all climatological phenomenon, and thus devoid of additional insights. Moreover, this single year recursive model run always starts the simulation from a particular point of year. Since, the spin-up process is strongly associated with the atmospheric forcing and surface conditions (Yang et al., 1995; Chen and Mitchell, 1999; Cosgrove et al., 2003; Rodell et al., 2005; de Goncalves et al., 2006), the spin-up behaviour would be different when the model simulations starts from a different time of year. Keeping the same initial conditions and employing different starting climatology certainly affects the speed of spin-up process. Recently, Rahman et al. (2015) discussed about the seasonality of soil moisture memory (SMM). Therefore, it is intuitive for any model spin-up time to show certain seasonality.

This study attempted to analyse the seasonality of model spin-up time using the Xinanjiang model (XAJ) (Ren-Jun, 1992). The XAJ model is a conceptual hydrological model discussed in section 3.2.4. Unlike existing literature, this study uses multi-year climatology instead of a single year recursive simulation. We believe that use of multi-year climatologies, train the model in a better way and the outcomes are more realistic. Moreover, to detect the seasonality of model spin-up time, we perform series of simulations that starts from different time of year (details are given in section 6.2.4). Using multi-year forcing climatology also requires the model equilibrium condition to be defined differently from the above mentioned spin-up studies (Yang et al., 1995; Cosgrove et al., 2003; Rodell et al., 2005; Lim et al., 2012; Ajami et al., 2014b; Seck et al., 2015; Rahman and Lu, 2015). Generally, in recursive

simulation based spin-up studies, the model equilibrium condition was defined mainly based on the percent cutoff-based time (PC time). PC time is the time requires for the yearly changes in daily/monthly model output to decrease to a pre-defined threshold values (Cosgrove et al., 2003; de Goncalves et al., 2006). In a recursive simulation, the system receives same temporal dynamics for a particular point. This actually allows to detect the progress of adjustment process based on user defined resolution. In contrary, using multi-year forcing employs varying temporal dynamics and the equilibrium state could be different from year to year. Therefore, we performs two simulations; one initialised with completely "dry", and another initialised with completely "saturated".

In XAJ model, the soil moisture is represented in three layers. When the XAJ model is calibrated with two extreme initial conditions ("saturated" and "dry"), the soil moisture stores of each simulation will gradually converge towards a common state of equilibrium, and thus will show correlations. This equilibrium model state can be detected by estimating Mahalanobis Distance (MD) (Mahalanobis, 1930) between the soil moisture states (prognostic variable for this study) of two simulations. MD has been applied in many fields to solve the classification problems, where there are several groups and concerns of affinities between the groups are present (McLachlan, 1999; De Maesschalck et al., 2000). MD has been used to detect the outliers (Martens and Naes, 1992; Leroy and Rousseeuw, 1987), to select the calibration samples from a large set of measurements (Shenk and Westerhaus, 1991), to investigate the representativity between two data sets (Jouan-Rimbaud et al., 1997, 1998; Wilson and Atkinson, 2007) and similarity between two river flow series (Corduas, 2011). MD is useful to measure the divergence or distance between groups in terms of multiple characteristics. MD weights the variables with their covariance, which attributes less weight to strongly correlated variables.

Calibrating XAJ model using multi-year forcing climatologies and "saturated" and "dry" initial conditions, this study investigated the seasonality of spin-up time for 18 river basins across the USA. This study holds at least three major comparative advantages over the existing spin-up literatures. Firstly, it uses multi-year forcing that overcome the limitations contains in single year recursive simulation in the sense of representativeness to the actual phenomenon. Secondly, it detects the model equilibrium state based on MD that is widely acceptable in the presence of co-linearity of datasets. Thirdly, it provides useful insights about the seasonality of model spin-up time that is missing in the available spin-up studies.

6.2 Materials and Methods

6.2.1 Study area

This study analyses 18 river basins across the USA. Stream gauge locations of the analysed river basins are shown in Fig. 6.1.

For the sake of consistency with the previous spin-up analysis discussed in Chapter 5, this chapter opted to select the same river basins. Analysing same river basins allows comparing

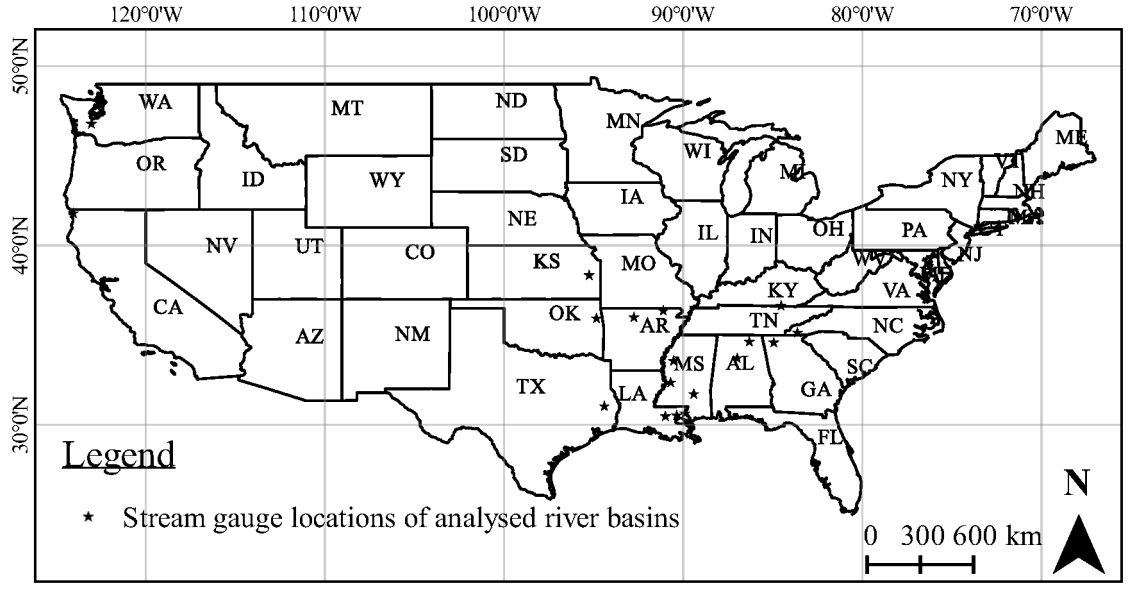

Figure 6.1: Stream gauge location map of studied river basins over USA mainland.

between the model spin-up outcomes derived from two different methodologies. Moreover, it enables to relate the model spin-up time and soil moisture memory. As it was mentioned in Chapter 4 and 5, these river basins have good data records, and their discharge could be simulated by the XAJ model with good accuracies. Analysed basins are located in nearly snow-free areas. Based on 30-year climate normal (1981-2010) released by NOAA's National Climatic Data Centre (available at http://www.ncdc.noaa.gov/oa/climate/normals/usnormals.html; accessed on November 13, 2013), the basins have less than 7 snow-days (a snow-day is a day that records at least 2.5mm snow/day) and receive less than 200mm of total new snow per year. A summary of the analysed basins' physical and hydro-climatic characteristics is presented in Table 6.1.

6.2.2 Data

The basin scale daily precipitation, P (daily mean areal precipitation calculated from ground based gauge precipitation), potential evaporation, PE (developed from NOAA Evaporation Atlas), and streamflow, Q data (developed from USGS hydro-climatic data) were obtained from the U.S. Model Parameter Estimation Project (MOPEX) data sets (Schaake et al., 2006). These are freely available at ftp://hydrology.nws.noaa.gov/ (accessed on October 19, 2013).

6.2.3 XAJ model parameters, calibration and validation

The XAJ model (discussed in section 3.2.4) was calibrated with the aid of a web-based application (available at http://lmj.nagaokaut.ac.jp/~khin/; last accessed on June 09, 2015) (Khin et al., 2015). This web platform not only allows the user to calibrate the XAJ model

Table 6.1: Studied MOPEX basins, locations and basic characteristics. Dry basins (aridity index >0.9) are marked with bold face, Italic font style.

MOPEX ID	Location			Ave. P (mm/year)	Ave. PE (mm/year)	Ave. snow-days (day/year)	Ave. total new snow (mm/year)	Ave. soil moisture saturation (%)
	Long.	Lat.	State					
11532500	-124.05	41.79	CA	2687	740	0.00	0	82
12027500	-123.03	46.78	WA	1599	579	3.00	127	75
03504000	-83.62	35.13	NC	1893	762	3.90	193	90
03410500	-84.53	36.63	TN	1389	817	6.20	160	74
02387500	-84.94	34.58	GA	1480	901	0.70	18	73
03574500	-86.31	34.62	AL	1467	941	0.80	41	74
07378500	-90.99	30.46	LA-MS	1594	1077	0.60	23	63
07375500	-90.36	30.51	LA-MS	1633	1074	0.60	23	64
02492000	-89.90	30.63	LA-MS	1583	1071	0.60	23	47
02456500	-86.98	33.71	AL	1425	982	0.80	41	66
02472000	-89.41	31.71	MS	1492	1060	0.60	23	64
07290000	-90.70	32.35	MS	1435	1073	0.60	23	57
07056000	-92.75	35.98	AR	1180	916	3.80	132	68
07288500	-90.54	33.55	MS	1381	1112	0.60	23	62
07072000	***-91.11***	***36.35***	***AR***	***1114***	***964***	***3.80***	***132***	***62***
07197000	***-94.84***	***35.92***	***OK***	***1162***	***1113***	***5.6***	***198***	***58***
08033500	***-94.40***	***31.02***	***TX***	***1100***	***1308***	***1.30***	***4***	***60***
06914000	***-95.25***	***38.33***	***KS***	***957***	***1206***	***10.0***	***373***	***61***

in a user friendly environment, but also provides: firstly, helpful calibration support by suggesting parameter settings developed by Li and Lu (2014); and secondly, hydrograph visualisation and calculating Nash-Sutcliffe ($NASH$) efficiency (Nash and Sutcliffe, 1970). $NASH$ efficiency was calculated based on Eq. 3.1. Inputs to the XAJ model are areal mean precipitation and potential evaporation. Input data sets throughout this chapter indicate time series of daily precipitation and potential evaporation. A list of XAJ model parameters and their calibrated ranges are presented in Table 6.2.

