ASSESSING CLIMATE AND TERRESTRIAL WATER STORAGE CONTROLS ON EVAPOTRANSPIRATION VARIABILITY:
TOWARDS IMPROVED UNDERSTANDING OF WATERSHEDS AS COUPLED NATURE-HUMAN SYSTEMS

BY

RUIJIE ZENG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor Ximing Cai, Chair
Professor Murugesu Sivapalan
Professor Albert J. Valocchi
Professor Richard M. Vogel, Tufts University

ABSTRACT

Terrestrial evapotranspiration (ET) is an important eco-hydrologic process the couples the land surface water and energy budgets, links the water, carbon and nutrient cycle, and represents the largest water consumption from agricultural sector. Although advances have been made in monitoring and simulating terrestrial ET in last decades, there are still challenges in reconciling and cross-validating ET observation and numerical model simulation results. In particular, due to human interferences (such as agricultural irrigation), existing knowledge obtained under natural conditions is inapplicable to intensively managed watersheds. Therefore, there is a pressing need to develop hydrologic theory that depicts watersheds as coupled nature-human systems, and to apply knowledge derived from the complex system to validate and diagnose existing hydrologic observations and models, and explore the inter-connects of hydrologic dynamics across scales.

This dissertation focuses on the ET temporal variability as a signature of watersheds as coupled nature-human systems, since ET variability is driven by the climatic fluctuations and modulated by hydrologic processes such as vegetation, snow dynamics and human water use. Based on general hydrologic laws on land surface water-energy coupling, this dissertation derives an Evapotranspiration Temporal VARiance Decomposition (ETVARD) framework for better understanding of both the climatic and hydrologic controls on ET temporal variability. Utilizing best available hydrologic observations, ETVARD quantifies the contributions from the variances and co-variances of climatic and terrestrial water storage change factors to ET variance at various temporal scale (e.g., monthly, seasonal and annual) for watersheds across a wide spectrum of climatic conditions (from humid to arid) under both natural and managed conditions.

As such, we derive hydrologic knowledge from the congruence among theories, observations and models. For multi-variable and multi-source hydroclimatic observations, ETVARD provides an independent diagnosis tool to detect the possible biases and uncertainties in observations and land surface models. Using ETVARD as a benchmark for inter-comparison of observation and models and through five systematically designed experiments, this dissertation identifies the inconsistencies in ET variance estimates among theories, observations and models, assesses the quality of multiple ET products, and provides guidelines to improve land surface model structure in capturing ET variance for the contiguous United States.

In particular, ETVARD identifies the temporal and spatial ET pattern changes due to extensive groundwater-based irrigation through a rea-world case study in the High Plains. The relation between ET and crop yield signatures (i.e., mean and variability) in rain-fed and irrigated crops reflects farmers' irrigation behavior heterogeneity in the formation of ET patterns, depending on farmers' preferences

between profit-maximization and risk-aversion. In addition, a power-law statistical relationship between ET mean and variability is developed from independent ET observations. While the differences in climate conditions and vegetation structures are reflected by ecosystems' water use preferences between consumption and variability, these water use preferences cluster on the same a power-law statistical relationship.

The comprehensive assessment on ET variance in this dissertation provides a synthesis from existing theories, observations and simulations towards improved understanding of ET variance at the watershed system level. The knowledge discovered in the dissertation also provides guidelines for conjointly managing the mean and variability of watershed responses to both natural and human driving forces in the context of coupled nature-human systems.

ACKNOWLEDGMENTS

First and for most, I would like to thank my supervisor, Professor Ximing Cai, for his invaluable guidance, mentoring and support throughout my graduate study over the past seven years. I still remember that it was a sunny cold winter afternoon when I met Dr. Cai for the first time. He told me that my graduate study would focus on human interferences on hydrologic cycle. That started my research journey! Over the years, his enthusiasm and devotion to research and teaching accompany me through the journey of discovery and inspire me to pursue a career in academia.

I gratefully acknowledge my Ph.D. committee members, Dr. Murugesu Sivapalan, Dr. Albert Valocchi and Dr. Richard Vogel. Specifically, I would like to thank Dr. Sivapalan for his in-depth thinking when I took my first course at University of Illinois, Dr. Vogel for all the encouragement and constructive discussions since 2012 when I just derived out the ETVARD framework, and Dr. Valocchi for all the inspiring discussions and tremendous support.

I would also like to take this chance to thank all my friends at the Ven Te Chow Hydrosystems Laboratory. They made this PhD journey much easier and more enjoyable. Finally, special thanks to my parents for their never-ending love and support, and to my wife, who is always ready to discuss the research ideas.

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

1.1 Background and motivation

Evapotranspiration (ET) is an important eco-hydrologic process the couples the land surface water and energy budget [*H Yang et al.*, 2008], links carbon and nutrient cycle [*Porporato et al.*, 2015] and represents water consumption in food and biomass production [*Housh et al.*, 2014]. As one of the largest hydrologic fluxes, ET accounts for two thirds of precipitation and consumes a significant amount of surface net radiation [*Brutsaert*, 2005]. In terms of water consumption, FAO estimated that agricultural irrigation uses about 70 percent of the world's total freshwater withdrawals via ET [*FAO*, 2016]. Some studies have argued that water resources management is essentially "*ET* management" [*Foster and Garduño*, 2004]. Therefore, a better understanding of the effects of climatic, hydrologic and anthropogenic factors on ET process is vitally important for hydrologic predictions and water resources management, especially in the context of non-stationarity [*Milly et al.*, 2008].

In addition to the natural forcings, anthropogenic factors play an increasingly role in changing the land surface water and energy budget. Human interferences (HI), including deforestation, irrigation, and urbanization have altered the terrestrial water flux distribution spatial and temporal patterns [*Gordon et al.*, 2005; *Vogel*, 2011]. Several studies have demonstrated that intensive farming, land use conversion and groundwater exploitation have affected regional and global ET flux [*Weiskel et al.*, 2007] and even climate systems [*DeAngelis et al.*, 2010; *Ferguson and Maxwell*, 2010]. Those studies addressed HI on ET at large spatial scales and long temporal scales, and demonstrated that in heavily managed agricultural watersheds, agricultural land use and/or irrigation possibly plays an important role on the inter-annual variability of ET, especially in semi-arid regions [*Allen et al.*, 2005; *Qiuhong Tang et al.*, 2009; *Cheng et al.*, 2011]. We argue that watersheds under intensive management should be considered as coupled nature-human systems (CNHS) [*Liu et al.*, 2007] that are featured by strong interaction and interdependence among natural processes and human land and water uses. Correspondingly, as the largest water consumption component, ET should be studied in the context of CNHS.

From a system perspective, there are many signatures that characterizes ET. For example, the long-term average captures the amount of ET in the water budget; the ET extreme value indicates how hydrologic system responses to drought events; the trend of ET shows how land surface processes evolve given the changing factors such as net surface radiation, wind speed and land use. Among these signatures, the temporal variability provides unique information of ET processes and is closely related to the management of agricultural water use. First, the temporal variability exhibits across a range of scales from decadal, annual, seasonal, daily, hourly to sub-hourly. The variability at each temporal scale shows specific properties of hydrologic dynamics. For example, the different water use patterns during a growing season and a non-growing season is embedded in the seasonal variability. Second, the temporal variability of ET reflects the intrinsic hydro-climatic variability and provides valuable information about ET sensitivity to changing environmental variables [*Milly and Dunne*, 2001; *Niemann and Eltahir*, 2005; *Roderick and Farquhar*, 2011; *McVicar et al.*, 2012; *Renner et al.*, 2012; *Yang et al.*, 2014b]. In addition to climate change, interferences introduced by human activities, such as conversion from natural vegetation to bio-fuel crops [*Le et al.*, 2011] and the expansion of irrigated crop land [*Zeng and Cai*, 2014], also significantly affect ET pattern. Therefore, understanding ET variability will unveil how ET processes change given the changing climatic and anthropogenic forcings, the key to improve weather and climate forecasting [*Ozdogan et al.*, 2010] and provide better guidelines for climate change adaptation [*Yohe et al.*, 2004]. Third, as the average of water consumption (i.e., irrigation) would help increase crop yield and farmers' profit, the variability of ET is a good indicator of the variability of crop yield and farmers' risk. In the context of "ET management", a robust water resources management practice should take the variability into account, in addition to the average profit. Hydroclimatic factors shape the ET variability, at the same time, farmers' water use behaviors modulate ET variability according to their preferences (e.g., risk-aversion, benefit-maximization). Therefore, the variance of ET captures both the fluctuation of a hydro-climatic system and human land and water use footprints.

1.2 Objectives

Advances have been made in monitoring and simulating ET over several decades. At the observation side, the efforts include remote-sensing signal retrieval [*Zhang et al.*, 2010; *Mu et al.*, 2011], fluxtower network development and data assimilation [*Pan and Wood*, 2006; *Munier et al.*, 2015; *Rodell et al.*, 2015]. Meanwhile, the land surface modelling community has developed many numerical models that include ET simulation, with different process representations, parameterizations, data requirements and model structures, such as Global Land-Atmosphere Coupling Experiment (GLACE) [*Koster et al.*, 2004] and Land Data Assimilation System (LDAS) [*Rodell et al.*, 2004]. Hydrologic observations and numerical simulations play complementary and inter-dependent roles in advancing our knowledge about the various hydrological processes and systems.

However, compared to the advances in ET observation and simulation, the ET theory development and application are limited in hydrology. Developing new theories and/or making better use of existing theories to underpin current models and data are urgently needed [*Kirchner*, 2006; *Beven*, 2012; *Clark et al.*, 2016]. A hydrologic theory that represents falsifiable conceptualizations of the real world is needed to justify the observation and/or model evaluation and diagnose the biases or errors involved in either observations or models, or both. We argue that hydrologic theories play a critical role in bridging the gap between model results and observations. Although observations, models and theories are not perfect, each contains complementary information about the real world. Observations can capture a broad range of hydrologic dynamics driven by climatic, biophysical and anthropogenic forcings [*Rodell et al.*, 2015]; models can predict the hydrologic responses to either stationary or non-stationary forcings and explore the feasible space of a hydrologic variable [*Kumar*, 2011]; theories are used to synthesize our understanding of hydrologic phenomena and expand hydrologic knowledge [*Kirchner*, 2006; *Clark et al.*, 2016].

Following current research advances, this dissertation aims to provide a theoretical framework to analyze ET temporal variability and apply the framework to address a number of theoretical and practical questions. The theoretical framework will both be a complementary approach to existing ET data and models and a bridge to reconcile the ET observations and simulations. The goal of this dissertation is to **provide a better understanding of the climatic and hydrologic controls on ET temporal variability, for both natural and managed watersheds, through an evapotranspiration temporal variance decomposition framework.**

1.3 Tasks

To achieve the research goal, specific tasks includes:

1) Deriving a theoretical framework to quantify contributions of climate (e.g., precipitation and net radiation) and terrestrial water storage change (by either natural processes or human activities) factors to ET variability based on generally hydrologic principle on land surface energy-water coupling.

2) Assessing watershed ET variability at various temporal scale (e.g., monthly, seasonal, annual) for watersheds across a wide spectrum of climatic conditions (from humid to arid) and with significant human interferences (e.g., irrigation), and identifying the dominant controlling factors on ET variability with the theoretical framework.

3) Applying the theoretical framework as a bridge to reconciling existing multi-source multi-variable hydroclimatic observations and multiple land surface models to find the congruence among the theory-observation-simulation triplet in term of ET variability.

4) With the knowledge obtained from Task 2 and 3, exploring the connections between ET signatures (specifically mean and temporal variance) and watershed properties (e.g., land use and vegetation types) and identifying possible general statistical law governing both natural and managed watersheds.

Through these specific tasks, the dissertation will synthesize the understanding of ET variability from observations, simulations, and theories to understand and manage a watershed in the context of coupled natural human system.

1.4 Hypotheses

This dissertation starts from deriving an evapotranspiration temporal variance decomposition framework, which is based on an empirical or semi-analytic theory on watershed energy-water coupling, namely, the Budyko curve. *Budyko* [1974] pioneered the estimation of long-term ET based on a coupled hydrologic cycle and the terrestrial energy budget. He asserted that a region's ET is largely controlled by two climatic factors: precipitation (P) and incident energy (usually represented by potential evaporation, PET). In arid regions (i.e., $PET/P \gg 1$), ET is mainly constrained by P; in humid regions (i.e., $PET/P \ll 1$), ET is controlled by energy supply (associated with PET); in between, ET is affected by both P and PET. Based on Budyko curve, *Koster and Suarez* [1999] attributed ET variance to the variance of precipitation. However, *Koster and Suarez* [1999] implicitly assumed that the watershed storage change is negligible in each period and their work only captured the climatic control on ET variance. Therefore, the hydrologic controls, such vegetation responses to climate fluctuation, farmers' irrigation decision, groundwater dynamics and snow thawing/melting, are not addressed in their study.

With the availability of terrestrial water storage data (from satellite observations and model simulations), this dissertation will extend Budyko curve and the ET variance relation developed by *Koster and Suarez* [1999] by incorporating terrestrial water storage into the derivation for ET variance. This study hypothesizes that terrestrial water storage (including soil moisture, groundwater and snow) provides a complimentary water buffer to atmospheric precipitation to sustain ET through related processes such as vegetation root water up-take, anthropogenic irrigation, groundwater recharge/discharge and snow thawing/melting. The responses of terrestrial water storage to climate variability (precipitation and potential evaporation) are captured by the co-variances among these hydrologic and climatic variables. Therefore, it is essential to incorporate terrestrial water storage into the analysis of ET variance.