6.2.4 Simulation design

Unlike recursive simulation with a single year climatologies (as was done in earlier analysis discussed in Chapter 5), this time XAJ model was simulated with full length available observed data sets with two initial conditions (saturated and dry). To detect the seasonality of model spin-up time, this study performs a series of XAJ model simulations with varying simulation start time. The first simulation started from the 1st of January, 1st year and the successive simulations were done with a simulation loop that shifts the simulation starting time by 10-days forward until it completes the loop at 21st December of last year. To maintain the consistency in length of the input data sets among the simulations, the shifted climatologies are placed at the end of the input climatologies, thus total number of data records remains the same for every simulations. Figure 6.2 explains the input data loop that shifts 10-days climatology for a data records from 1st January 1948 to 31st December 1999. In every step, the model was simulated twice using the same input file with two

Table 6.2: Range of calibrated parameters in the Xinanjiang model.

Parameter	Physical meaning	Parameter value
C_p	Ratio of measured precipitation to actual precipitation	0.92-1.1
C_{ep}	Ratio of potential evapotranspiration to pan evaporation	0.9-1.29
b	Exponent of the tension water capacity curve	0.1-0.3
imp	Ratio of the impervious to the total area of the basin	0-0.0001
WUM	Water capacity in the upper soil layer (mm)	20
WLM	Water capacity in the lower soil layer (mm)	50-90
WDM	Water capacity in the deeper soil layer (mm)	20-80
C	Coefficient of deep evapotranspiration	0.1-0.3
SM	Areal mean free water capacity of the surface soil layer (mm)	51-55
EX	Exponent of the free water capacity curve	0.5-1.5
KI	Outflow coefficient of the free water storage to inter-flow	0.1-0.65; KI+KG=0.7
KG	Outflow coefficient of the free water storage to groundwater	0.08-0.6; KI+KG=0.7
c_s	Recession constant for channel routing	0.5-0.88
c_i	Recession constant for the lower inter-flow storage	0.3-0.82
c_g	Daily recession constant of groundwater storage	0.982-0.998

different initial conditions (saturated and dry). Initially, the XAJ model was calibrated with saturated initial condition and thereafter the daily streamflow was validated against those of observed. Later, the calibrated parameter sets were used for rest of the simulations. $NASH$ efficiency reported in Table 6.3 represents only the first simulation.

6.2.5 Definition of model spin-up time

We assume that the model is in equilibrium state when two sets of soil moisture state (from "saturated" and "dry"simulation) become similar. The similarity is measured based on MD. The model was said to be in equilibrium state when the MD score is zero (0). The spin-up time is defined as the number of days require for this MD becomes zero (0). The MD was calculated based on Eq. (6.1).

$$MD\left(\overrightarrow{x_S}, \overrightarrow{y_D}\right) = \sqrt{\left(\overrightarrow{x_S} - \overrightarrow{y_D}\right)^T S^{-1} \left(\overrightarrow{x_S} - \overrightarrow{y_D}\right)} \tag{6.1}$$

where $MD\left(\overrightarrow{x_S}, \overrightarrow{y_D}\right)$ is the MD between the random vectors $\overrightarrow{x_S}$ (states of three soil moisture layers from "saturated" simulation) and $\overrightarrow{y_D}$ (states of three soil moisture layers from "dry" simulation). T and S^{-1} is the matrix transpose and covariance matrix (non-singular) between $\overrightarrow{x_S}$ and $\overrightarrow{y_D}$ respectively.

6.2.6 Calculation of monthly and basin scale model spin-up time and corresponding aridity index

A basin that has 52-year (1948-1999) long observed data requires approximately 1899 simulations. The spin-up time was estimated for every simulations and grouped into months based on the simulation starting time. The monthly spin-up time was then computed by

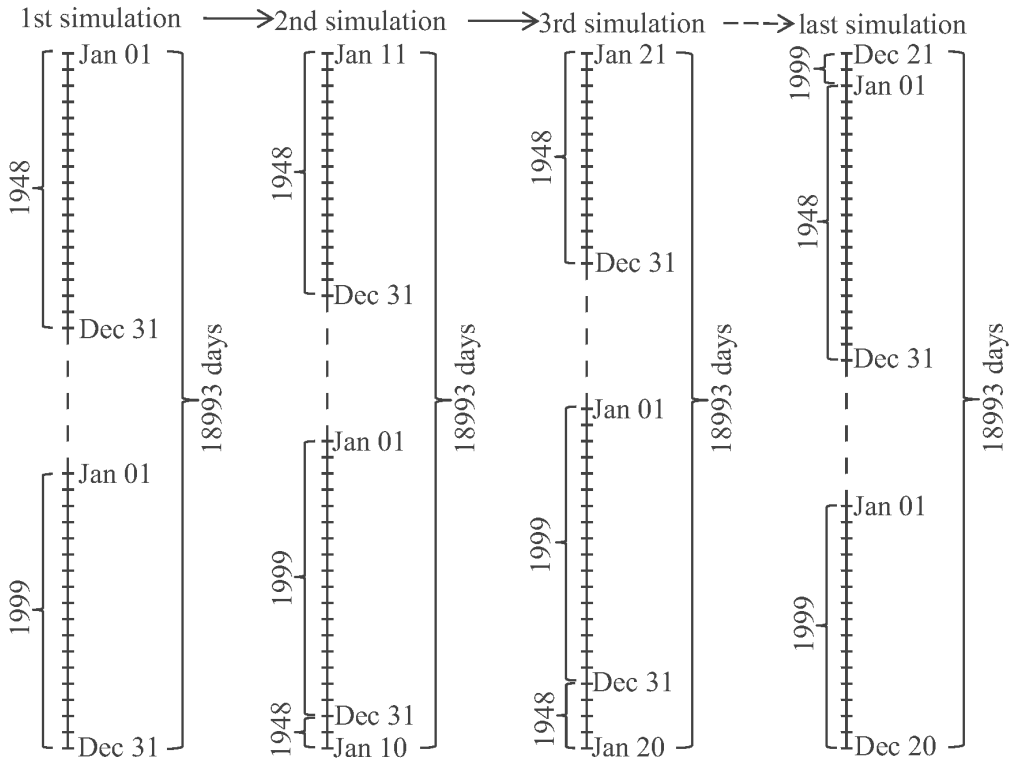

Figure 6.2: Figure explains the input data loop for the XAJ model.

averaging all the spin-up times for the respective months. The basin average spin-up time is the arithmetic mean of all months (January to December).

The monthly spin-up time calculation was followed by the computation of aridity index for the corresponding spin-up period. Aridity index of the corresponding spin-up period for all months of all years were calculated. Finally, monthly means were calculated from the corresponding aridity index of all years.

6.2.7 Calculation of annual aridity index and SMM timescale

The basin annual aridity index was estimated from an independent data sets by interpolating aridity index values of 400 MOPEX river basins across the USA (discussed in section 4.3.4). This annual aridity index has been used to discuss the relationship between the model spin-up time and the aridity index. Consistent with Chapter 4 and 5, the river basins are grouped as wet ($\zeta < 0.9$) and dry ($\zeta > 0.9$) basins for the simplicity of analysis. Basin-wise SMM timescale was computed based on Eq. 4.3.

6.2.8 Optimization of regression equation

On the basis of conclusions regarding the SMM under saturated soil moisture condition (model achieves equilibrium state quickly under less SMM condition; discussed in Chapter 4 and 5), the regression equation that shows the relationship between the spin-up time and

Table 6.3: Summary of the XAJ model spin-up time analysis.

MOPEX ID	Area (sq.km)	Data length (year)	Daily NASH	Annual aridity index (ζ)	τ_{SMM} (day)	τ_{Xsp}(day)		
						Min. (month)	Max. (month)	Mean
11532500	1577	52	0.79	0.29	10.82	18 (Nov)	156 (May)	72.31
12027500	2318	52	0.81	0.39	15.56	25 (Nov)	205 (Apr)	98.11
03504000	135	52	0.81	0.40	16.06	36 (Dec)	147 (Apr)	82.78
03410500	2471	52	0.65	0.58	26.36	51 (Nov)	227 (Mar)	137.02
02387500	4144	52	0.78	0.61	28.32	58 (Dec)	224 (Dec)	135.09
03574500	829	52	0.65	0.64	30.34	48 (Nov)	216 (Apr)	121.34
07378500	3315	51	0.66	0.70	34.64	100 (Nov)	154 (Apr)	184.57
07375500	1673	51	0.67	0.71	35.38	134 (Oct)	295 (Mar)	215.82
02492000	3142	52	0.61	0.71	36.38	69 (Nov)	222 (May)	155.48
02456500	2292	52	0.80	0.72	36.14	74 (Dec)	246 (Apr)	152.73
02472000	1924	52	0.71	0.76	39.26	69 (Nov)	215 (Apr)	144.38
07290000	7283	50	0.67	0.80	42.54	116 (Nov)	258 (Mar)	197.34
07056000	2147	52	0.66	0.81	43.39	82 (Oct)	230 (Mar)	169.19
07288500	1987	42	0.70	0.86	47.79	135 (Oct)	272 (Mar)	205.70
07072000	***1134***	***46***	***0.71***	***0.90***	***51.51***	184 (Oct)	257 (Jun)	***216.09***
07197000	***795***	***52***	***0.65***	***0.94***	***55.42***	145 (Sep)	232 (Jun)	***192.83***
08033500	***9417***	***52***	***0.61***	***1.19***	***84.83***	205 (Nov)	305 (May)	***249.99***
06914000	***865***	***52***	***0.62***	***1.34***	***107.43***	237 (Aug)	290 (Jan)	***272.63***

aridity index was optimised so that the model's behaviour in arid conditions beyond the examined basins could be better understood. It assumes that the XAJ model will take little or no time to achieve an equilibrium state under highly wet conditions (aridity index approaches zero).

6.3 Results and Discussion

6.3.1 NASH efficiency and SMM timescale

The daily $NASH$ efficiencies of the analysed basins suggest that the simulated streamflow has a good agreement with that of observed data sets. Basin-wise NASH efficiency and SMM timescales are reported in Table 6.3.

6.3.2 XAJ model spin-up time and corresponding aridity index

The spin-up time ranged from 1 to 1265 days. The corresponding aridity index ranged from 0.002 to 2.16. All spin-up times are plotted against the corresponding aridity index of that spin-up period in Fig. 6.3. Figure 6.3 reveals that spin-up time is exponentially related with the corresponding aridity index. Although the relationship looks weaker in summer months, all the relationships are statistically significant at 0.0001% (N>2600).

Mean monthly spin-up times disclose a distinct variations in wet ($\zeta < 0.9$) and dry ($\zeta > 0.9$) basins (Fig. 6.4). In wet basins, the XAJ model requires longer time to be

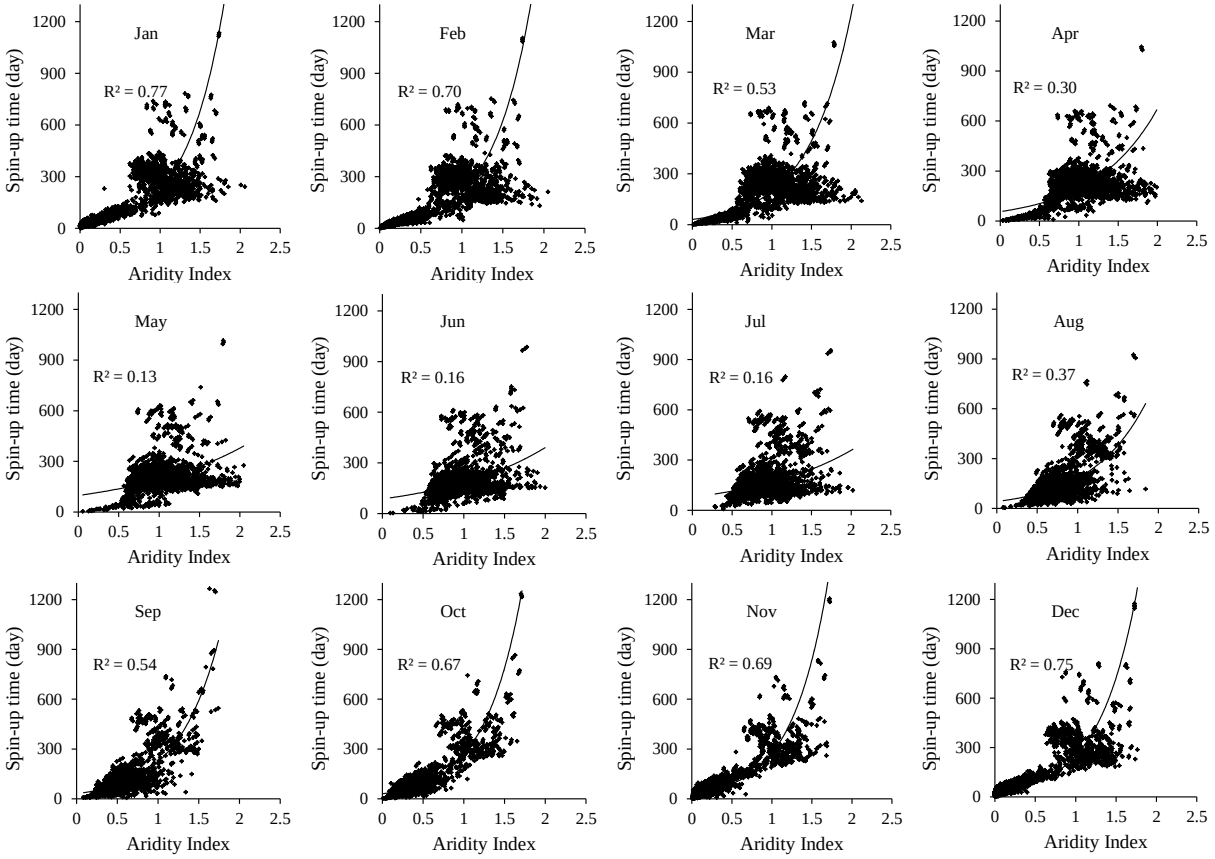

Figure 6.3: XAJ model spin-up times and their corresponding aridity index plotted based on the simulation starting time.

equilibrated when the model simulation starts from the spring months (March-May). While, it achieves equilibrium quickly for late autumn or early winter (October-December). In contrast, in dry basins, the XAJ model equilibrated quickly in early spring (March-April) and autumn (August-October) and it takes longer time for the equilibrium in late spring to summer (May-June). Overall, in all basins spin-up time is highest in spring (March-May) and lowest in late autumn (October-November). This implies that starting simulations from the hydrological year (1 October) could save the spin-up time.

Theoretically, it is believed that the model spin-up time is mainly controlled by the persistence characteristics of soil moisture (as the model equilibrium was detected for soil moisture state). A low SMM implies the soil moisture anomalies are short lived and diminishes quickly to reach in equilibrium condition. The shorter the memory the shorter the spin-up period. Basin-wise SMM timescale and the model spin-up time shows strong correspondence with a $R^2 = 0.79$ (Fig. 6.5). Analysis indicates that model spin-up times are 3-7 times longer than SMM timescale. Yang et al. (1995) showed that the spin-up time of PILPS (Project for Intercomparison of Land Surface Parameterization Scheme) experiment is three times larger than the e-folding time (SMM) at 0.1% PC threshold. In another LSM

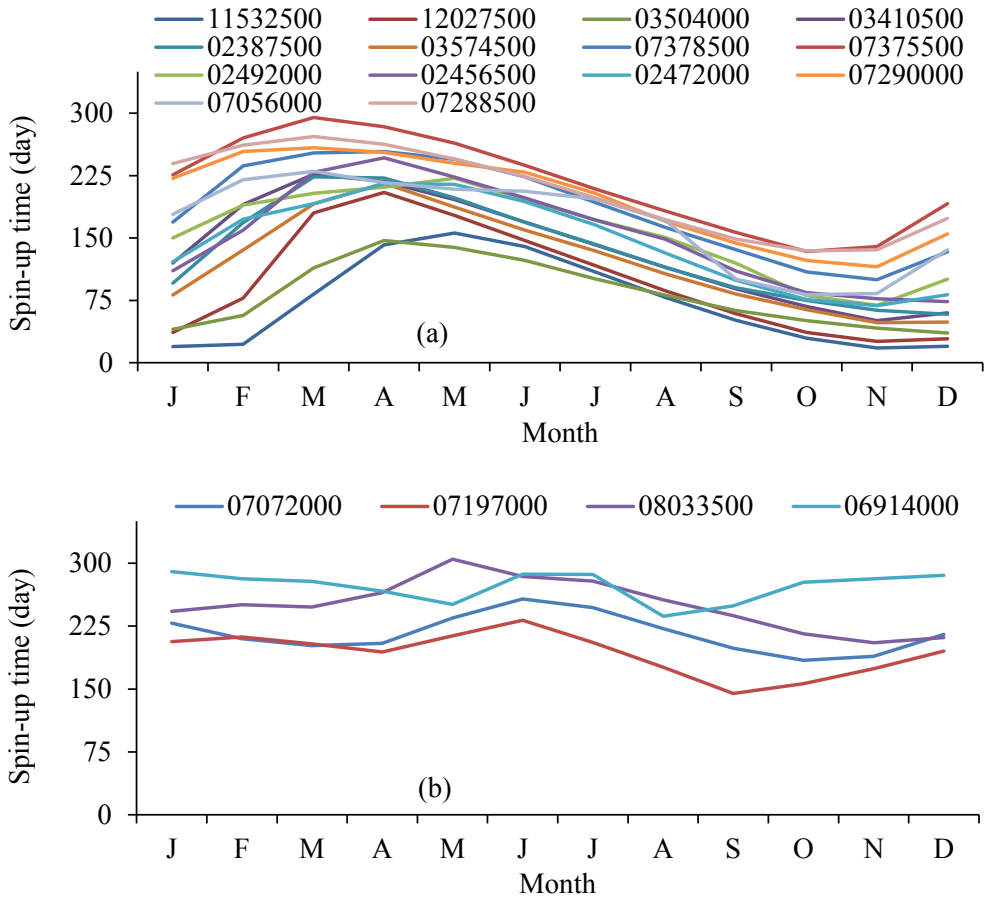

Figure 6.4: Basin-wise mean monthly spin-up times for (a) wet basins ($\zeta < 0.9$) (b) dry basins ($\zeta > 0.9$).

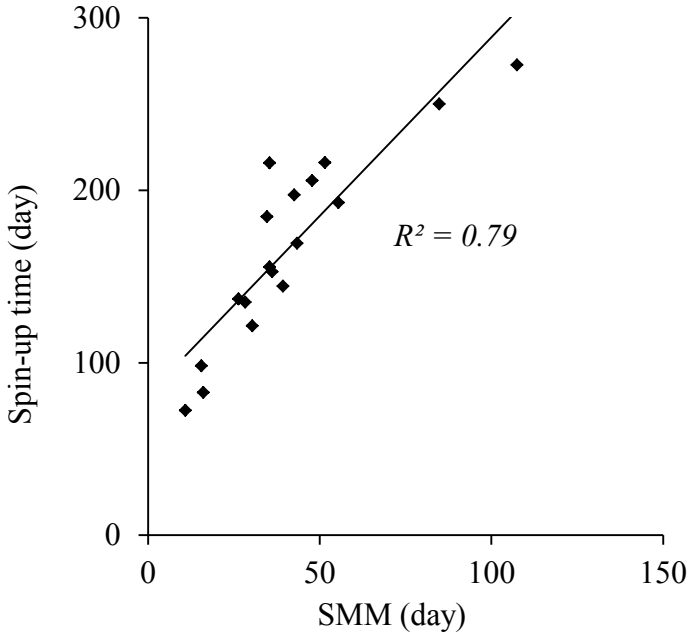

Figure 6.5: SMM timescale and the model spin-up time.