If an intensively managed watershed is regarded as a coupled natural-human system, farmers' preferences, such as risk-aversion and benefit-maximization, determines their water use behaviors to reduce the climatic impact on crop yield. The goal to achieve better crop production, with higher crop yield and more stable crop yield, will propagate to the signatures of ET. Therefore, though the analysis of ET signatures (i.e., mean and variance in this study), we can capture the spatial and temporal pattern of ET changes due to intensive irrigation and explore how climatic, hydrologic and anthropogenic factors jointly shape ET in the context of coupled nature-human system. Farmers' irrigation behaviors reflect their preferences on crop production and responses to climate fluctuation, so do various ecosystems.

Different vegetation covers (e.g., trees, grasses, and shrubs) exist in various ecosystems with different climate conditions and soil properties. As the result of the co-evolution between vegetation and climate, different ecosystems will develop different water use strategies to optimize their goals. Although the goals and strategies of different natural and managed systems are different, they all respond to and are limited by the climate fluctuations. In the context of CNHS, we hypothesize that their water use strategies will be captured by the ET signatures (i.e., mean and variance). By analyzing the correspondence between ET signatures and the ecosystem properties, we can unveil the underlying water use strategies of different ecosystems.

1.5 Organization of the dissertation

With the objectives and tasks stated above, the rest of this dissertation consists of four parts in six main chapters.

Chapter 2 lays the foundation of this dissertation by deriving the general evapotranspiration temporal variance decomposition framework. Chapter 2 is built on the first hypothesis that terrestrial water storage and its responses to climate fluctuations modulate the viability of ET. By incorporating terrestrial storage change, Chapter 2 extends the Budyko hypothesis and the ET variance relation by *Koster and Suarez* [1999] and provide a comprehensive function for assessing ET temporal variance.

With the analytical framework developed in Chapter 2, Chapters 3 and 4 in the second part will assess the quantify the ET variances at different temporal scales and spatial scales with different controlling factors. Chapter 3 assesses the ET variances in 32 global basins with a wide range of climate conditions at both intra- and inter- annual scale (i.e., monthly, and annual ET variance). Chapter 3 identifies which climatic and hydrologic component dominates ET variance at intra- and inter- annual scale, respectively. Since the 32 global basins cover a large spatial extension where human interferences are relatively small, Chapter 3 focuses on natural watersheds. While Chapter 4 focuses on small watersheds using those in the High Plains of the CONUS with extensive groundwater-fed irrigation. This chapter therefore provides a detailed case study for ET variance under intensively managed systems, including the spatial and temporal change of ET pattern over 70 years at the seasonal scale (in accordance with the crop grow and non-growing season). Given the similarity of climate conditions in small basins in the High Plain, Chapter 4 shows an example of how irrigation behavior heterogeneity affects the ET signature, ending with implications for better watershed management in the context of as CNHS.

The third part addresses the methodological aspect of hydrologic knowledge discovery by assessing the congruence in the observation-model -theory triplet. The theory for ET variance derived in Chapter 2 serves in this chapter as a complementary approach to ET monitoring and modelling approach and a bridge to reconcile ET observations and simulations. In this part, Chapter 5 assess multi-variable multi-source observations by identifying the possible biases and uncertainties of multiple ET products in terms of their capability and compatibility in capturing of ET variance under the theoretical framework presented in Chapter 2. Chapter 6 focuses on the simulation side with the theoretical framework as a diagnosis tool and benchmark for multiple land surface model cross-evaluation, inter-comparison, and implication derivation for model improvement. Being companioned together, Chapters 5 and 6 illustrate how a generic hydrologic theory can be effectively used to bridge the gap between hydrologic observations and simulations, which illustrate our philosophy that hydrologic knowledge is advanced by finding the congruence among observation-simulation-theory triplet.

Being the last part of this thesis, Chapter 7 syntheses all studies in natural and managed watersheds in Chapters 2 to 6 and explores the connections between the signatures of ET (e.g., the relationship between mean and temporal variance). Hydrologic processes are inter-connected across scales, so do their signatures. The relationship between ET mean and temporal variance discovered in intensively managed watershed in Chapter 4 provides a clue for synthesis among different ecosystems with coupled nature and human components. Chapter 7 illustrates a general statistical law between ET mean and temporal variance across scales and discusses the role of ecosystem water use strategies in affecting the ET signatures. The identified relationship between ET mean and temporal variance will provide implications for understanding and possibly management of coupled nature-human system by trading off between system-wide water consumption and variability.

CHAPTER 2

DERIVING THE EVAPOTRANSPIRATION TEMPORAL VARIANCE DECOMPOSITION FRAMEWORK

This chapter lays the foundation of this dissertation and derives the Evapotranspiration Temporal VARiance Decomposition (ETVARD) framework.

2.1 Introduction

Budyko [1974] pioneered the estimation of long-term ET based on a coupled hydrologic cycle and the terrestrial energy budget. He asserted that a region's ET is largely controlled by two climatic factors: precipitation (*P*) and incident energy (usually represented by potential evaporation, *PET*). In arid regions (i.e., $PET/P \gg 1$), ET is mainly constrained by *P*; in humid regions (i.e., $PET/P \ll 1$), ET is controlled by energy supply (associated with *PET*); in between, ET is affected by both *P* and *PET*. The Budyko Hypothesis has been validated by observations all over the world [*Choudhury*, 1999; *Zhang et al.*, 2001]. Based on the Budyko Hypothesis, *Fu* [1981] and *H Yang et al.* [2008] derived analytical expressions, known as Budyko equation, which provides a framework to quantify long-term ET.

The Budyko equation has been used for ET sensitivity and variability analysis due to its explicit function form. For example, *Roderick and Farquhar* [2011] evaluated the derivatives of ET with respect to *P*, *PET* and a catchment property parameter to predict the effect of climate condition change on catchment water balance. *Niemann and Eltahir* [2005] studied the sensitivity of regional hydrology to climate change using Budyko equation and a physical model in the Illinois River basin and found that ET tends to dampen the signals in *P* and *PET*. *Han et al.* [2011] assessed long-term and annual water balances in Tarim Basin in China and found that influences on ET variability became increasingly apparent with the increase of irrigation in the arid basin. Especially, besides those assessments using models or data, *Koster and Suarez* [1999] proposed an analytical framework based on Budyko equation to quantify ET variance as below:

$$\sigma_{ET}^2 = [F(\bar{\phi}) - F'(\bar{\phi})\bar{\phi}]^2 \sigma_P^2 \qquad \text{Eqn.(2.1)}$$

where $\bar{\phi}$ is the long-term average arid index defined as *PET*/*P*; $F(\bar{\phi})$ is the Budyko equation. According to the results from this equation and a general circulation model, they found that water and energy availability appears to be critical factors controlling the inter-annual ET variance. Later, they validated Eqn.(2.1) using a global observation dataset for its predictability of ET variance [*Koster et al.*, 2006]. Following that, Eqn.(2.1) has been applied to assessing ET variance by many studies [*Arora*, 2002; *Sankarasubramanian and Vogel*, 2002; *Koster et al.*, 2006]. However, Eqn.(2.1) does not consider many other factors that are also important for ET variance. For example, based on the assessment of to 1337 catchments in the United States, *Sankarasubramanian and Vogel* [2002] found that the buffer effect of soil storage capacity could be an important factor on ET variability. *Koster et al.* [2006] pointed out that Eqn.(2.1) performs well in dry climate but the temporal coincidence of *P* and surface energy can affect ET variance in wetter climates, but these factors are not considered by Eqn.(2.1).

This study addresses the limitation of Eqn.(2.1) by re-examining its assumptions. First, Eqn.(2.1) is based on long-term average water balance assuming negligible storage change (i.e., *P* is the only water source for ET). At the annual or monthly time scale, however, *P* is not the sole source of water availability, since catchment storage change plays an important role to balance the water budget. The estimation of annual ET was found biased without considering subsurface water storage change [*Wang et al.*, 2009; *Istanbulluoglu et al.*, 2012]. Even at the long-term scale, successive groundwater exploitation provides an additional source for ET [*Siebert et al.*, 2010; *Döll et al.*, 2012]. Thus, incorporating catchment storage change caused by both natural factors and human activities will improve the understanding of ET temporal variance. Second, Eqn.(2.1) assumes that ET variance is driven by the fluctuation of *P* only and does not capture the effects from *PET* variance and the temporal coincidence between *P* and *PET*. As a result, Eqn.(2.1) is limited to arid regions where *P* dominates the hydrologic processes; however in moderate and wet climates, the effect of *P* on ET variance diminishes. An analysis of world-wide ET during the period of 1961–1999 [*Ukkola and Prentice*, 2013] shows that *P* accounts for 95% of the ET variance in dry basins, but only 55% in wet basins. Particularly, in cold areas, the accumulation and melting of snow pack is controlled by radiative energy, which further affects vegetation growth and ET flux [*Lute and Abatzoglou*, 2014]. As a result, *PET* becomes an essential factor in understanding ET variance in basins with limited energy supply. Furthermore in arid regions with intensive irrigation, *P* would not dominate ET as a result of irrigation application to maintain crop yield [*Han et al.*, 2011]. In such case, ET variance is closely related to farmers' response to climate fluctuation.

The goal of this study is to identify climate and human factors governing ET temporal variance by extending the relationship analytical framework of *Koster and Suarez* [1999] to a more comprehensive one, which incorporates the response of basins and human activities to climatic variability. The questions to address include: 1) how the fluctuation of climatic variables shapes ET variance in a wide spectrum of climate conditions; 2) how climate and human water use affect ET variance at various time scales (i.e., inter- and intra- annual scale). In session 2, we develop the framework for ET variance analysis based on Budyko Hypothesis and water balance and discuss the dominant factors affecting ET under various conditions. In session 3, we apply the framework in Murry-Darling River Basin to assess inter- and intra-annual ET variance. In session 4, we discuss some implications of the proposed framework and end with conclusions.

2.2 Theoretical framework for ET temporal variance

2.2.1 Catchment water balance in the Budyko equation

The water balance lumped over a catchment over a time interval of ΔT_i is:

$$\Delta S_i = P_i - ET_i - Q_i \quad \text{Eqn.(2.2)}$$

where ΔS is catchment storage change; P is precipitation; ET is actual evapotranspiration; Q is the runoff; and the subscript *i* represents the time interval ΔT_i, which can range from a month to decades. Over a long period when the catchment reaches equilibrium (i.e., flux-in balances flux-out and ΔS is negligible), P is the water source for ET and Q. At a small temporal scale (e.g., month), the water availability for ET and Q is adjusted by catchment storage. When catchment storage increases (i.e., ΔS is positive, such as snow pack accumulation and aquifer recharge), less available water is left for ET and Q. On the other hand, when catchment storage releases (i.e., ΔS is negative, such as snow melting and aquifer discharge), it provides additional water for ET and Q [*Wang et al.*, 2009]. As time scale becomes smaller, the role of catchment storage on water balance becomes significant in Eqn.(2.2). To account for the complementary effect of storage, the total available water (*P'*) for ET and Q is defined by rearranging Eqn.(2.2), which yields:

$$P_i' = P_i - \Delta S_i = ET_i + Q_i \quad \text{Eqn.(2.3)}$$

The total available water for ET does not only depend on the system input (i.e., atmospheric water supply), but also determined by catchment storage. Vegetation, soil moisture condition, groundwater table, and catchment management practices all affect the total water availability. The human impact items are not explicitly shown in Eqns. (1) and (2), however, the consumptive use is included in ET, and the return flow in Q or/and ΔS. In catchments where trans-boundary water is provided for use such as irrigation, the inflow to one catchment or the outflow to another catchment be added to or subtracted from P in Eqn.(2.2). Moreover, the spring flow or aquifer recharge will be accounted in ΔS.

The original Budyko hypothesis focuses on geographical zonality (i.e., spatial comparison) and is validated for long-term average over many catchments. *Fu* [1981] and *H Yang et al.* [2008] derived analytical solutions expressed as the long-term arid index ($\bar{\phi} = \overline{PET}/\bar{P}$) and evaporation index ($\overline{ET}/\bar{P}$) based on dimensional analysis and mathematical reasoning. Hereinafter, variables with over-bar denote long-term average. For example, the analytical solution obtained by *Fu* [1981] is:

$$\frac{\overline{ET}}{\bar{P}} = F(\bar{\phi}) = F\left(\frac{\overline{PET}}{\bar{P}}\right) = 1 + \frac{\overline{PET}}{\bar{P}} - \left[1 + \left(\frac{\overline{PET}}{\bar{P}}\right)^{\varpi}\right]^{1/\varpi} \quad \text{Eqn.(2.4)}$$

where ϖ is a parameter representing catchment characteristics. Since Eqn.(2.4) is based on long-term average, it assumes negligible catchment storage change (i.e., $\overline{\Delta S} = 0$) and atmospheric water is the only source for ET and Q. Although some studies have applied Fu's equation to smaller time scales and found a reasonable fit to observed data [*Choudhury*, 1999], the assumption that storage does not change has not been tested as being true at a short time scale.

To account for the storage change and satisfy the water balance, replacing atmospheric water supply (P) by total available water (P') in Eqn.(2.4), Fu's equation becomes:

$$\frac{ET_i}{P_i'} = F(\phi_i) = F\left(\frac{PET_i}{P_i'}\right) = 1 + \frac{PET_i}{P_i'} - \left[1 + \left(\frac{PET_i}{P_i'}\right)^{\varpi}\right]^{1/\varpi} \qquad \text{Eqn.(2.5)}$$

Or multiplying both side by P_i', which yields:

$$ET_i = P_i' + PET_i - \left({P_i'}^{\varpi} + {PET_i}^{\varpi}\right)^{1/\varpi} \qquad \text{Eqn.(2.6)}$$

The validity has been explored at annual scale in catchment with significant storage change [*Wang*, 2012]. Here, we adopt the catchment characteristics parameter ϖ in Eqn.(2.5) and Eqn.(2.6) same as that in the long-term Eqn.(2.3). The effect of ϖ is discussed in later section.

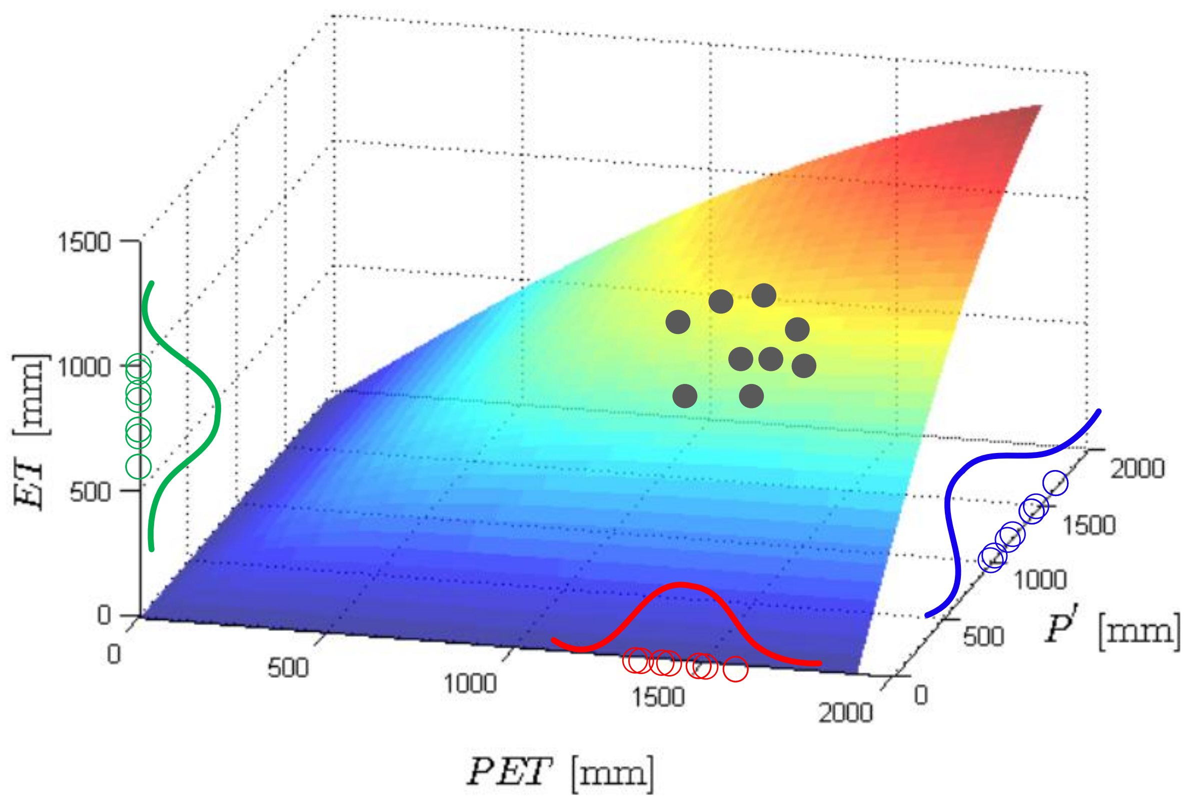


Figure 2.1. The hydro-climatic system state over a time interval is represented as one point on the surface. ET variability is represented by ET variance, that is, the second moment of the scatter points projected to ET axis.

Eqn.(2.5) and Eqn.(2.6) expresses the Budyko Hypothesis for time interval *i* by incorporating the water balance. Thus, the steady-state assumption (i.e., long-term storage change $\overline{\Delta S} = 0$) in the long-term Budyko equation Eqn.(2.4) is replaced by the water balance in Eqn.(2.3) to account for the role of catchment storage change. Within each time interval, the total water availability (P_i') is the sum of atmospheric water and catchment storage change. If the time period is large enough and the long term catchment storage change is negligible, Eqn.(2.4) and Eqn.(2.5) are essentially the same. In the space of (P_i', PET_i, ET_i), Eqn.(2.6) describes a system state surface for one catchment under various hydro-climatic conditions, as shown in Figure 2.1. For a specific catchment, given the energy supply (PET_i) and total water availability (P_i') over a time interval, the value of ET_i can be identified from the system state surface (P_i', PET_i, ET_i). Since Eqn.(2.6) does not involve the changing rate of ET (i.e., $\frac{\partial ET}{\partial t}$), it cannot describe the dynamics of catchment hydro-climatic variables. That is, Eqn.(2.6) tells where the points should be located on the state surface according to various water and energy supply combinations, but does not depict the temporal trajectory of ET. This limitation is beyond the scope of this paper and some promising approaches to handle the hydro-climatic system dynamics have been discussed in recent studies [*Donohue et al.*, 2010].

2.2.2 Derivation for ET temporal variance

For a specific catchment, the catchment ET over interval *i* is a point on the system state surface under total available water (*P'*) and energy supply (*PET*). For many intervals, there are cloud of points on the system state surface, as shown in Figure 2.1 for one particular basin. By projecting the sample points of system states to each axis, Eqn.(2.6) allows us to assess the statistics of ET_i given the statistics of P_i' and PET_i. For example, the first moment (i.e., expected value $\overline{ET}$) of ET_i distribution is captured by the long-term $\overline{PET}$ and $\bar{P}$ from the original Budyko equation in Eqn.(2.4). ET temporal variance, represented by second central moment of ET_i samples, is derived as follows.

The approximation of Eqn.(2.5) by the Taylor series expansion near the long-term mean climate condition $\bar{\phi}$ can be obtained by neglecting the higher order terms $\mathcal{O}[(\phi_i - \bar{\phi})^2]$:

$$F(\phi_i) = F(\bar{\phi}) + F'(\bar{\phi})(\phi_i - \bar{\phi}) + \mathcal{O}[(\phi_i - \bar{\phi})^2] \approx F(\bar{\phi}) + F'(\bar{\phi})\Delta\phi_i \qquad \text{Eqn.(2.7)}$$

where the deviation of the arid index with a specific time interval to its long-term mean is:

$$\Delta\phi_i = \phi_i - \bar{\phi} = \frac{PET_i}{P_i'} - \frac{\overline{PET}}{\bar{P}} = \frac{PET_i\bar{P} - \overline{PET}P_i'}{P_i'\bar{P}} \qquad \text{Eqn.(2.8)}$$

By the adding and subtracting a term $\bar{P} \cdot \overline{PET}$ in the numerator, $\Delta\phi_i$ can be represented by the deviation of PET and P'. Since the long-term storage change $\Delta\bar{S} = 0$, $\Delta P'_i = (P_i - \Delta S_i) - (\bar{P} - \Delta\bar{S}) = P'_i - \bar{P}$, then Eqn.(2.8) becomes:

$$\Delta\phi_i = \frac{PET_i\bar{P} - \bar{P}\cdot\overline{PET} + \bar{P}\cdot\overline{PET} - \overline{PET}P_i'}{P_i'\bar{P}} = \frac{\bar{P}(PET_i - \overline{PET}) - \overline{PET}(P_i' - \bar{P})}{P_i'\bar{P}} = \frac{\bar{P}\Delta PET_i - \overline{PET}\Delta P_i'}{P_i'\bar{P}} = \frac{\Delta PET_i - \bar{\phi}\Delta P_i'}{P_i'} \qquad \text{Eqn.(2.9)}$$

The ET deviation ΔET_i can be expressed as:

$$\Delta ET_i = ET_i - \overline{ET} = P_i' F(\phi_i) - \bar{P}F(\bar{\phi}) \qquad \text{Eqn.(2.10)}$$

Substituting $F(\phi_i)$ from Eqn.(2.7) into Eqn.(2.10) yields:

$$\Delta ET_i = P_i'[F(\bar{\phi}) + F'(\bar{\phi})\Delta\phi_i] - \bar{P}F(\bar{\phi}) = F(\bar{\phi})\Delta P_i' + F'(\bar{\phi})P_i'\Delta\phi_i \qquad \text{Eqn.(2.11)}$$

Substituting $\Delta\phi_i$ from Eqn.(2.9) into Eqn.(2.11) and cancelling $P'_i\Delta\phi_i$ yields:

$$\Delta ET_i = F(\bar{\phi})\Delta P_i' + F'(\bar{\phi})(\Delta PET_i - \phi_i \Delta P_i') = \Delta P_i'[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})] + \Delta PET_i F'(\bar{\phi})$$
$$= \Delta P_i[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})] - \Delta S_i[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})] + \Delta PET_i F'(\bar{\phi}) \qquad \text{Eqn.(2.12)}$$

Eqn.(2.12) expresses ET deviation in terms of deviation of P, PET, ΔS and long-term arid index $\bar{\phi}$.

The unbiased sample variance of ET is defined as

$$\sigma_{ET}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(ET_i - \overline{ET})^2 = \frac{1}{N-1}\sum_{i=1}^{N}\Delta ET_i{}^2 \qquad \text{Eqn.(2.13)}$$

where N is the sample size. Taking square of Eqn.(2.12), summing over N sample and scaled by N-1, the sample variance of ET is:

$$\sigma_{ET}^2 = [F(\bar{\phi}) - F'(\bar{\phi})\bar{\phi}]^2\sigma_P^2 + [F(\bar{\phi}) - F'(\bar{\phi})\bar{\phi}]^2\sigma_{\Delta S}^2 + [F'(\bar{\phi})]^2\sigma_{PET}^2 + 2[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})]F'(\bar{\phi})cov(P, PET) - 2[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})]^2 cov(P, \Delta S) - 2[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})]F'(\bar{\phi})cov(PET, \Delta S) \qquad \text{Eqn.(2.14)}$$

Or:

$$\sigma_{ET}^2 = w_P\sigma_P^2 + w_{\Delta S}\sigma_{\Delta S}^2 + w_{PET}\sigma_{PET}^2 + w_{P,PET}cov(P, PET) + w_{P,\Delta S}cov(P, \Delta S) + w_{PET,\Delta S}cov(PET, \Delta S) \qquad \text{Eqn.(2.15)}$$

where the terms in Eqn.(2.14) before the variances/covariances can be expressed as weighting functions related to long-term arid index $\bar{\phi}$ and catchment characteristics parameter ϖ:

$$w_P = [F(\bar{\phi}) - F'(\bar{\phi})\bar{\phi}]^2 \qquad \text{Eqn.(2.16.a)}$$

$$w_{\Delta S} = [F(\bar{\phi}) - F'(\bar{\phi})\bar{\phi}]^2 \qquad \text{Eqn.(2.16.b)}$$

$$w_{PET} = [F'(\bar{\phi})]^2 \qquad \text{Eqn.(2.16.c)}$$

$$w_{P,PET} = 2[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})]F'(\bar{\phi}) \qquad \text{Eqn.(2.16.d)}$$

$$w_{P,\Delta S} = -2[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})]^2 \qquad \text{Eqn.(2.16.e)}$$

$$w_{PET,\Delta S} = -2[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})]F'(\bar{\phi}) \qquad \text{Eqn.(2.16.f)}$$

Thus the total variance of ET is decomposed into the variances/covariances of P, PET and ΔS. The sources of ET variance include the variance of climate forcing (i.e., σ_P and σ_{PET}), the coincidence of water and energy supply $cov(P, PET)$ (e.g., seasonality) and the catchment's response to climate forcing (i.e., the variance and covariance terms associated with ΔS). Therefore, catchment ET variance depends on both the mean and variance of hydro-climate variables. The contributions from these variance/ covariance terms vary by climate and catchment condition, as captured by the weighting functions in Eqn.(2.16).

The ET variance also depends on the scale of interval ΔT_i. Ideally, Eqn.(2.15) can be used to assess ET variance across any temporal scale if the variances/covariances terms can be calculated from the scale. In this study, ΔT_i is chosen at annual and monthly scales, representing inter-annual and intra-annual variance, respectively.