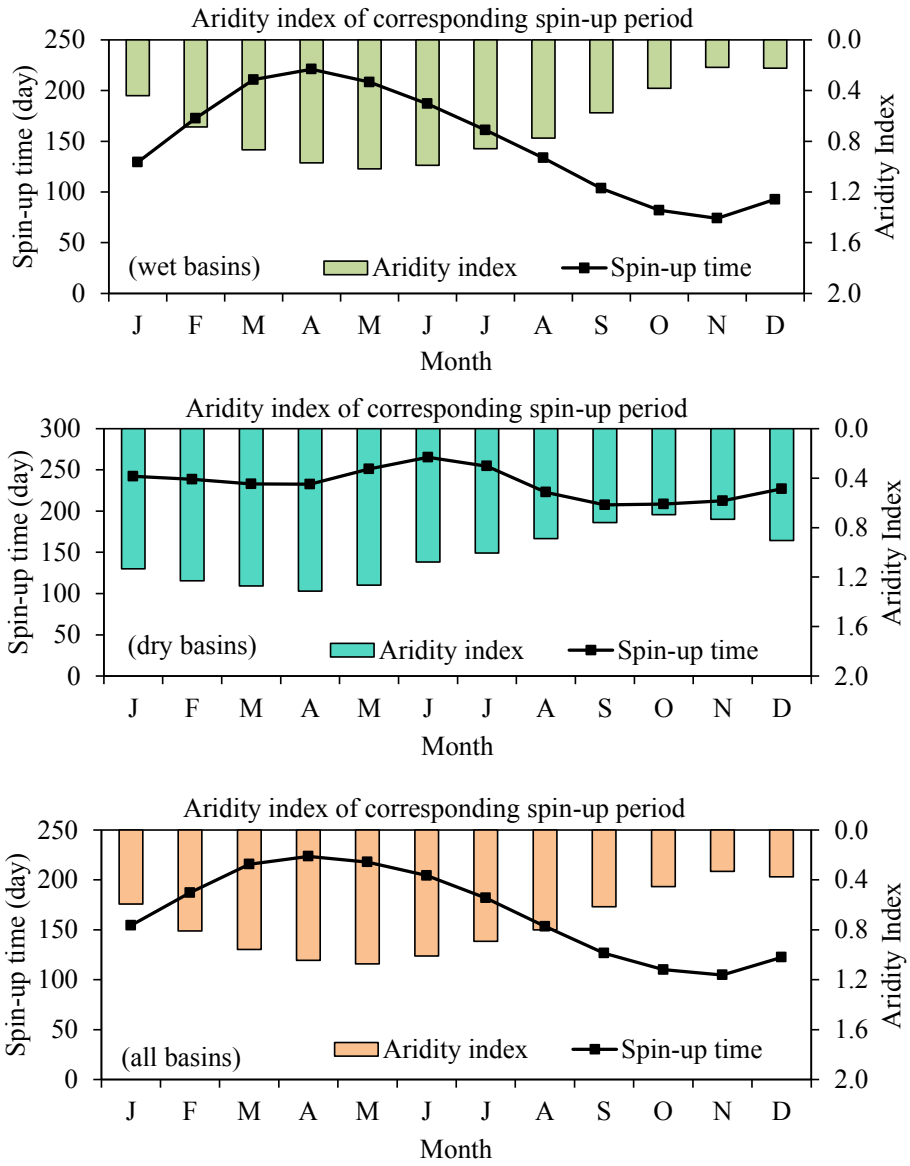

Figure 6.6: Seasonal variations of mean spin-up times and their corresponding aridity index.

study, Cosgrove et al. (2003) provided that the spin-up times are two to 17 times larger depending on the nature of initialisation and PC threshold. Although the methodologies and studied models are completely different, these literature agrees the overall comparative weight of SMM timescale and the spin-up time. The SMM timescale accounts the number of lag days that requires for the soil moisture autocorrelations to drop below the threshold significance at 95% confidence level. Therefore, in SMM timescale calculation, a complete shedding of soil moisture anomalies is not counted. On the other hand, in spin-up analysis a complete equilibrium was hunted thus requires longer times.

However, the seasonal cycles of model spin-up time particularly for the wet basins seems to be inconsistent with that of SMM timescale presented in Chapter 4 and Rahman et al. (2015). The seasonal cycles of wet basins' model spin-up time shows almost opposite cycles that of SMM timescale (Fig. 6.5 and 6.6). This variation might be resulted due to the

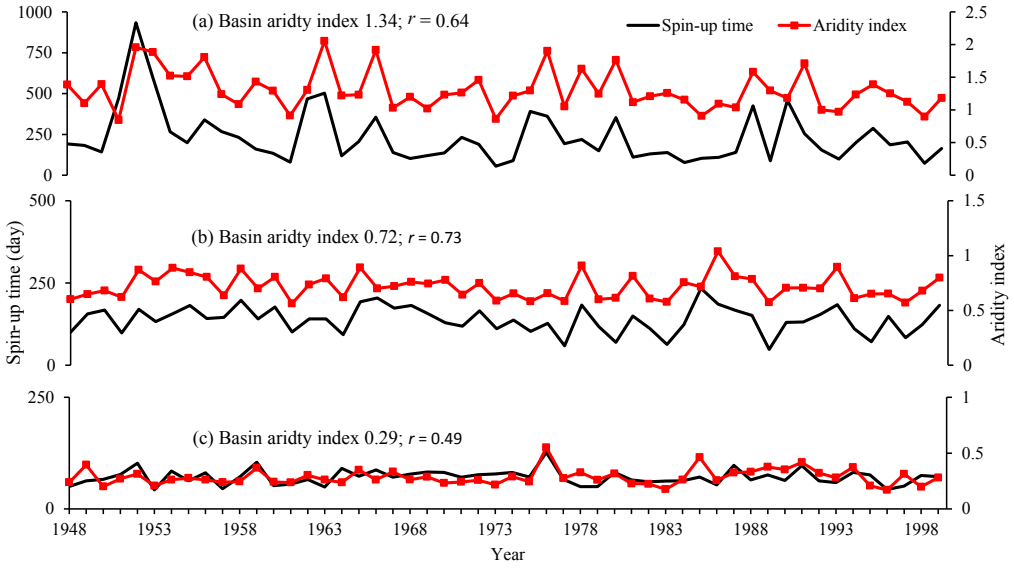

Figure 6.7: Year-wise model spin-up time and the aridity index. r is the correlation coefficient between yearly model spin-up time and aridity index.

difference in approaches of computations. Two cycles are estimated based on two separate methodologies and contain certain assumptions. Moreover, SMM timescale seasonality were drawn based on the SMM timescale of the first day of every month. SMM timescale study assumes that SMM timescale of the first day of any month is representative to the whole month. The spin-up time seasonality are drawn from large number of data sets, and thus is believed to be better representative to the reality. SMM timescale seasonality calculated from all the days of the year would achieve better agreement between the cycles. Additionally, the model spin-up study considering equilibrium for different state variable (sensible/latent heat flux, total soil moisture, root zone soil moisture, depth of water table, discharge, ground water storage etc.) would provide different results. On the other hand, the seasonal cycles of model spin-up time are quite consistent with that of SMM timescales for dry basins. Both the cycles show two peaks in summer and winter.

The model spin-up time highly varies (mean range is 154 days) throughout the year for wet basins. On the other hand, in dry basins, the spin-up time varies moderately (mean range 78 days) from month to month. Basin-wise monthly mean spin-up times are presented in Table 6.3.

The overall spin-up time is shorter in wet basins (wet basins mean 148 days) than those of dry basins (dry basins mean 233 days). This is consistent with Rahman et al. (2015); Rodell et al. (2005); Lim et al. (2012); Cosgrove et al. (2003). Figure 6.6 suggests that the model spin-up is mainly influenced by the aridity index during the corresponding spin-up period.

6.3.3 XAJ model spin-up time and basin aridity index

Yearly mean model spin-up time shows high association with the corresponding year's aridity index for all the basins. Figure 6.7 displays the yearly time series of model spin-up time and

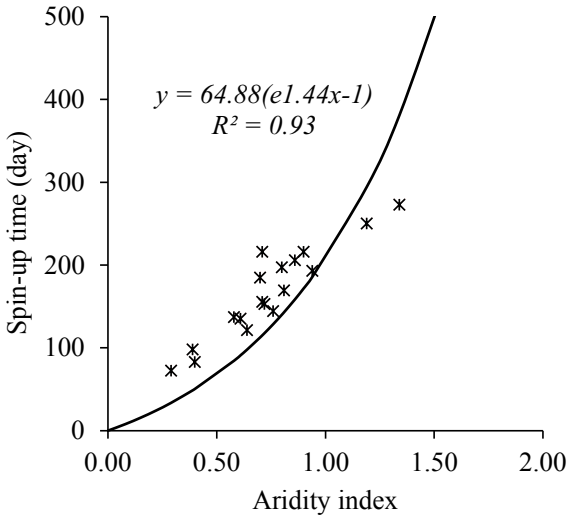

Figure 6.8: Relationship between XAJ model spin-up time and corresponding basin's annual aridity index.

the aridity index for three river basins representing three aridity conditions. This figure also suggests that aridity index is a dominant factor that affect the speed of model equilibrium.

Computed basin-wise XAJ model spin-up time reveals an exponential relationship with basin annual aridity index (calculated from independent data sets) with a R^2 value of 0.93 (Fig. 6.8). The relationship between the basin aridity index and model spin-up time can be expressed by Eq. (6.2).

$$\tau_{Xsp} = 64.88\left(e^{1.44\zeta} - 1\right) \tag{6.2}$$

where τ_{Xsp} is the XAJ model spin-up time in days and ζ is the basin aridity index.

This relationship could be useful for roughly estimating the XAJ model spin-up time and could be handy for simulations with better confidence. However, it should be noted that this relationship is based on the daily scale model simulation only, thus the XAJ model spin-up time for shorter or longer scale might be different.

6.4 Conclusions

When a model is calibrated with an unusual initial condition, the model undergoes some adjustment process to reach the normal equilibrium state. Model outputs during this spin-up process are highly affected by the initial conditions, and often unrealistic or misleading. Therefore, understanding this spin-up period has been the interest of modelling communities, particularly for the LSMs. Most spin-up studies are done based on a recursive simulations using a single year climatologies. Arguably, conclusions based on this recursive model runs might be erroneous due the lack of representativeness in the climatological extremes within the single year climatology. Moreover, researchers used different thresholds to define the model equilibrium conditions, and thus lost the comparability or uniformity. Furthermore,

recursive simulations based spin-up outcomes does not provide any insight about the seasonality of spin-up time.

Aiming to solve these limitations, this study detects and analyses the seasonality of spin-up time using multi-year climatologies adopting new techniques of model equilibrium definition. The spin-up time shows high seasonality and mainly controlled by the aridity index of model forcing. This analysis suggests that model spin-up time could vary based on the simulation start time of year. The simulation that starts from month of January might achieves the equilibrium quickly as compared that starts from the month of May. However, this conclusions are based on the American climatic conditions and it might show different seasonal cycles elsewhere.

The spin-up time displays exponential relationship with the aridity index. This relationship allows estimating the XAJ model spin-up time using only precipitation and evaporation information only. Estimation of the XAJ model spin-up time could be valuable to reduce uncertainty associated with guessing spin-up time, simply based on feeling or experience. Prior information about model spin-up time would allow us to fully use the information included in short data records under data scarce situation.

Appendix A

Development of soil moisture autocorrelation equation

The water balance of the soil column for a typical land surface model (LSM) under no snow condition, for the month n of year y can be written as Eq. (A.1)

$$C_s w_{n+lag,y} = C_s w_{n,y} + P_{n,y} - E_{n,y} - Q_{n,y} \tag{A.1}$$

where C_s, $w_{n,y}$ and $w_{n+lag,y}$ are the water holding capacity of the soil column, degree of soil moisture saturation at day n and day $n + lag$ of year y respectively. $P_{n,y}$, $E_{n,y}$ and $Q_{n,y}$ are the accumulated precipitation, evapotranspiration and streamflow during the time steps $(n, n + lag)$ of year y respectively.

Based on the results of Koster and Milly (1997), a constant linear relationship was assumed between the evaporation and runoff fluxes with the mean soil moisture of the month. The mean soil moisture was calculate as the average of the iniatial and final values, thus, the the assumptions looks like the following Eqs. (A.2-A.3):

$$\frac{Q_{n,y}}{P_{n,y}} = a_n \frac{w_{n,y} + w_{n+lag,y}}{2} + b_n \tag{A.2}$$

$$\frac{E_{n,y}}{R_{n,y}} = c_n \frac{w_{n,y} + w_{n+lag,y}}{2} + d_n \tag{A.3}$$

where $w_{n,y}$, $w_{n+lag,y}$, $E_{n,y}$ and $P_{n,y}$ express the same meaning as mentioned above. $\bar{R}_n$ is the radiation. a_n, b_n, c_n and $d_n$are parameters after Koster and Milly (1997).

Substituting Eq. (A.2) and (A.3) into Eq. (A.1) will yield Eq. (A.4):

$$C_s w_{n+lag,y} = C_s w_{n,y} + P_{n,y} - R_{n,y} c_n \left(\frac{w_{n,y} + w_{n+lag,y}}{2} \right) - d_n R_{n,y} - P_{n,y} a_n \left(\frac{w_{n,y} + w_{n+lag,y}}{2} \right) - b_n P_{n,y} \tag{A.4}$$

By re-arranging we get Eq. (A.5):

$$\left(C_s + \frac{c_n R_{n,y}}{2} + \frac{a_n P_{n,y}}{2}\right) w_{n+lag,y} = \left(C_s - \frac{c_n R_{n,y}}{2} - \frac{a_n P_{n,y}}{2}\right) w_{n,y} + (P_{n,y} - d_n R_{n,y} - b_n P_{n,y}) \quad \text{(A.5)}$$

Separating the soil moisture, precipitation and net radiation into their mean components (shown by overline) for the month n of year y and the corresponding interannual anomalies (shown by primes) we get Eqs. (A.6-A.9):

$$w_{n,y} = \bar{w}_n + \acute{w}_{n,y} \quad \text{(A.6)}$$

$$w_{n+lag,y} = \bar{w}_{n+lag} + \acute{w}_{n+lag,y} \quad \text{(A.7)}$$

$$P_{n,y} = \bar{P}_n + \acute{P}_{n,y} \quad \text{(A.8)}$$

$$R_{n,y} = \bar{R}_n + \acute{R}_{n,y} \quad \text{(A.9)}$$

By inserting Eqs. (A.6-A.9) into Eq. (A.5), we get Eq. (A.10):

$$\left(C_s + \tfrac{c_n \bar{R}_n}{2} + \tfrac{c_n \acute{R}_{n,y}}{2} + \tfrac{a_n \bar{P}_n}{2} + \tfrac{a_n \acute{P}_{n,y}}{2}\right)(\bar{w}_{n+lag} + \acute{w}_{n+lag,y}) =$$
$$\left(C_s - \frac{c_n \bar{R}_n}{2} - \frac{c_n \acute{R}_{n,y}}{2} - \frac{a_n \bar{P}_n}{2} - \frac{a_n \acute{P}_{n,y}}{2}\right)(\bar{w}_n + \acute{w}_{n,y}) \quad \text{(A.10)}$$

By solving, we get Eq. (A.11)

$$2C_s\bar{w}_{n+lag} + c_n\bar{R}_n\bar{w}_{n+lag} + c_n\acute{R}_{n,y}\bar{w}_{n+lag} + a_n\bar{P}_n\bar{w}_{n+lag} + a_n\acute{P}_{n,y}\bar{w}_{n+lag} + 2C_s\acute{w}_{n+lag} +$$
$$c_n\bar{R}_n\acute{w}_{n+lag,y} + c_n\acute{R}_{n,y}\acute{w}_{n+lag,y} + a_n\bar{P}_n\acute{w}_{n+lag,y} + a_n\acute{P}_{n,y}\acute{w}_{n+lag,y} = 2C_s\bar{w}_n - c_n\bar{R}_n\bar{w}_n -$$
$$c_n\acute{R}_{n,y}\bar{w}_n - a_n\bar{P}_n\bar{w}_n - a_n\acute{P}_{n,y}\bar{w}_n + 2C_s\acute{w}_{n,y} - c_n\bar{R}_n\acute{w}_{n,y} - c_n\acute{R}_{n,y}\acute{w}_{n,y} - a_n\bar{P}_n\acute{w}_{n,y} -$$
$$a_n\acute{P}_{n,y}\acute{w}_{n,y} + 2\bar{P}_n + 2\acute{P}_{n,y} - 2d_n\bar{R}_n - 2d_n\acute{R}_{n,y} - 2b_n\bar{P}_n - 2b_n\acute{P}_{n,y} \quad \text{(A.11)}$$

Now by ignoring the higher order terms and rearranging we get Eqs. (A.12) and (A.13):
{all mean components (i.e. $2C_s\bar{w}_{n+lag}, c_n\bar{R}_n\bar{w}_{n+lag}, a_n\bar{P}_n\bar{w}_{n+lag}, 2C_s\bar{w}_n, c_n\bar{R}_n\bar{w}_n, a_n\bar{P}_n\bar{w}_n, 2\bar{P}_n$, $2d_n\bar{R}_n, 2b_n\bar{P}_n$) and the terms in the form of $\acute{w}_{n,y} - \acute{w}_{n+lag,y}$; i.e. $c_n\acute{R}_{n,y}(\acute{w}_{n,y} - \acute{w}_{n+lag,y})$ and $c_n\acute{P}_{n,y}(\acute{w}_{n,y} - \acute{w}_{n+lag,y})$}
$2C_s\acute{w}_{n+lag,y} + c_n\bar{R}_n\acute{w}_{n+lag,y} + a_n\bar{P}_n\acute{w}_{n+lag,y} = 2C_s\acute{w}_{n,y} - c_n\bar{R}_n\acute{w}_{n,y} - a_n\bar{P}_n\acute{w}_{n,y} + 2\acute{P}_{n,y}$-

$$a\acute{P}_{n,y}\bar{w}_n - a_n\acute{P}_{n,y}\bar{w}_{n+lag} - 2b_n\acute{P}_{n,y} - c_n\acute{R}_{n,y}\bar{w}_n - c_n\acute{R}_{n,y}\bar{w}_{n+lag} - 2d_n\acute{R}_{n,y} \quad \text{(A.12)}$$

$$\left(2C_s + c_n\bar{R}_n + a_n\bar{P}_n\right)\acute{w}_{n+lag,y} = \left(2C_s - c_n\bar{R}_n - a_n\bar{P}_n\right)\acute{w}_{n,y} + \left(2 - a_n\bar{w}_n - a_n\bar{w}_{n+lag} - 2b_n\right)\acute{P}_{n,y}$$
$$- \left(c_n\bar{w}_n - c_n\bar{w}_{n+lag} - 2d_n\right)\acute{R}_{n,y} \quad \text{(A.13)}$$

Substituting $\bar{w}_{mid} = \frac{\bar{w}+\bar{w}_{n+lag}}{2}$, into Eq. (A.13) we get:

$$\left(2C_s + c_n\bar{R}_n + a_n\bar{P}_n\right)\acute{w}_{n+lag,y} = \left(2C_s - c_n\bar{R}_n - a_n\bar{P}_n\right)\acute{w}_{n,y} + \left\{2 - 2\left(a_n\bar{w}_{mid} + b_n\right)\right\}\acute{P}_{n,y}$$
$$\left(c_n\bar{w}_{mid} + d_n\right)2\acute{R}_{n,y} \quad \text{(A.14)}$$