Figure 2.2 shows the weighting functions with respect to long-term atmosphere water supply (i.e., $\bar{P}$) and energy supply (i.e., $\overline{PET}$). The weighting functions can also be plotted with respect to $\bar{\phi}$, as shown in Figure 2.4. Since the long-term arid index $\bar{\phi}$ of a point on the $(\bar{P}, \overline{PET})$ plane is the slope of the line across that point and origin, the derivative $F'(\bar{\phi})$ on the $(\bar{P}, \overline{PET})$ plane is the directional derivative along the $\bar{\phi}$ contour line. The following analysis on weighting functions is mainly based on $(\bar{P}, \overline{PET})$ plane rather than the $\bar{\phi}$ axis for two reasons. First, the arid index $\bar{\phi}$ is not symmetric for the humid climate condition (i.e., $\bar{\phi} \in (0,1)$) and the arid climate condition (i.e., $\bar{\phi} \in (1, \infty)$). On the $(\bar{P}, \overline{PET})$ plane, arid and humid conditions are symmetric and separated by the 1:1 line. Second, the effect of storage change can be better assessed on the $(\bar{P}, \overline{PET})$ plane, which captures both the ratio and magnitude of $\bar{P}$ and $\overline{PET}$, while the arid index $\bar{\phi}$ only represents the ratio of the two variables. For example, two points along the 1:1 line on the $(\bar{P}, \overline{PET})$ plane have the same arid index (i.e., $\bar{\phi} = 1$), while the absolute magnitude of $\bar{P}$ and $\overline{PET}$ at the point located at the southwest corner is smaller than that at the northeast corner. The magnitudes of $\bar{P}$ and $\overline{PET}$ are important to assess the catchment's response to climate forcing, since catchment has limited storage capacity.

Before any further discussion, it is worth providing a summary of this section. We first discussed how to incorporate a water balance over a short time interval into the Budyko Hypothesis. The extended Budyko equation specifies the catchment system state in hydro-climatic space; each point on this plane represents a system state given the energy supply and water availability during a time interval. From a statistical perspective, the first moment of the ET scatter points over various states (associated with various time points) describes the ET long-term average, and the second moment represents the variance of ET at a specified time scale. Based on the Taylor series expansion of the extended Budyko equation, we decompose ET variance into various components from climate and catchment storage change. In the following, we address the weighting functions.

2.2.3 Climatic control on ET variance

The weighting function of ET variance from P variance under various climate conditions is shown if Figure 2.2a. Thereafter, climate condition is qualitatively expressed as "dry" or "wet" to refer the abundance of P, as "hot" or "cold" to denote the level of PET, and "moderate" to represent condition where P and PET has comparable magnitudes. Given $\overline{PET}$, the value of w_P decreases as the catchment becomes wetter (i.e., $\bar{P}$ increases) and vice versa. Given $\bar{P}$, w_P increases as the catchment becomes hotter (i.e., $\overline{PET}$ increases) and vice versa. The contribution of P variance to ET variance is significant under the dry-hot climate (i.e., low P and high PET) and becomes trivial under the humid and cool climate (i.e., high P and low PET). w_P then represents the climate control on the contribution of P variance to ET variance. For example, two catchments may have the same σ_P, but the ET variance in a humid-cool catchment is less affected by P variance than that in a dry-hot catchment. As shown by numerical experiments from *Fatichi and Ivanov* [2014], ET inter-annual variation is more sensitive to precipitation fluctuation in water-limited environments than other conditions.

Figure 2.2b shows the contribution of PET variance to ET variance under various climate conditions. w_{PET} is close to one under the humid-cool climate and approaches to zero under the dry-hot climate. As indicated by Eqn.(2.17) and Eqn.(2.19), P variance and PET variance controls ET variance in basins located in the lower-right and upper-left region of $(\bar{P}, \overline{PET})$ plane, respectively. Thus, the mean climate condition determines the contribution of climatic variance to ET variance.

$w_{PET,P}$, the contribution of ET variance from the covariance of PET and P, is shown in Figure 2.2c. The covariance of PET and P indicates the temporal coincidence between water and energy supply, such as the phasing between storm season and warm season or the concurrence of dry periods and heat waves. Note that the covariance can be either positive or negative. When $cov(P, PET) > 0$, P changes in-phase with PET (i.e., a wet period comes together with a hot period) and the covariance increases ET variance. When $cov(P, PET) < 0$, P evolves out-of-phase with PET (e.g., a wet period coincides with a cool period) and the covariance reduces ET variance. For example at the seasonal scale when the rainfall season is out-of-phase with a warm season, ET is then limited by energy supply although there is abundant water supply, or vice versa. As a result, this "out-of-phase" seasonality pattern dampens ET variance. The weight is large along the 1:1 line on the $(\bar{P}, \overline{PET})$ plane, which means that the contribution of the covariance between PET and P to ET variance is important under conditions where rainfall and energy supply have comparative magnitudes.

Figure 2.2d shows the weighting function related to catchment storage change. The contribution of ΔS variance to ET variance is similar to that of w_P, since P represents the water supply from the atmosphere and ΔS represents the water supply from the catchment. The range for $w_{P,\Delta S}$ is from -2 to 0 in Figure 2.2e, which indicates that the response of catchment storage change to P will significantly affect ET variance. For example, in a dry year, catchment storage change decreases with reduced rainfall (since farmers pump groundwater for irrigation). Thus $cov(P, \Delta S) > 0$ and $w_{P,\Delta S} < 0$ yield $w_{P,\Delta S} cov(P, \Delta S) < 0$. ET variance is dampened by the interaction between natural water supply and human water use. Similarly, ET variance is dampened by the interaction between energy supply and catchment water redistribution as shown in Figure 2.2f, and the buffer effect is significant under climate conditions when P and PET have similar magnitudes.

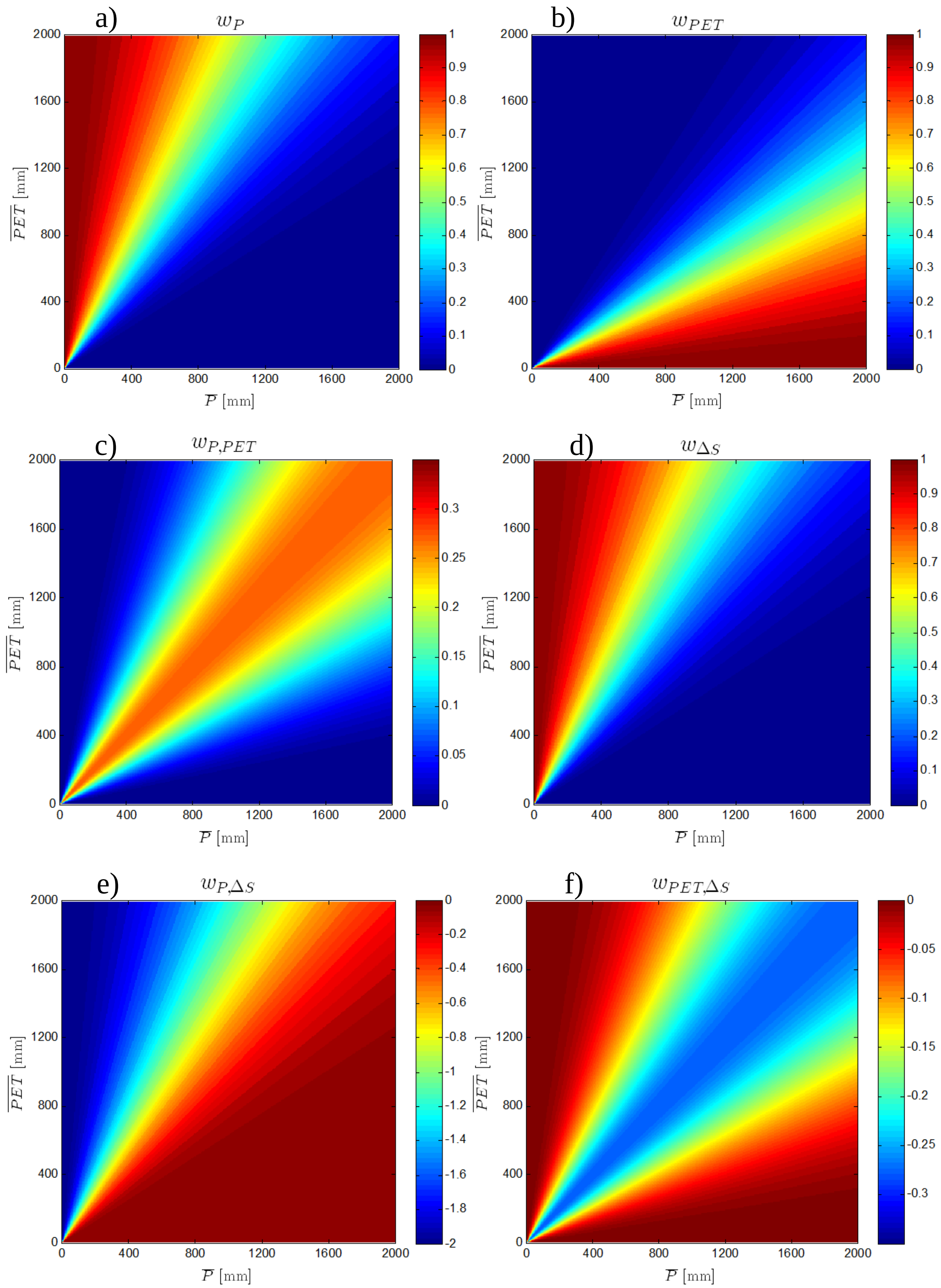


Figure 2.2. Weighting functions for ET variance from: a) variance of *P*; b) variance of *PET*; c) covariance of *PET* and *P*; d) variance of *ΔS*; e) covariance of *P* and *ΔS*; f) covariance of *PET* and *ΔS*.

2.2.4 ET variance under various climate conditions and storage capacities

As discussed above, the weights indicate the contribution from different components to ET variance. Under some climate conditions, there exist some dominant factors on ET variance. Thus the ET variance in Eqn.(2.15) can be simplified accordingly. This section will discuss how to obtain simpler expressions for ET variance based on assumptions about climate conditions.

In hot-dry regions (i.e., the upper-left part of the $(\bar{P}, \overline{PET})$ plane), we can find that weights associated with PET (i.e., $w_{PET}, w_{PET,\Delta S}, w_{PET,P}$) are negligible. These negligible weights are associated with the conditions under which the hydrological cycle is dominated by water supply, and then the ET variance is not sensitive to the variance of energy supply. Furthermore, if a catchment has a limited storage capacity (i.e., shallow soil profile, small aquifer porosity and thickness, or limited engineering storage infrastructure), the effect of storage change is also limited, which will lead to the following cases.

Case 1: With the hot-dry climate and limited catchment storage capacity, Eqn.(2.15) can be simplified as:

$$\sigma_{ET}^2 = w_P \sigma_P^2 \qquad \text{Eqn.(2.17)}$$

which is the same as that provided by *Koster and Suarez* [1999], i.e., the variance of ET only comes from the variance of P. Qualitatively, the condition of Eqn.(2.17) corresponds to domain Ω_{I} in Figure 2.3, where $\bar{\phi}$ is large and the magnitude of $\overline{PET}$ and $\bar{P}$ are also large (compared to catchment storage capacity).

Case 2: In a hot-dry climate and the catchment storage change is large enough to redistribute water and affect ET flux, Eqn.(2.15) becomes:

$$\sigma_{ET}^2 = w_P \sigma_P^2 + w_{\Delta S} \sigma_{\Delta S}^2 + w_{P,\Delta S} cov(P, \Delta S) \qquad \text{Eqn.(2.18)}$$

Compared to Case 1, Eqn.(2.18) includes the variance of ΔS and its interaction with P, corresponding to domain Ω_{II} as shown in Figure 2.3, where $\bar{\phi}$ is large and the magnitude of $\overline{PET}$ and $\bar{P}$ is small and the catchment has sufficiently large storage capacity to redistribute water. For example, the groundwater table declines during a drought event to sustain vegetation growth by natural processes (e.g., deep roots uptake) or by human activity (e.g., groundwater pumping for irrigation).

In cool-wet regions (i.e., lower-right part of $(\bar{P}, \overline{PET})$ plane), we can find that weights associated with P (i.e., $w_P, w_{P,\Delta S}, w_{P,PET}$) are negligible. In such a climate, the hydrological cycle is dominated by energy supply and the variance of water supply is not significant to ET variance. Different from the hot-dry regions, the weights associated with ΔS (i.e., $w_{\Delta S}, w_{PET,\Delta S}$) are also negligible under the cool-wet climate. This implies that the catchment storage can effectively redistribute water but not surface energy. Accordingly, there are two cases as follows:

Case 3: Under the cool-wet climate, the hydrological cycle is limited in energy supply, and the variance of P is negligible. In a catchment with limited storage capacity, we can further drop the terms associated with ΔS and Eqn.(2.15) can be reduced to:

$$\sigma_{ET}^2 = w_{PET}\sigma_{PET}^2 \qquad \text{Eqn.(2.19)}$$

i.e., the variance of ET only comes from the variance of PET under the cool-wet climate, corresponding to domain Ω_{III} shown in Figure 2.3.