Rearranging Eq. (A.14), we get Eq. (A.15):

$$\acute{w}_{n+lag,y} = \frac{\left(2C_s - c_n\bar{R}_n - a_n\bar{P}_n\right)\acute{w}_{n,y}}{\left(2C_s + c_n\bar{R}_n + a_n\bar{P}_n\right)} + \frac{\left\{2 - 2\left(a_n\bar{w}_{mid} + b_n\right)\right\}\acute{P}_{n,y}}{\left(2C_s + c_n\bar{R}_n + a_n\bar{P}_n\right)} - \frac{\left(c_n\bar{w}_{mid} + d_n\right)2\acute{R}_{n,y}}{\left(2C_s + c_n\bar{R}_n + a_n\bar{P}_n\right)} \quad \text{(A.15)}$$

Now substituting $A_n = \frac{\left(2C_s - c_n\bar{R}_n - a_n\bar{P}_n\right)}{\left(2C_s + c_n\bar{R}_n + a_n\bar{P}_n\right)}$, $B_n = \frac{\left\{2-2\left(a_n\bar{w}_{mid}+b_n\right)\right\}}{\left(2C_s + c_n\bar{R}_n + a_n\bar{P}_n\right)}$ and $H_n = \frac{\left(c_n\bar{w}_{mid}+d_n\right)2}{\left(2C_s + c_n\bar{R}_n + a_n\bar{P}_n\right)}$ into Eq. (A.15), we get Eq. (A.16)

$$\acute{w}_{n+lag,y} = A_n\acute{w}_{n,y} + B_n\acute{P}_{n,y} - H_n\acute{R}_{n,y} \quad \text{(A.16)}$$

Again substituting $\acute{F}_{n,y} = B_n\acute{P}_{n,y} - H_n\acute{R}_{n,y}$, into Eq. (A.16), we get Eq. (A.17):

$$\acute{w}_{n+lag,y} = A_n\acute{w}_{n,y} + \acute{F}_{n,y} \quad \text{(A.17)}$$

Equation (A.17) is now the basis for autocorrelation equation.

Multiplying both sides of Eq. (A.17) by $\acute{w}_{n,y}$, we get Eq. (A.18):

$$\acute{w}_{n,y}\acute{w}_{n+lag,y} = A_n\acute{w}_{n,y}\acute{w}_{n,y} + \acute{F}_{n,y}\acute{w}_{n,y} \quad \text{(A.18)}$$

By replacing anomalies with their corresponding values and means (i.e. $w_{n,y} = \bar{w}_n + \acute{w}_{n,y}$) into Eq. (A.18), we get Eq. (A.19):

$$\left(w_{n,y} - \bar{w}_n\right)\left(w_{n+lag,y} - \bar{w}_{n+lag}\right) = A_n\left(w_{n,y} - \bar{w}_n\right)^2 + \left(F_{n,y} - \bar{F}_n\right)\left(w_{n,y} - \bar{w}_n\right) \quad \text{(A.19)}$$

Taking time mean for N years, the Eq. (A.19) can be written as Eq. (A.20):

$$\frac{1}{N}\sum\left(w_{n,y} - \bar{w}_n\right)\left(w_{n+lag,y} - \bar{w}_{n+lag}\right) = \frac{1}{N}\sum A_n\left(w_{n,y} - \bar{w}_n\right)^2 + \frac{1}{N}\sum\left(F_{n,y} - \bar{F}_n\right)\left(w_{n,y} - \bar{w}_n\right) \quad \text{(A.20)}$$

Based on the definition i.e. variance of X $\{\sigma_x^2 = \frac{1}{N}\sum\left(x_i - \bar{x}\right)^2\}$and covariance of X, Y $\{COV\left(x,y\right) = \frac{1}{N}\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)\}$, we can rewrite the Eq. (A.20) as Eq. (A.21):

$$COV\left(w_n, w_{n+lag}\right) = A_n\sigma_{w_n}^2 + COV\left(w_n, F_n\right) \quad \text{(A.21)}$$

Where, $\sigma_{w_n}^2$refers the variance of $w_{n,y}$ and $COV\left(X,Y\right)$ represents to the covariance between variables X and Y.

Now based on the definition we can write the autocorrelation equation between initial soil moisture (w_n)and the soil moisture at the end of the time step ($w_{n+lag,y}$) as Eq. (A.22):

$$\rho\left(w_n, w_{n+lag}\right) = \frac{COV\left(w_n, w_{n+lag}\right)}{\sigma_{w_n}\sigma_{w_{n+lag}}} \tag{A.22}$$

By substituting Eq. (A.21) into Eq. (A.22) we get Eq. (A.23):

$$\rho\left(w_n, w_{n+lag}\right) = \frac{A_n\sigma_{w_n}^2 + COV\left(w_n, F_n\right)}{\sigma_{w_n}\sigma_{w_{n+lag}}} \tag{A.23}$$

Rearranging we get Eq. (A.24),

$$\rho\left(w_n, w_{n+lag}\right) = \frac{\sigma_{w_n}}{\sigma_{w_{n+lag}}}\left(A_n + \frac{COV\left(w_n, F_n\right)}{\sigma_{w_n}^2}\right) \tag{A.24}$$

Equation (A.24) is now the final form of Koster and Suarez (2001).

Considering the limitations in Eq. (A.24) by using the soil moisture variability at the $n + lag$ time step Seneviratne et al. (2006a), Seneviratne and Koster (2012) revised the Eq. (A.24) to replace $\sigma_{w_{n+lag}}$term.

By replacing anomalies with their corresponding values and means (i.e. $w_{n,y} = \bar{w}_n + \acute{w}_{n,y}$) into Eq. (A.17), we get Eq. (A.25):

$$w_{n+lag,y} - \bar{w}_{n+lag} = A_n\left(w_{n,y} - \bar{w}_n\right) + \left(F_{n,y} - \bar{F}_n\right) \tag{A.25}$$

Taking square,

$$\left(w_{n+lag,y} - \bar{w}_{n+lag}\right)^2 = \left\{A_n\left(w_{n,y} - \bar{w}_n\right) + \left(F_{n,y} - \bar{F}_n\right)\right\}^2 \tag{A.26}$$

$$\left(w_{n+lag,y} - \bar{w}_{n+lag}\right)^2 = A_n^2\left(w_{n,y} - \bar{w}_n\right)^2 + 2A_n\left(w_{n,y} - \bar{w}_n\right)\left(F_{n,y} - \bar{F}_n\right) + \left(F_{n,y} - \bar{F}_n\right)^2 \tag{A.27}$$

Taking time mean for N years,

$$\frac{1}{N}\sum\left(w_{n+lag,y} - \bar{w}_{n+lag}\right)^2 = \frac{1}{N}\sum A_n^2\left(w_{n,y} - \bar{w}_n\right)^2 + \frac{1}{N}\sum 2A_n\left(w_{n,y} - \bar{w}_n\right)\left(F_{n,y} - \bar{F}_n\right) + \frac{1}{N}\sum\left(F_{n,y} - \bar{F}_n\right)^2 \tag{A.28}$$

Again, based on the definition, we can rewrite the Eq. (A.28) as:

$$\sigma_{w_{n+lag}}^2 = A_n^2\sigma_{w_n}^2 + 2A_nCOV\left(w_n, F_n\right) + \sigma_{F_n}^2 \tag{A.29}$$

By rearranging and substituting $COV\left(w_n, F_n\right)$ term into Eq. (A.29), we can write as:

$$\sigma_{w_{n+lag}}^2 = A_n^2\sigma_{w_n}^2 + 2A_n\sigma_{w_n}\sigma_{F_n}\rho\left(w_n, F_n\right) + \sigma_{F_n}^2 \tag{A.30}$$

Where,

$\rho\left(w_n, F_n\right)$ is the correlation between the variables w_n and F_n by definition {correlation between X, Y $\rho\left(x, y\right) = \frac{COV(x,y)}{\sigma_x\sigma_y}$}

Equation (A.30) again can be rewritten as:

$$\sigma_{w_{n+lag}} = \sqrt{A_n^2 \sigma_{w_n}^2 + 2A_n \sigma_{w_n} \sigma_{F_n} \rho\left(w_n, F_n\right) + \sigma_{F_n}^2} \tag{A.31}$$

Therefore, with assumptions in Eq. (A.6) to (A.9), using semi-implicit form of those equations, and replacing $\sigma_{w_{n+lag}}$ term, the revised form of Eq. (A.24) Seneviratne and Koster (2012) becomes:

$$\rho\left(w_n, w_{n+lag}\right) = \frac{\sigma_{w_n} A_n + \sigma_{F_n} \rho\left(w_n, F_n\right)}{\sqrt{A_n^2 \sigma_{w_n}^2 + 2A_n \sigma_{w_n} \sigma_{F_n} \rho\left(w_n, F_n\right) + \sigma_{F_n}^2}} \tag{A.32}$$

Appendix B

Supplementary figures

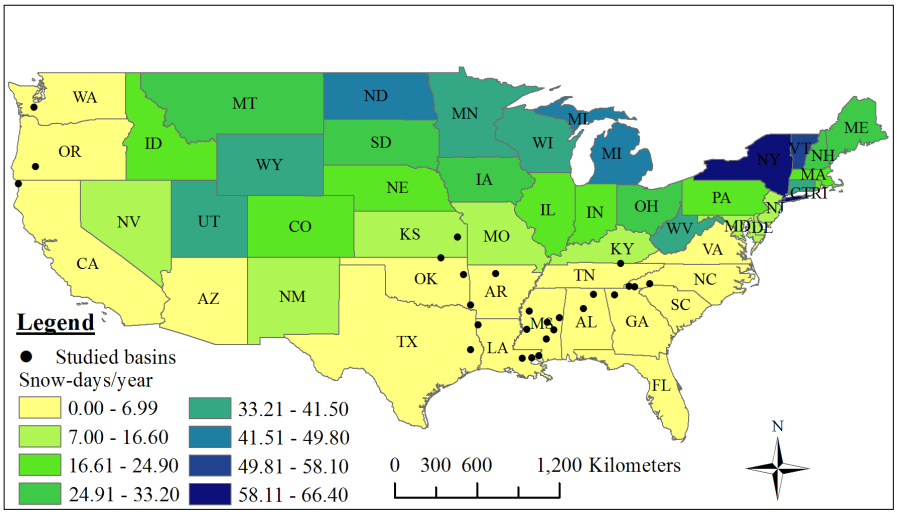

Figure B.1: Snow-day map of USA mainland and stream gauge locations of SMM studied river basins.