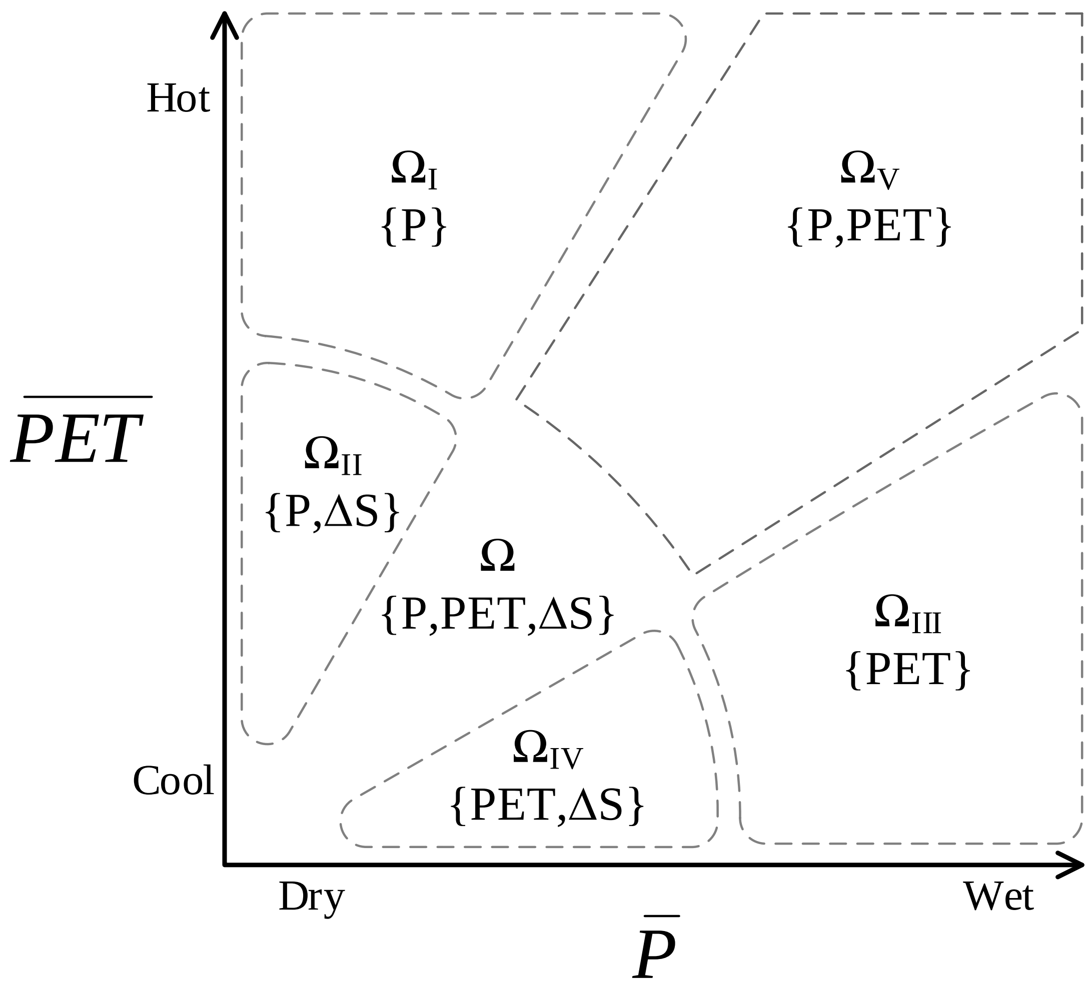


Figure 2.3. Schematic plot for controlling factors of ET variance under different climate conditions and catchment storage capacity. The variables in the curly bracketsindicates the controlling factor for ET variance in subdomain Ω_i.

Case 4: Under the cool-wet climate with significant storage capacity, P can be stored for ET until energy supply becomes large. For example, snow pack accumulates in cold periods and melts in warm periods. Thus Eqn.(2.15) is reduced to the following form:

$$\sigma_{ET}^2 = w_{\Delta S}\sigma_{\Delta S}^2 + w_{PET}\sigma_{PET}^2 + w_{PET,\Delta S}cov(PET, \Delta S) \qquad \text{Eqn.(2.20)}$$

Eqn.(2.20) shows that the variance of PET and ΔS together affects ET variance in cool-wet regions, corresponding to domain Ω_{IV} shown in Figure 2.3.

Under a moderate climate, the magnitude of $\bar{P}$ and $\overline{PET}$ is approximately the same, and they both contribute to ET variance. If the magnitudes of $\bar{P}$ and $\overline{PET}$ are large compared to catchment storage capacity, the catchment storage variance is then negligible. With such conditions, the terms associated with catchment storage change (i.e., $w_{\Delta S}, w_{P,\Delta S}, w_{PET,\Delta S}$) can be dropped:

$$\sigma_{ET}^2 = w_P\sigma_P^2 + w_{PET}\sigma_{PET}^2 + w_{P,PET}cov(P, PET) \qquad \text{Eqn.(2.21)}$$

Eqn.(2.21) shows that variance of PET and ET jointly affects ET variance in catchments with limited storage capacity under a moderate climate condition, corresponding to domain Ω_V shown in Figure 2.3.

In summary, the total ET variance is decomposed into different components, and the complete forms of these components are shown in Eqn.(2.15). Simpler forms are obtained under specific climate conditions and catchment storage capacities, as shown in Eqn.(2.17) to Eqn.(2.21).

2.3 Case study

We assess the annual and monthly scale water balance in the Murray-Darling Basin as a case study to illustrate the analytical framework presented above. Detailed ET variance analysis in more basins with different climate conditions is under work. The monthly time series of *P*, ET, *PET*, and *ΔS* during 1984-2006 in the Murray-Darling Basin is obtained from multiple sources and the errors with each item are handled through a data assimilation procedure based on constrained Kalman Filter [*Pan et al.*, 2012]. The annual time series is aggregated from the monthly dataset, as shown in Figure 2.4a. At the annual scale, catchment storage fluctuates according to the annual rainfall. In dry years, catchment storage decreases to provide a complementary water source for ET. During dry years such as 1994, 2001 and 2002, ET, with the complementary catchment storage supply, is larger than the atmospheric water supply. In wet years such as 1988, catchment storage recovers due to rainfall. At the monthly scale, catchment storage change is significant, as shown in the monthly average plot in Figure 2.4b. ET has a clearly opposite pattern with the catchment storage change. Catchment storage recovers from rainfall during May to July when ET is low and decreases from September to November to supply ET.

The Budyko curve for the annual and monthly dataset is plotted in Figure 2.4c and 4d, respectively. ET and *PET* are scaled by rainfall (P) and total water availability (P') for comparison. For the annual scale in Figure 2.4c, the points generally follow the Budyko curve. The points scaled by P (i.e., without ΔS) are more scattered than the points scaled by P' (i.e., with ΔS). It is noted that there are four points violating the "water limit" if catchment storage is neglected. At the monthly scale, without considering storage change, the points follow an approximately linear trend and there are many points violating the "water limit". After the monthly data are scaled by the total available water (P'), all points follow the extended Budyko curve as expressed by Eqn.(2.5).

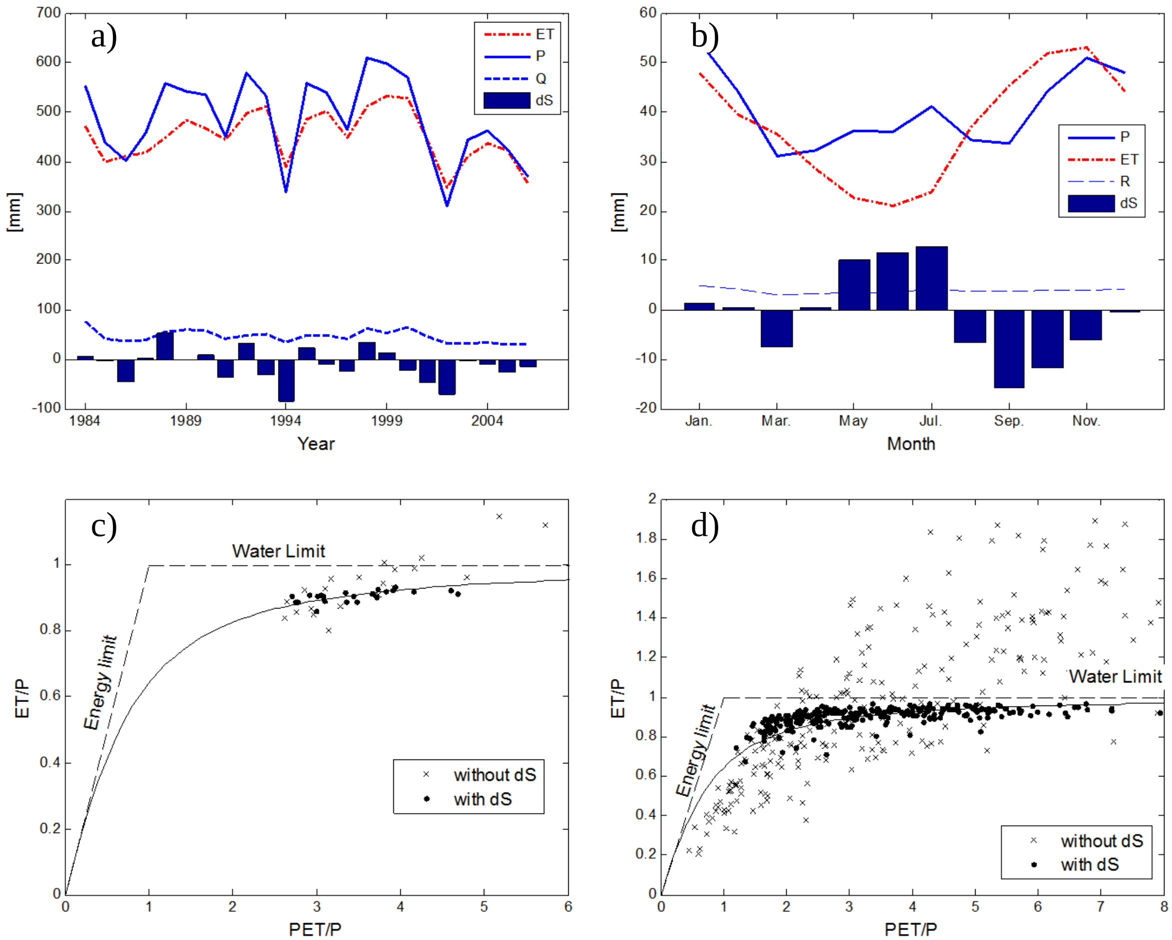


Figure 2.4. a) Annual and b) monthly water balance fluxes for Murray-Darling Basin from 1984-2006. Plots for arid index vs. evaporation index scaled by atmospheric water supply (neglecting ΔS) and total water availability (including ΔS) for c) annual and d) monthly series.

ET inter annual variance calculated by Eqn.(2.15) is 2567 mm^2, which is close to the observed one at 2682 mm^2, with a relative error of 4%. The contributions to the ET variance from different components are plotted in Figure 2.5a. The climate control on ET variance mainly come from P, and the contribution from *PET* variance is negligible, since the basin is under arid climate (i.e., $\bar{\phi}$=3.48). Specifically the variance contribution from *P* is 4960 mm^2 by Eqn.(2.1), about twice large as the observed ET variance. Even in this arid basin, neglecting the effect of storage and estimating ET variance from *P* alone would over-estimate the ET variance.

At the monthly scale, the intra annual ET variance estimated from Eqn.(2.15) is 167 mm^2, which is close to the observed one at 171 mm^2. Both *P* and *ΔS* variance contribute to ET variance; while *P* and *ΔS* covariance reduces the ET variance, as shown in Figure 2.5b. Again, it shows the importance of incorporating storage change to estimate ET variance, since the estimation from *P* only is highly over-estimated.

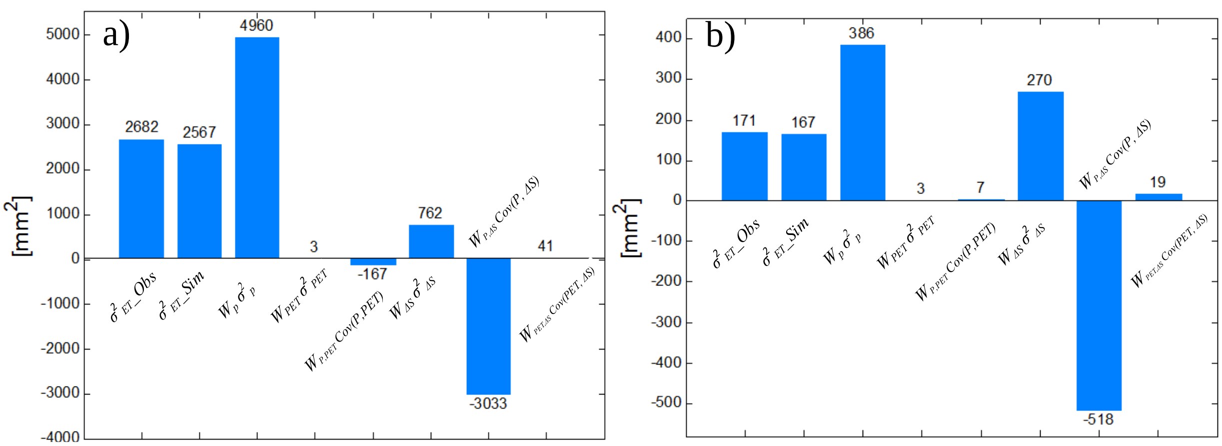


Figure 2.5. ET a) inter-annual variance and b) intra-annual variance from observed data, calculated value from Eqn.(2.15) and contribution from each components.

2.4 Discussion

2.4.1 Impact of catchment characteristics on ET variance

The catchment characteristics are represented by parameter ϖ in Eqn.(2.4). The parameter is introduced by obtaining an analytical solution for Budyko Hypothesis and has no specific physical meaning (Fu, 1981). From Eqn.(2.4), one can see that ϖ balances the partition between ET and *Q*: a large ϖ value leads to higher ET and lower *Q*, and vice versa. Thus a catchment with limited storage capacity, less vegetation coverage, and a steep hill slope would expect to have a smaller ϖ value. Moreover, some studies obtained empirical relationships for ϖ to relate ϖ to various catchment characteristics. For example, *Yang et al.* [2007] related ϖ to infiltration capacity, soil water storage capacity and average hill slope. *Li et al.* [2013] estimated ϖ assuming it has a linear relationship with the normalized difference vegetation index (NDVI).

The effect of ϖ on the weights in Eqn.(2.16) is plotted with respect to the arid index ($\bar{\phi}$) for the convenience of comparison in Figure 2.6. For the weighting function of *P* in Figure 2.6a, large ϖ results in more contribution from *P* to ET variance than small ϖ, and the difference is especially significant in arid regions. For example, grassland ET under arid climate is more sensitive to precipitation fluctuations than under humid climate [*Y Yang et al.*, 2008]. Similar patterns can be found with *ΔS* in Figure 2.6d, since large ϖ implies that vegetation can use more catchment storage. For the weighting function of *PET*, large ϖ results in more contribution from *PET*, especially in humid climates. The contribution from the covariance of energy and water supply (i.e., $w_{P,\Delta S}$ and $w_{PET,\Delta S}$) is significant under moderate climate (i.e., $\bar{\phi}$ around 1) for large ϖ, which implies that vegetation contributes to ET variance. However, the effect of the vegetation type on the catchment scale ET still remains unclear. *Williams et al.* [2012] assessed ET measurements across global fluxtower network and found that grasslands on average have higher ET than forest. This contrasts the convention that forest has higher ET due to higher canopy interception, deeper and more extensive root system and higher leaf area than grassland. Further work has been suggested to relate catchment physical properties to parameter ϖ in order to fully understand the scale-dependence role of vegetation in water and energy cycle [*Brooks et al.*, 2011].

2.4.2 Stationarity and ET variance change

The ET variance is derived based on the assumption that hydro-climatic time series are stationary. Here we assume stationarity in a wide sense, that is, the first moment and second moment of the hydro-climatic sequence is irrespective to time. The theoretical framework developed in this study can be applied to assessing the ET variance change between two periods. It is acknowledged that the hydro-climatic processes are interdependent and coevolve with each other at a long-time scale. A practical question to ask is: what implications can we draw from the framework to assess the ET variance change under catchment management and climate adaptation?

The nonlinear and concave form of the Budyko equation exhibits similar properties as utility functions or vegetation productivity functions [*Hsu et al.*, 2012; *Fatichi and Ivanov*, 2014]. Due to the nonlinear form, the change of climate forcing (e.g., *PET* or *P*) variance has different impacts on ET variance under various climates. Following Eqn.(2.15), factors affecting ET variance can be categorized into three situations: First, climate conditions directly affect ET variance via the variance of *P* and *PET* and indirectly by the weighting functions via the mean of *P* and *PET*. Second, natural processes (e.g., glacier and aquifer dynamics) and anthropogenic interferences (e.g., groundwater utilization and irrigation) determine the covariance of catchment storage to climatic variables. Third, catchment properties such as vegetation pattern, soil property and land use, affect the weighting functions through

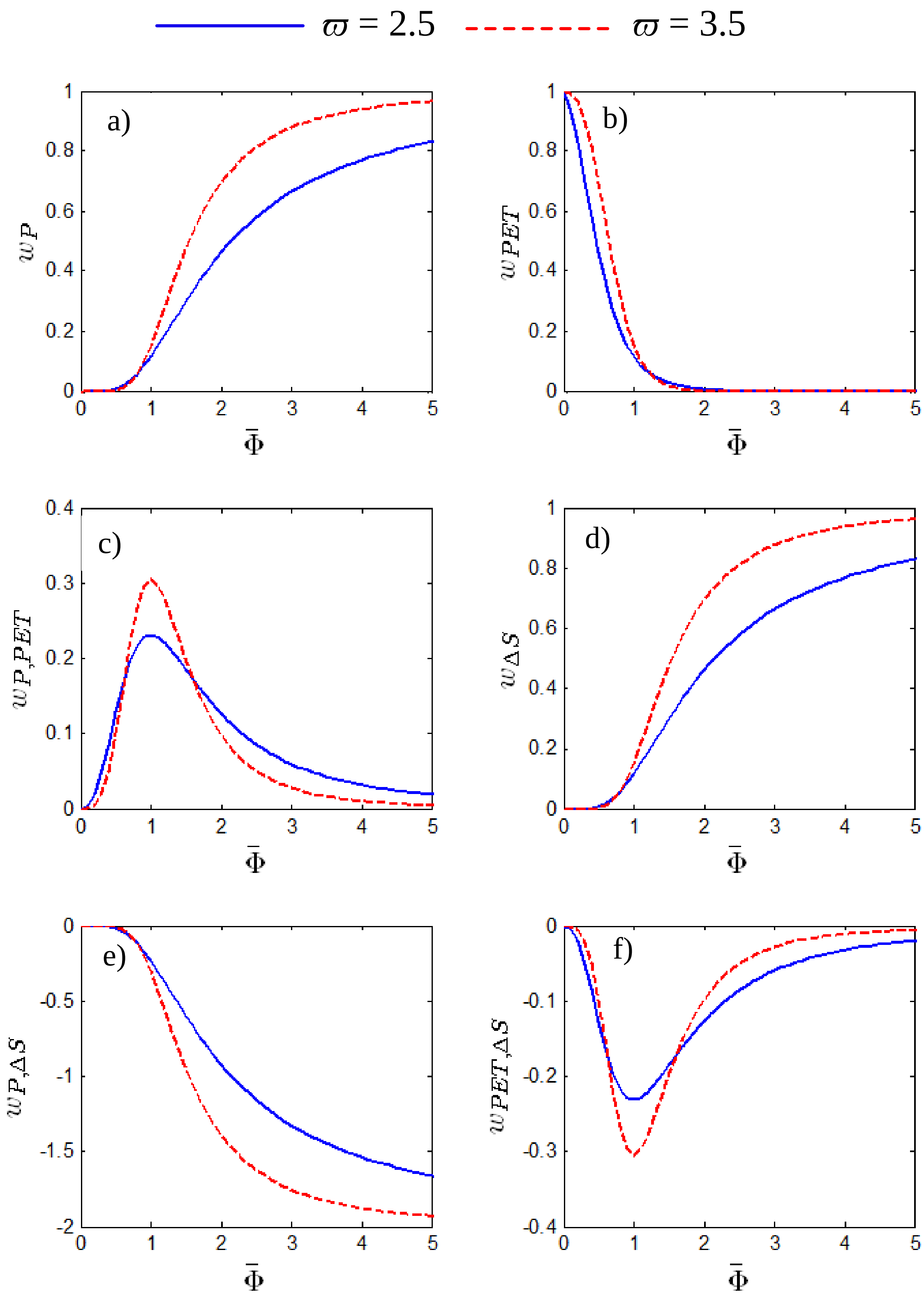


Figure 2.6. Effect of catchment characteristics parameter on weighting functions.

catchment characteristics parameter ϖ. Thus the ET variance change can be quantified by changes of the three categories listed above.

2.4.3 Catchment storage change responding to climate change and human impacts

By incorporating the water balance to the Budyko Hypothesis, this study highlights the important role of catchment storage change in water and energy cycle dynamics. As shown above, ET variance is affected by catchment storage change via the covariance terms in Eqn.(2.15). The vegetation use of catchment storage, in the form of soil moisture, groundwater, surface reservoir , has been extensively studied in ecohydrology [*Eagleson*, 2002]. However, the process of snow accumulating and melting has rarely been studied in the context of the Budyko curve. A recent study by *Lute and Abatzoglou* [2014] shows the importance of extreme snowfall events in shaping the inter-annual variability in water resources. The retreat of glaciers due to climate change would result in the loss of capacity to “carry-over” the effect of winter precipitation for summer ET. Thus the response of snow dynamics to climate should be incorporated to improve the ET variance assessment in future studies. Human impact is another important factor to ET variance. The amount and timing of irrigation is closely related to climate fluctuation [*Vico and Porporato*, 2013]. The human adaptation to climate could be characterized by the use of catchment storage to couple with climate variables via the covariance of catchment storage change and climate (i.e., $cov(PET, \Delta S)$ and $cov(P, \Delta S)$). In this way, this framework quantifies human impacts on ET variance.

2.5 Conclusions

This study extends the analytical framework by *Koster and Suarez* [1999] to assess the temporal variance of ET based on water balance and the Budyko Hypothesis. The framework incorporates the response of basins and human activities to climatic variance so that it can be used to quantify the key controlling factors on catchment ET variance. The ET variance is decomposed variances/covariances from P, PET and ΔS, and each component is scaled by a weighting function, which is a function of long-term climate condition and catchment characteristics. Thus the framework can attribute ET variance to both the mean and variance of climate variables.

The sources of ET variance are identified under various climate conditions and human activities. ET variance is dominated by P variance under the hot-dry climate; by PET variance under the cold-wet climate; by both P and PET variance under moderate climate conditions. Moreover the “in-phase” of P and PET under moderate climate conditions increases ET variance via the coincidence of water and energy supply for ET. Besides climate conditions, a catchment’s response to climate fluctuation also shapes ET variance. Vegetation, human adaptation to climate, and snow and aquifer storage dynamics can be represented by the covariance terms to quantify their impact on ET variance. The framework can be applied for assessing ET variance over various temporal scales, which is illustrated by the inter- and intra-annual variance of ET in the Murray-Darling Basin. It is shown that catchment storage change plays an important role to buffer ET temporal variance in this arid basin, and overlooking storage change and estimating ET variance only from P variance yields over-estimated ET variance.

The ET variance decomposition framework is derived from Budyko hypothesis with some extensions and assumptions. The original Budyko curve captures ET flux for long-term average, while the decomposition framework is based on each time period (e.g., monthly, seasonal and annual). Therefore, it is essential to incorporate terrestrial water storage (or to satisfy the water balance) for the short time scale analysis. Further, we assume the terrestrial water storage change is caused by vertical fluxes (e.g., recharge and pumping). If lateral fluxes exist, such as inter-basin water transfer and regional groundwater flow, these lateral fluxes should be added or subtracted to get the terrestrial water storage change due to vertical fluxes. Both Budyko curve and the ET variance decomposition framework quantify the watershed flux signatures over a long period of time, with the former focusing on the mean (the first-order statistics) and the latter focusing on the variance (the second-order statistics). The climate conditions (i.e., aridity index) play a similar role in ET variance decomposition framework as that in Budyko curve. The long-term climate conditions determine how each variance/co-variance component contribute to total ET variance, as quantified by the weighting factors.

The assessment of ET variance in catchments with various climate conditions will be conducted in future work. The framework can be applied to interpreting the historical ET variance change and to understanding the alteration of ET variance with future climate. The insights obtained in this study can be used to provide improved understanding of the stochasticity of hydro-climatic systems, more accurate hydrologic prediction in ungauged basins, and guidelines on adaptation to climate change.

CHAPTER 3

INTER- AND INTRA-ANNUAL ET VARIANCE IN GLOBAL BASINS

In this chapter, ETVARD derived in Chapter 2 is applied to 32 basins that are distributed globally and represent a wide range of climate conditions. In this way, we quantify the climate and terrestrial water storage controls on ET variance. We also examine how main components vary at inter- and intra-annual scales.

3.1 *Introduction*

Koster and Suarez [1999] made the first attempt to derive ET inter-annual variance from precipitation variance and compare the theoretical framework with climate model simulations. Later, it was found that other climatic factors (e.g., net radiation) and their variation at smaller time scales (e.g., seasonality) also play important roles in ET variance [*Koster et al.*, 2006]. Since ET is the output of climatic forcings filtered by hydrologic systems, hydrologic state variables (e.g., terrestrial storage, such as glacier, snow, soil moisture, groundwater and reservoirs) also contribute to ET through both natural processes and anthropogenic interferences such as snow thawing-melting [*Lettenmaier and Milly*, 2009], vegetation root up-take [*Nepstad et al.*, 1994] and agricultural irrigation [*Condon and Maxwell*, 2014]. Our recent study [*Zeng and Cai*, 2015] extended the theoretical framework of *Koster and Suarez* [1999] by accounting for the effect of terrestrial storage and derived the climatic-hydrologic controls on ET variance. Using the new theoretical framework, this paper investigates the ET temporal variance pattern in large basins with various climatic conditions. Specifically, this study conducts benchmark assessment of ET variance at both inter- and intra- annual scales in thirty-two big river basins with the best available data.

Issues to address in this study include: 1) From where do the sources of ET variance originate, and which factors dominate ET variance? Assessment of the similarities and differences in the sources of ET variance over various regions may help capture the main factors governing catchment hydro-climatic variation. 2) What is the role of terrestrial storage change in ET variance under arid/humid climates? It is assumed that the terrestrial storage control on ET variance would be different under various climates, since hydrologic responses to the variability in climatic forcings involve different processes, such as snow thawing-melting and vegetation growth, under different climate conditions. 3) How is ET variance controlled by climate and terrestrial storage at different temporal scales, especially at annual and monthly scales? A better knowledge of these issues would help hydrologists, climate scientists, and water resources managers better predict the effects of climate change and adapt to a changing environment.