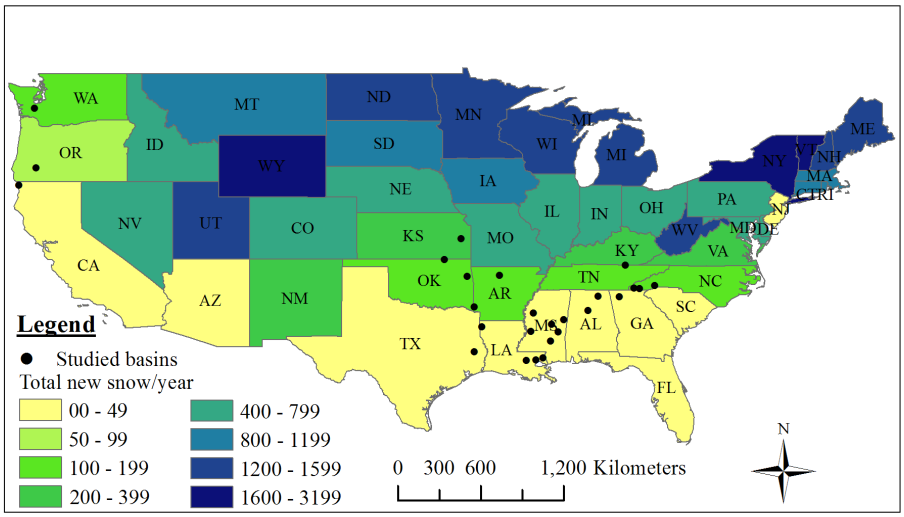

Figure B.2: Total new snow map of USA mainland and stream gauge locations of SMM studied river basins.

Bibliography

Ajami, H., Evans, J., McCabe, M., and Stisen, S. (2014a). Technical note: Reducing the spin-up time of integrated surface water–groundwater models. *Hydrology and Earth System Sciences Discussions*, 11(6):6969–6992.

Ajami, H., McCabe, M. F., Evans, J. P., and Stisen, S. (2014b). Assessing the impact of model spin-up on surface water-groundwater interactions using an integrated hydrologic model. *Water Resources Research*, 50(3):2636–2656.

Berthet, L., Andréassian, V., Perrin, C., and Javelle, P. (2009). How crucial is it to account for the antecedent moisture conditions in flood forecasting? comparison of event-based and continuous approaches on 178 catchments. *Hydrology and Earth System Sciences Discussions*, (13):p–819.

Bonan, G. B. and Stillwell-Soller, L. M. (1998). Soil water and the persistence of floods and droughts in the mississippi river basin. *Water Resources Research*, 34(10):2693–2701.

Castillo, V., Gomez-Plaza, A., and Martınez-Mena, M. (2003). The role of antecedent soil water content in the runoff response of semiarid catchments: a simulation approach. *Journal of Hydrology*, 284(1):114–130.

Chen, F. and Mitchell, K. (1999). Using the gewex/islscp forcing data to simulate global soil moisture fields and hydrological cycle for 1987-1988. *Journal of the Meteorological Society of Japan*, 77(1 B):167–182.

Chen, T. H., Henderson-Sellers, A., Milly, P., Pitman, A., Beljaars, A., Polcher, J., Abramopoulos, F., Boone, A., Chang, S., Chen, F., et al. (1997). Cabauw experimental results from the project for intercomparison of land-surface parameterization schemes. *Journal of Climate*, 10(6):1194–1215.

Corduas, M. (2011). Clustering streamflow time series for regional classification. *Journal of Hydrology*, 407(1):73–80.

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Sheffield, J., Duan, Q., Luo, L., et al. (2003). Land surface model spin-up behavior in the north american land data assimilation system (nldas). *Journal of Geophysical Research: Atmospheres (1984–2012)*, 108(D22).

de Goncalves, L., Shuttleworth, W. J., Burke, E. J., Houser, P., Toll, D. L., Rodell, M., and Arsenault, K. (2006). Toward a south america land data assimilation system: Aspects of land surface model spin-up using the simplified simple biosphere. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 111(D17).

De Maesschalck, R., Jouan-Rimbaud, D., and Massart, D. L. (2000). The mahalanobis distance. *Chemometrics and intelligent laboratory systems*, 50(1):1–18.

Delworth, T. L. and Manabe, S. (1988). The influence of potential evaporation on the variabilities of simulated soil wetness and climate. *Journal of Climate*, 1(5):523–547.

Dirmeyer, P. A., Schlosser, C. A., and Brubaker, K. L. (2009). Precipitation, recycling, and land memory: An integrated analysis. *Journal of Hydrometeorology*, 10(1):278–288.

Entin, J. (1999). Evaluation of global soil wetness project soil moisture simulations. *J. Meteor. Soc. Japan*, 77:183–191.

Entin, J. K., Robock, A., Vinnikov, K. Y., Hollinger, S. E., Liu, S., and Namkhai, A. (2000). Temporal and spatial scales of observed soil moisture variations in the extratropics. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 105(D9):11865–11877.

Georgakakos, K. P., Bae, D.-H., and Cayan, D. R. (1995). Hydroclimatology of continental watersheds: 1. temporal analyses. *Water Resources Research*, 31(3):655–675.

Goodrich, D., Schmugge, T., Jackson, T., Unkrich, C., Keefer, T., Parry, R., Bach, L., and Amer, S. (1994). Runoff simulation sensitivity to remotely sensed initial soil water content. *Water Resources Research*, 30(5):1393–1405.

Gudmundsson, L., Tallaksen, L., Stahl, K., and Fleig, A. (2011). Low-frequency variability of european runoff. *Hydrology and Earth System Sciences*, 15(9):2853–2869.

Henderson-Sellers, A., McGuffie, K., and Pitman, A. (1996). The project for intercomparison of land-surface parametrization schemes (pilps): 1992 to 1995. *Climate Dynamics*, 12(12):849–859.

Hollinger, S. E. and Isard, S. A. (1994). A soil moisture climatology of illinois. *Journal of Climate*, 7(5):822–833.

Hong, S.-Y. and Kalnay, E. (2000). Role of sea surface temperature and soil-moisture feedback in the 1998 oklahoma–texas drought. *Nature*, 408(6814):842–844.

Jouan-Rimbaud, D., Massart, D., Saby, C., and Puel, C. (1997). Characterisation of the representativity of selected sets of samples in multivariate calibration and pattern recognition. *Analytica chimica acta*, 350(1):149–161.

Jouan-Rimbaud, D., Massart, D., Saby, C., and Puel, C. (1998). Determination of the representativity between two multidimensional data sets by a comparison of their structure. *Chemometrics and intelligent laboratory systems*, 40(2):129–144.

Khin, H. K., Lu, M., and Li, X. (2015). Development of a user friendly web-based rainfall runoff model. In *International Conference on Climate Change and Water & Environmental Management in Monsoon Asia, Bangkok, Thailand, 28-30 January*.

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C., Kanae, S., Kowalczyk, E., Lawrence, D., et al. (2004). Regions of strong coupling between soil moisture and precipitation. *Science*, 305(5687):1138–1140.

Koster, R. D., Mahanama, S., Yamada, T., Balsamo, G., Berg, A., Boisserie, M., Dirmeyer, P., Doblas-Reyes, F., Drewitt, G., Gordon, C., et al. (2010a). Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophysical Research Letters*, 37(2):L02402.

Koster, R. D., Mahanama, S. P., Livneh, B., Lettenmaier, D. P., and Reichle, R. H. (2010b). Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. *Nature Geoscience*, 3(9):613–616.

Koster, R. D. and Milly, P. (1997). The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models. *Journal of Climate*, 10(7):1578–1591.

Koster, R. D. and Suarez, M. J. (1992). Modeling the land surface boundary in climate models as a composite of independent vegetation stands. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 97(D3):2697–2715.

Koster, R. D. and Suarez, M. J. (2001). Soil moisture memory in climate models. *Journal of hydrometeorology*, 2(6):558–570.

Leroy, A. M. and Rousseeuw, P. J. (1987). Robust regression and outlier detection. *Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1987*, 1.

Li, X. and Lu, M. (2014). Application of aridity index in estimation of data adjustment parameters in the xinanjiang model. *Annual Journal of Hydraulic Engineering, JSCE*, 58:163–168.

Lim, Y.-J., Hong, J., and Lee, T.-Y. (2012). Spin-up behavior of soil moisture content over east asia in a land surface model. *Meteorology and Atmospheric Physics*, 118(3-4):151–161.

Lin, C. A., Wen, L., Lu, G., Wu, Z., Zhang, J., Yang, Y., Zhu, Y., and Tong, L. (2006). Atmospheric-hydrological modeling of severe precipitation and floods in the huaihe river basin, china. *Journal of hydrology*, 330(1):249–259.

Liu, Y. and Avissar, R. (1999). A study of persistence in the land-atmosphere system using a general circulation model and observations. *Journal of climate*, 12(8):2139–2153.

Lu, G., Wu, Z., Wen, L., Lin, C. A., Zhang, J., and Yang, Y. (2008). Real-time flood forecast and flood alert map over the huaihe river basin in china using a coupled hydro-meteorological modeling system. *Science in China Series E: Technological Sciences*, 51(7):1049–1063.

Lu, M. and Li, X. (2014). Time scale dependent sensitivities of the xinanjiang model parameters. *Hydrological Research Letters*, 8(1):51–56.

Mahalanobis, P. C. (1930). On tests and measures of group divergence. *Journal of the Asiatic Society of Bengal*, 26:541–588.

Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D., and Reichle, R. (2012). Soil moisture, snow, and seasonal streamflow forecasts in the united states. *Journal of Hydrometeorology*, 13(1):189–203.

Mahanama, S. P. and Koster, R. D. (2003). Intercomparison of soil moisture memory in two land surface models. *Journal of Hydrometeorology*, 4(6):1134–1146.

Manabe, S. (1969). The atmospheric circulation and the hydrology of the earth's surface.

Martens, H. and Naes, T. (1992). *Multivariate calibration*. John Wiley & Sons.

McLachlan, G. J. (1999). Mahalanobis distance. *Resonance*, 4(6):20–26.

Minet, J., Laloy, E., Lambot, S., and Vanclooster, M. (2011). Effect of high-resolution spatial soil moisture variability on simulated runoff response using a distributed hydrologic model. *Hydrology and Earth System Sciences*, 15(4).

Mueller, B. and Seneviratne, S. I. (2012). Hot days induced by precipitation deficits at the global scale. *Proceedings of the National Academy of Sciences*, 109(31):12398–12403.

Nash, J. and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part iâĂŤa discussion of principles. *Journal of hydrology*, 10(3):282–290.

Nicholson, S. (2000). Land surface processes and sahel climate. *Reviews of Geophysics*, 38(1):117–139.

Nikolopoulos, E. I., Anagnostou, E. N., Borga, M., Vivoni, E. R., and Papadopoulos, A. (2011). Sensitivity of a mountain basin flash flood to initial wetness condition and rainfall variability. *Journal of Hydrology*, 402(3):165–178.

Orth, R. (2013). *Persistence of soil moisture-Controls, associated predictability and implications for land surface climate*. PhD thesis, Diss., Eidgenossische Technische Hochschule ETH Zurich, Nr. 21186.

Orth, R. and Seneviratne, S. (2012a). Analysis of soil moisture memory from observations in europe. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 117(D15).

Orth, R. and Seneviratne, S. (2012b). Propagation of soil moisture memory to runoff and evapotranspiration. *Hydrology and Earth System Sciences Discussions*, 9(10):12103–12143.

Orth, R. and Seneviratne, S. (2013). Propagation of soil moisture memory to streamflow and evapotranspiration in europe. *Hydrology and Earth System Sciences*, 17(10):3895–3911.

Rahman, M. M. and Lu, M. (2015). Model spin-up behaviour for wet and dry basins: A case study using xinanjiang model. *Water*, 7(8):4256–4273.

Rahman, M. M., Lu, M., and Kyi, K. H. (2015). Variability of soil moisture memory for wet and dry basins. *Journal of Hydrology*, 523:107–118.

Ren, L., Huang, Q., Yuan, F., Wang, J., Xu, J., Yu, Z., and Liu, X. (2006). Evaluation of the xinanjiang model structure by observed discharge and gauged soil moisture data in the hubex/game project. *IAHS publication*, 303:153.

Ren-Jun, Z. (1992). The xinanjiang model applied in china. *Journal of Hydrology*, 135(1):371–381.

Robock, A., Schlosser, C. A., Vinnikov, K. Y., Speranskaya, N. A., Entin, J. K., and Qiu, S. (1998). Evaluation of the amip soil moisture simulations. *Global and Planetary change*, 19(1):181–208.

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., Liu, S., and Namkhai, A. (2000). The global soil moisture data bank.

Rodell, M., Houser, P., Berg, A., and Famiglietti, J. (2005). Evaluation of 10 methods for initializing a land surface model. *Journal of Hydrometeorology*, 6(2):146–155.

Schaake, J., Cong, S., and Duan, Q. (2006). The u.s. mopex data set. *IAHS Publication*.

Schlosser, C. A. and Milly, P. C. D. (2002). A model-based investigation of soil moisture predictability and associated climate predictability. *Journal of Hydrometeorology*, 3(4).

Schlosser, C. A., Slater, A. G., Robock, A., Pitman, A. J., Vinnikov, K. Y., Henderson-Sellers, A., Speranskaya, N. A., and Mitchell, K. (2000). Simulations of a boreal grassland hydrology at valdai, russia: Pilps phase 2 (d). *Monthly Weather Review*, 128(2):301–321.

Seck, A., Welty, C., and Maxwell, R. M. (2015). Spin-up behavior and effects of initial conditions for an integrated hydrologic model. *Water Resources Research*.

Senarath, S. U., Ogden, F. L., Downer, C. W., and Sharif, H. O. (2000). On the calibration and verification of two-dimensional, distributed, hortonian, continuous watershed models. *Water Resources Research*, 36(6):1495–1510.

Seneviratne, S. I. and Koster, R. D. (2012). A revised framework for analyzing soil moisture memory in climate data: Derivation and interpretation. *Journal of Hydrometeorology*, 13(1):404–412.

Seneviratne, S. I., Koster, R. D., Guo, Z., Dirmeyer, P. A., Kowalczyk, E., Lawrence, D., Liu, P., Mocko, D., Lu, C.-H., Oleson, K. W., et al. (2006a). Soil moisture memory in agcm simulations: analysis of global land-atmosphere coupling experiment (glace) data. *Journal of Hydrometeorology*, 7(5):1090–1112.

Seneviratne, S. I., Lüthi, D., Litschi, M., and Schär, C. (2006b). Land–atmosphere coupling and climate change in europe. *Nature*, 443(7108):205–209.

Shenk, J. and Westerhaus, M. (1991). Population definition, sample selection, and calibration procedures for near infrared reflectance spectroscopy. *Crop science*, 31(2):469–474.

Simmonds, I. and Lynch, A. H. (1992). The influence of pre-existing soil moisture content on australian winter climate. *International journal of climatology*, 12(1):33–54.

van den Hurk, B., Doblas-Reyes, F., Balsamo, G., Koster, R. D., Seneviratne, S. I., and Camargo Jr, H. (2012). Soil moisture effects on seasonal temperature and precipitation forecast scores in europe. *Climate dynamics*, 38(1-2):349–362.

Vinnikov, K. Y., Robock, A., Speranskaya, N. A., and Schlosser, C. A. (1996). Scales of temporal and spatial variability of midlatitude soil moisture. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 101(D3):7163–7174.

Vinnikov, K. Y. and Yeserkepova, I. (1991). Soil moisture: Empirical data and model results. *Journal of climate*, 4(1):66–79.

Wilson, M. and Atkinson, P. (2007). The use of remotely sensed land cover to derive floodplain friction coefficients for flood inundation modelling. *Hydrological processes*, 21(26):3576–3586.

Wood, E. F., Lettenmaier, D. P., Liang, X., Lohmann, D., Boone, A., Chang, S., Chen, F., Dai, Y., Dickinson, R. E., Duan, Q., et al. (1998). The project for intercomparison of land-surface parameterization schemes (pilps) phase 2 (c) red–arkansas river basin experiment:: 1. experiment description and summary intercomparisons. *Global and Planetary change*, 19(1):115–135.

Wu, W. and Dickinson, R. E. (2004). Time scales of layered soil moisture memory in the context ofland-atmosphere interaction. *Journal of climate*, 17(14):2752–2764.

Wu, Z., Lu, G., Wen, L., Lin, C. A., Zhang, J., and Yang, Y. (2007). Thirty-five year (1971–2005) simulation of daily soil moisture using the variable infiltration capacity model over china. *Atmosphere-ocean*, 45(1):37–45.

Yang, Z.-L., Dickinson, R., Henderson-Sellers, A., and Pitman, A. (1995). Preliminary study of spin-up processes in land surface models with the first stage data of project for intercomparison of land surface parameterization schemes phase 1 (a). *Journal of Geophysical Research: Atmospheres (1984–2012)*, 100(D8):16553–16578.

Zehe, E., Becker, R., Bárdossy, A., and Plate, E. (2005). Uncertainty of simulated catchment runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation. *Journal of hydrology*, 315(1):183–202.

Zhang, Y., Wei, H., and Nearing, M. (2011). Effects of antecedent soil moisture on runoff modeling in small semiarid watersheds of southeastern arizona. *Hydrology and Earth System Sciences*, 15(10):3171–3179.

Acknowledgement

I am sincerely grateful to Professor Dr. Lu Minjiao who provided me new insights to scientific thinking and personal views through his critical comments and suggestions. He spent his valuable time offering me discussions, reviews, straight advice and kept his patience and faith about my limited ability to perform this difficult task on time. I also like to convey my special gratitude to all the faculty members, for their valuable cooperations and flexibilities all the way throughout the coursework and weekly seminars.

Exceptional appreciation goes to Ms. Khin Thay Kyi for her crucial help and suggestions right through this research. Her moral and technical supports are highly valuable for every steps of this research. It would be difficult for me to finish this work on time without her timely developed web-platform and time to time facilitation.

I am most thankful to NUT friends, especially to Mr. Md. Sohrab Hossain, Mr. Balavenkatesh Rengaraj, Mr. Md. Shahed Uz Zaman, Ms. Shreejana Prajapati and many others Japanese lab mates whose friendship made my stay in Nagaoka extraordinary.

Particular acknowledgments to Mr. Ishii Tomoki and Mr. Dominic Joy who not only supported me during my initial stay at Nagaoka by all means but also offered a friendship to cherish rest of my life.

My sincere gratitude goes to all the staff members of International Affairs Division (Kokusaika) for their special services to me particularly for translating documents and letters written in Japanese. Their cordial and smiley efforts made my life easy by breaking the linguistic barriers.

My heartfelt thankfulness also goes to MEXT. This study would not be possible without the financial support given by MEXT.

I like to take this opportunity to show the supreme appreciation to my wife for her continuous moral support, encouragement and patience stay without me for such a long duration. Equally, I extend thankfulness to my parents who have prayed for my success.