3.2 Methods and data

3.2.1 ET variance decomposition framework

In Chapter 3, we extended the ET variance assessment framework by *Koster and Suarez* [1999] which was based on the Budyko's hypothesis on climatic primary control on ET. However, terrestrial water storage is missing in this picture. Since the climatic control on ET in Budyko's hypothesis is based on a long-term equilibrium, in order to account for the effect of terrestrial storage change on ET we introduced total water availability (*P'*), the sum of precipitation (*P*) and terrestrial storage change (*ΔS*), to represent the available water source for ET. The original Budyko hypothesis is tested with the long-term average hydroclimatic variables with basins around the world to capture the similarity among the basins. However, the extension here applies the Budyko hypothesis to temporal analysis of a single basin. That is, during each time interval *i*, the two items $\left(\frac{\mathrm{PET}}{\mathrm{P'}}\right)_{\mathrm{i}}$ and $\left(\frac{\mathrm{ET}}{\mathrm{P'}}\right)_{\mathrm{i}}$ fit the Budyko hypothesis (i.e., the boundary conditions specified for dry/wet conditions [*Zhang et al.*, 2004; *H Yang et al.*, 2008]). The extension from the primary spatial analysis to temporal analysis follows the idea of hydrologic similarity and space-time symmetry [*Sivapalan et al.*, 2011].

As detailed in [*Zeng and Cai*, 2015], ET variance is decomposed into variance/covariance of precipitation, energy supply (represented by potential ET, *PET*) *and* terrestrial storage change*:*

$$\sigma_{ET}^2 = w_P\sigma_P^2 + w_{PET}\sigma_{PET}^2 + w_{\Delta S}\sigma_{\Delta S}^2 + w_{P,PET}cov_{P,PET} + w_{P,\Delta S}cov_{P,\Delta S} + w_{PET,\Delta S}cov_{PET,\Delta S} \qquad \text{Eqn.(3.1)}$$

where σ represents the standard deviation, cov represents the covariance, and w represents the weighting factors, which quantify the contribution from different variance/covariance sources to ET variance and can be analytically calculated from the aridity index ($\bar{\phi} = \overline{PET}/\bar{P}$) [*Zeng and Cai*, 2015]:

$$w_P = [F(\bar{\phi}) - F'(\bar{\phi})\bar{\phi}]^2 \qquad \text{Eqn.(3.2.a)}$$

$$w_{\Delta s} = [F(\bar{\phi}) - F'(\bar{\phi})\bar{\phi}]^2 \qquad \text{Eqn.(3.2.b)}$$

$$w_{PET} = [F'(\bar{\phi})]^2 \qquad \text{Eqn.(3.2.c)}$$

$$w_{P,PET} = 2[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})]F'(\bar{\phi}) \qquad \text{Eqn.(3.2.d)}$$

$$w_{P,\Delta S} = -2[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})]^2 \qquad \text{Eqn.(3.2.e)}$$

$$w_{PET,\Delta S} = -2[F(\bar{\phi}) - \bar{\phi}F'(\bar{\phi})]F'(\bar{\phi}) \qquad \text{Eqn.(3.2.f)}$$

where $F(\bar{\phi})$ and $F'(\bar{\phi})$ denote the Budyko equation and it first order derivative, respectively. We use Fu's equation [*Fu*, 1981; *Zhang et al.*, 2004] to depict the Budyko curve:

$$\frac{\overline{ET}}{\bar{P}} = F(\bar{\phi}) = F\left(\frac{\overline{PET}}{\bar{P}}\right) = 1 + \frac{\overline{PET}}{\bar{P}} - \left[1 + \left(\frac{\overline{PET}}{\bar{P}}\right)^{\varpi}\right]^{1/\varpi} \qquad \text{Eqn.(3.3)}$$

The value of w_P, $w_{\Delta S}$ and w_{PET} range between 0 and 1, $w_{P,\Delta S}$ between -2 and 0; $w_{P,PET}$ between 0 and 0.35; and $w_{PET,\Delta S}$ between -0.35 and 0 as identified in [*Zeng and Cai*, 2015]. The parameter ϖ is regressed for individual basins with annual data.

Furthermore, the effect of terrestrial storage change on ET variance can be analyzed for arid ($\bar{\phi} > 1$) and humid ($\bar{\phi} < 1$) climate, respectively. Since in arid climates the weighting functions associated with *PET* (i.e., w_{PET}, $w_{P,PET}$ and $w_{PET,\Delta S}$) are negligible which means the contribution from *PET* fluctuation to ET variance would be small, we have

$$\sigma_{ET}^2 = w_P \sigma_P^2 + w_{\Delta S} \sigma_{\Delta S}^2 + w_{P,\Delta S} cov(P, \Delta S). \qquad \text{Eqn.(3.4)}$$

If the terrestrial storage change is small and contributes little to ET variance in the arid climates, Eqn.(3.4) can be further reduced to:

$$\sigma_{ET}^2 = w_P \sigma_P^2 \qquad \text{Eqn.(3.5)}$$

where ET variance is explained solely by *P* variance. Note that Eqn.(3.5) is the equation derived by *Koster and Suarez* [1999].

In humid climates where weighting functions associated with *P* ($i.e., w_P$, $w_{P,PET}$ and $w_{P,\Delta S}$) are negligible, ET variance can be simplified as:

$$\sigma_{ET}^2 = w_{PET} \sigma_{PET}^2 + w_{\Delta S} \sigma_{\Delta S}^2 + w_{PET,\Delta S} cov(PET, \Delta S) \qquad \text{Eqn.(3.6)}$$

If terrestrial storage change is small and contributes little to ET variance in the humid climates, Eqn.(3.6) can be further reduced to:

$$\sigma_{ET}^2 = w_{PET} \sigma_{PET}^2 \qquad \text{Eqn.(3.7)}$$

where the ET variance is attributed to *PET*.

We can assess the effect of terrestrial storage change on ET variance by comparing ET variance estimated from Eqn.(3.4) and (5) for arid climates and Eqn.(3.6) and (7) for humid climates.

3.2.2 Data

The terrestrial water storage controls on ET variance have not been comprehensively examined before, mainly due to the lack of accurate observations of terrestrial water storage. The terrestrial water budget dataset used in this study was developed for thirty-two big river basins over the world during 1984-2006 by *Pan et al.* [2011]. This dataset is based on multiple sources, including *in situ* observations, remote sensing retrievals, land surface model simulations, and global reanalysis. Compared to ET estimation from a single source, this dataset has several advantages. Bias and errors from the various sources are compensated in a systematic way to achieve the best possible confidence by accounting for the observation network density, model accuracy and other factors. In addition, a constraint Kalman filter technique is applied to preserving the water budget for each month. Readers are referred to *Pan et al.* [2011] for a full description of data sources, data assimilation procedures and uncertainty quantification. This dataset has been previously applied to assessing the controlling factors of water and energy cycles within the context of the Budyko framework [*Li et al.*, 2013; *Xu et al.*, 2013].

The *PET* used in this study is calculated from the Penman equation. The meteorological data (e.g. wind speed, relative humidity, air temperature, incoming short-wave and long-wave radiation) are obtained from the Princeton University global forcing data [*Sheffield et al.*, 2006] and the Variable Infiltration Capacity land surface model simulation [*Sheffield and Wood*, 2007]. The daily *PET* is aggregated spatially for each of the 32 basins and temporally for the monthly and annual scale. Due to the uncertainty involved in the calculation of *PET*, which may affect the analysis result, we calculate *PET* using three additional methods (the Priesley-Taylor equation, the FAO Penman-Monteith equation and the Penman equation with zero energy flux). We find that the result from the ET decomposition framework is robust under the various *PET* calculation methods. The four calculation equations and detailed comparison of *PET* mean value, variance and its contribution to ET inter- and intra-annual variance can be found in the Appendix A.

Climate and storage data are available for river basins over a variety of geographic regions, as shown in Figure 3.4. These basins cover a wide climatic spectrum with aridity indices $\bar{\phi}$ ranging from 0.60 to 8.33. The catchment characteristics parameter of the Budyko curve is regressed from least squares error by fitting the Budyko curve. The ET variance calculated from the ET time series from data assimilation by *Pan et al.* [2011] is denoted by "assessed", and the ET variance calculated from Equation (1) is denoted as " prediction". The intra-annual variance is evaluated by the monthly data sequence, and inter-annual variance is calculated from annual aggregations of the monthly data.

3.3 *Results*

3.3.1 ET inter-annual variance sources

The ET inter-annual variances (represented by standard deviation σ_{ET} in mm) for the thirty-two basins as predicted by Eqn.(3.1) and from the assessed dataset are plotted in Figure 3.1a. In general, Eqn.(3.1) reasonably captures the ET variance at the annual scale. The r-squared and Nash-Sutcliffe coefficient between the two sets of estimates are 0.67 and 0.44, respectively, with an average error of 5.08 mm. It is noted that σ_{ET} in some basins is slightly underestimated by Eqn.(3.1) compared to the assessment, mainly due to the small sample size (i.e., the 23-year study period for each basin) used for the analysis.

The percentage contributions to ET inter-annual variance σ_{ET}^2 from each term in Eqn.(3.1) are shown in Figure 3.1b, where the basins are displayed from left to right according to the values of the long-

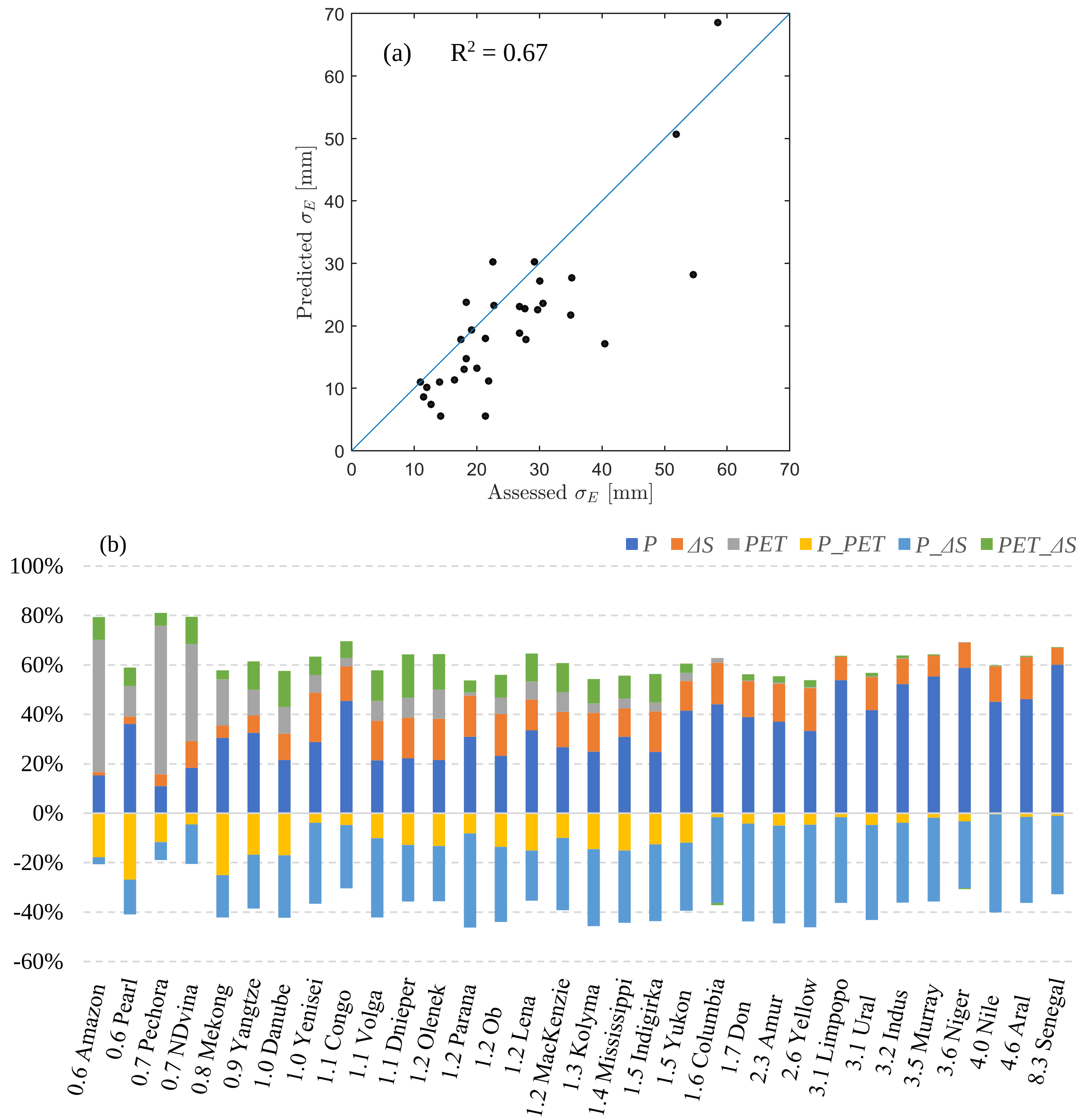


Figure 3.1. a) ET inter-annual standard deviation from assessed data and prediction by Eqn.(3.1). b) Percentage contribution to ET inter-annual variance from each component in Eqn.(3.1). Basins are listed from left to right with increasing aridity index, and the numbers in front of basin names are the aridity index.

term average aridity index $\bar{\phi}$. Note that we use σ_{ET}^2 instead of σ_{ET}, since the contribution from those covariance terms can be negative. Figure 3.1b shows that climate is the primary source for ET inter-annual variance, and the largest sources shift from *PET* to *P* with the increase of $\bar{\phi}$ (from humid to arid *basins*). For example, *PET* contributes more than 50% of the ET variance in the Amazon, Northern Dvina, and Pechora basins, all of which have small aridity indices (i.e., $\bar{\phi}$ *= 0.60, 0.74, and 0.63*, respectively). The result is consistent with the finding of *Karam and Bras* [2008], who found that ET in the Amazon is in-phase with the basin-averaged surface net radiation and concluded that Amazonian ET is prevalently limited by energy. In arid regions, such as in the Indus, Limpopo, Murray-Darling, Niger, and Senegal basins, where the aridity index $\bar{\phi} > 3$, *P* contributes more than half of the ET variance.

In addition to these climate variables, terrestrial storage change is also an important contributing factor to ET variance. The negative contribution from the term $w_{P,\Delta S}cov(P,\Delta S)$ in Figure 3.1b shows that ET variance is dampened, especially in arid and moderate climates. By examining the annual water budget in these thirty-two basins, it can be seen that *ΔS* follows *P* change (i.e., *ΔS* decreases during a dry year and recovers during a wet year), leading to a positive $cov(P,\Delta S)$. The negative contribution to ET variance results from a negative weighting function value, $w_{P,\Delta S}$. As a result, the complementary sources from the atmosphere (i.e., *P*) and catchment (i.e., *ΔS*) act jointly in dampening the ET inter-annual variance. This phenomenon is described in detail by *Wang et al*. [2009], who found that groundwater storage could be significant in buffering water balance inter-annual variance in catchments in Nebraska.

3.3.2 Terrestrial storage control on ET inter-annual variance

The effect of terrestrial storage change on ET inter-annual variance is assessed for arid and humid climates, respectively. ET variances of 25 arid basins (with $\bar{\phi} > 1$) calculated from Eqn.(3.4) (with considering *ΔS*) and Eqn.(3.5) (without considering *ΔS*) are compared with the assessed ET variance, as shown in Figure 3.3a. Both equations reasonably capture ET variance; the r-squared are 0.90 for Eqn.(3.4) (with considering *ΔS*) and 0.82 for Eqn.(3.5) (without considering *ΔS*), respectively. However, the slope of prediction without *ΔS* is much steeper than that with *ΔS* (i.e., 1.41 vs. 1.10, referring to the perfect slope 1.0). This implies that ET inter-annual variance is over-estimated in arid climates if terrestrial storage change is not considered.

In humid climates where *P* variance can be neglected, ET variances calculated by Eqn.(3.6) (with considering *ΔS*) and Eqn.(3.7) (without considering *ΔS*) for basins with $\bar{\phi} < 1$ are compared with the assessed ET variance, as shown in Figure 3.3b. Opposite to the results in arid climates, the regression slope between the assessed *ET* variance and the prediction without *ΔS* is much smaller than that with *ΔS* (i.e., 0.40 vs. 0.54). In other words, in humid climates, the ET variance would be under-estimated without considering terrestrial storage change. Note that there are only 7 humid basins in the dataset; the regression between the assessed and predicted ET variance is expected to be improved if more humid basins are represented.

ET variance is usually dampened in arid climates and enhanced in humid climates by increased terrestrial storage change (e.g., reservoirs and glaciers) because of the different roles of terrestrial storage change under these different climates. The buffer effect of terrestrial storage on ET in arid climate has been well recognized in previous studies [*Sankarasubramanian and Vogel*, 2002; *Potter et al.*, 2005; *Nicholas J. Potter and Lu Zhang*, 2009]. In humid climate, terrestrial storage carries *P* from an energy-limited period to a hot period as complementary water supply to sustain ET. That is to say, the terrestrial storage adjusts the temporal distribution of water availability for ET, increases ET in periods with high energy supply and thus enhances ET variance. This also results in larger contribution from *PET* to ET variance [*Karam and Bras*, 2008]. This contrast indicates the asymmetric role of terrestrial storage control on ET variance. Essentially, terrestrial storages such as aquifers and reservoirs (either natural or man-made) mitigate the impact of *P* variance on ET variance.

However, it is interesting to note that terrestrial storages do not significantly buffer the energy supply fluctuation at the annual scale, as shown in the primary analysis provided by Budyko [*Budyko*, 1974]. We may understand "terrestrial storage" as both terrestrial energy storage and terrestrial water storage. ET links water and energy budget in general if we neglect the fluxes across the boundary. Terrestrial energy storage can be transferred in other ways such as heat conduction, convection and radiation; however, terrestrial water storage can only be consumed via ET, if neglecting leakage to deep aquifer and/or lateral fluxes across the boundary. In this sense, the asymmetric role of terrestrial storage control on ET variance is illustrated in Figure 3.1b and Figure 3.2b for annual scale and monthly scale, respectively.

3.3.3 ET intra-annual variance sources

The ET intra-annual standard deviation σ_{ET} predicted by Eqn.(3.1) for the thirty-two basins are plotted in Figure 3.2a. The r-squared and Nash-Sutcliffe coefficient between assessed and predicted intra-annual σ_{ET} are 0.87 and 0.83, respectively, with an average error of 1.62 mm. The estimation of ET intra-annual variance is more accurate than inter-annual variance mainly due to more data in a monthly sequence (i.e., 276 months) in this study. In addition, the seasonality at the intra-annual scale may also introduce certain ET patterns.

The percentage contributions to ET intra-annual variance σ_{ET}^2 from each source in Eqn.(3.1) are displayed in Figure 3.2b. At the intra-annual scale, main climatic controls on ET variance are found in several basins. In arid regions, *ET* variance reflects *P* variance; it can be seen that more than half of ET variance is attributed to *P* variance in the Amur ($\bar{\phi}$ = *2.25*), Yellow ($\bar{\phi}$ = *2.64*), and Senegal ($\bar{\phi}$ = *8.33*) basins. In humid basins, ET variance is more related to *PET* variance, such as in the Pechora ($\bar{\phi}$ = 0.65),

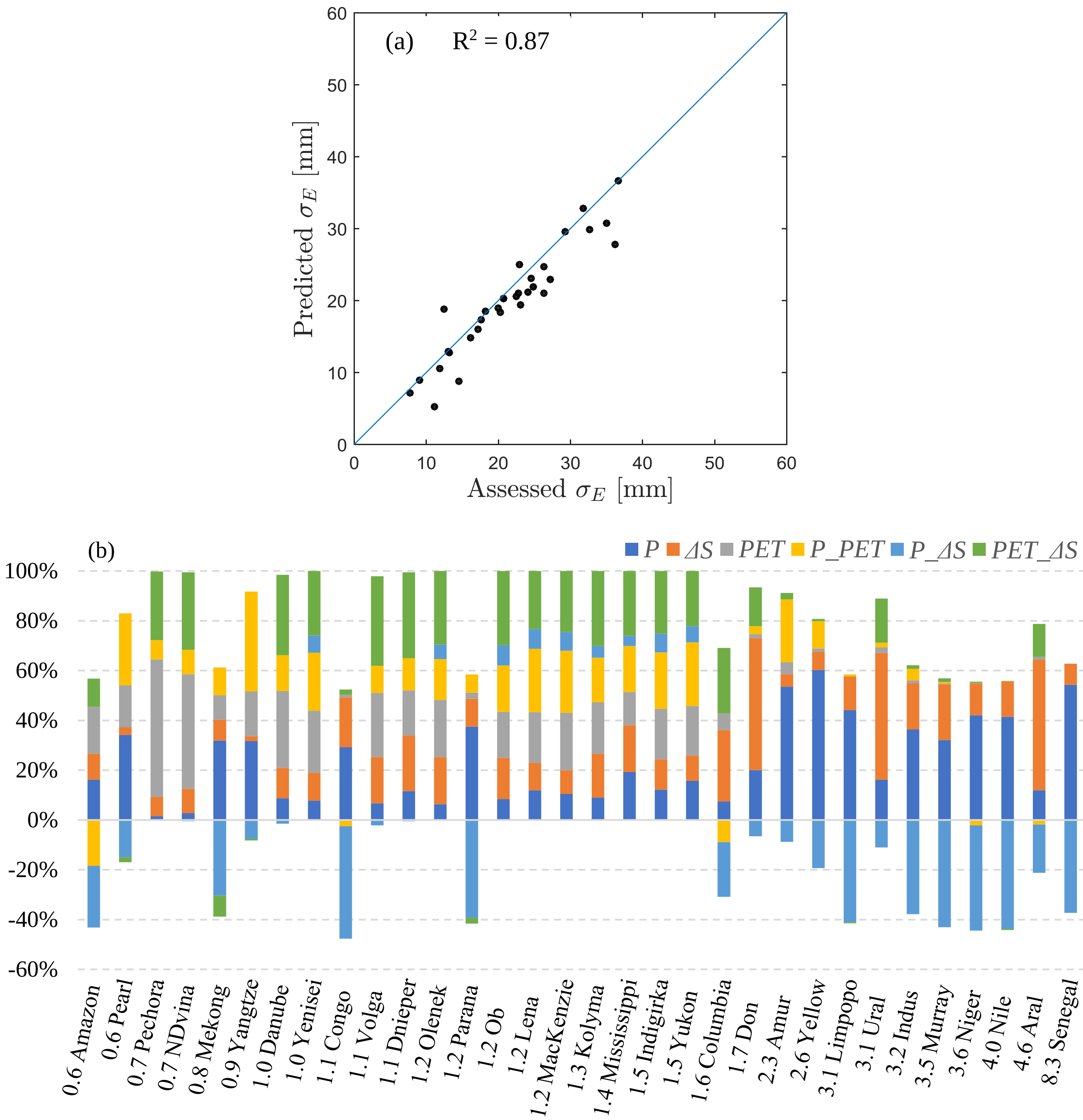


Figure 3.2. a) ET intra-annual standard deviation from assessed data and prediction by Eqn.(3.1). b) Percentage contribution to ET intra-annual variance from each component in Eqn.(3.1). Basins are listed from left to right with increasing aridity index.

where *PET* contributes more than 55% of ET variance. On the other hand, a main climate control on ET intra-annual variance is not exhibited in all the basins in Figure 3.2b. Instead, in some basins, all ET variance sources in Eqn.(3.1) jointly contribute to ET variance. Furthermore, terrestrial storage change plays a more important role in ET variance in some basins than others at the intra-annual scale. For example, the contribution from ΔS variance accounts for more than half of the ET variance in the Aral, Don, and Ural basins. This indicates that terrestrial storage becomes a major control of ET variance, and it is more capable of accommodating climate fluctuations at a finer time scale. The monthly ΔS time series in these basins show larger fluctuation than the *P*, since the storage is enough to accommodate the limited *P* at monthly scale. For instance, reservoirs with a relatively small capacity can still be used for regulating flow fluxes within a year rather over years.

Interestingly, catchments' responses to climate show opposite patterns in different climates in the case of intra-annual ET variance (Figure 3.2b). In catchments where $w_{PET,\Delta S}cov(PET,\Delta S)$ enhances ET variance (e.g., in the Dnieper, Northern Dvina, and Volga basins), the impact from $w_{P,\Delta S}cov(P,\Delta S)$ is trivial. These basins are located in cold regions where snow accumulation/melting processes are significant features in the water budget. This type of terrestrial storage change follows the *PET* cycle, retaining water from a cold season for ET water consumption in a warm season. On the other hand, in catchments where $w_{P,\Delta S}cov(P,\Delta S)$ dampens ET variance (e.g., the Congo, Niger, and Limpopo basins), the impact from $w_{PET,\Delta S}cov(PET,\Delta S)$ is negligible. These basins have arid climates, and *PET* is not a limiting factor on ET. Here, terrestrial storage (e.g., soil moisture content and groundwater) follows *P* cycles and holds water from a wet season to sustain ET during dry seasons.

The seasonality of *PET* and *P* (represented by $w_{P,PET}cov(P,PET)$) is another significant source of ET variance in some basins such as Yangtze, where *PET* and *P* seasonality contributes to more than 40% of ET variance. In the dataset, most basins have an in-phase *PET* and *P* seasonal pattern, which yields positive *PET* and *P* covariance; only Amazon and Columbia have slightly out-of-phase *PET* and *P* seasonality.

3.3.4 Terrestrial storage control on ET intra-annual variance

The effects of terrestrial storage on ET intra-annual variance are assessed for arid and humid climates, respectively. Figure 3.3c shows the ET variance from assessed dataset and predicted by Eqn.(3.4) (with considering ΔS) and Eqn.(3.5) (without considering ΔS) for 25 arid basins. As can be seen, the estimation without ΔS does not capture the assessed ET variance at all, with r-squared equal to 0.004. This implies that, even in arid climate ET variance cannot be solely explained by *P* variance at the intra-annual scale, since basins can enlarge storage capacity to store the whole annual P. As pointed out by *Koster et al.* [2006], some criteria should be adopted to exclude basins with significant terrestrial storage changes to avoid the bias in the estimate of ET intra-annual variance if storage change information is not available.

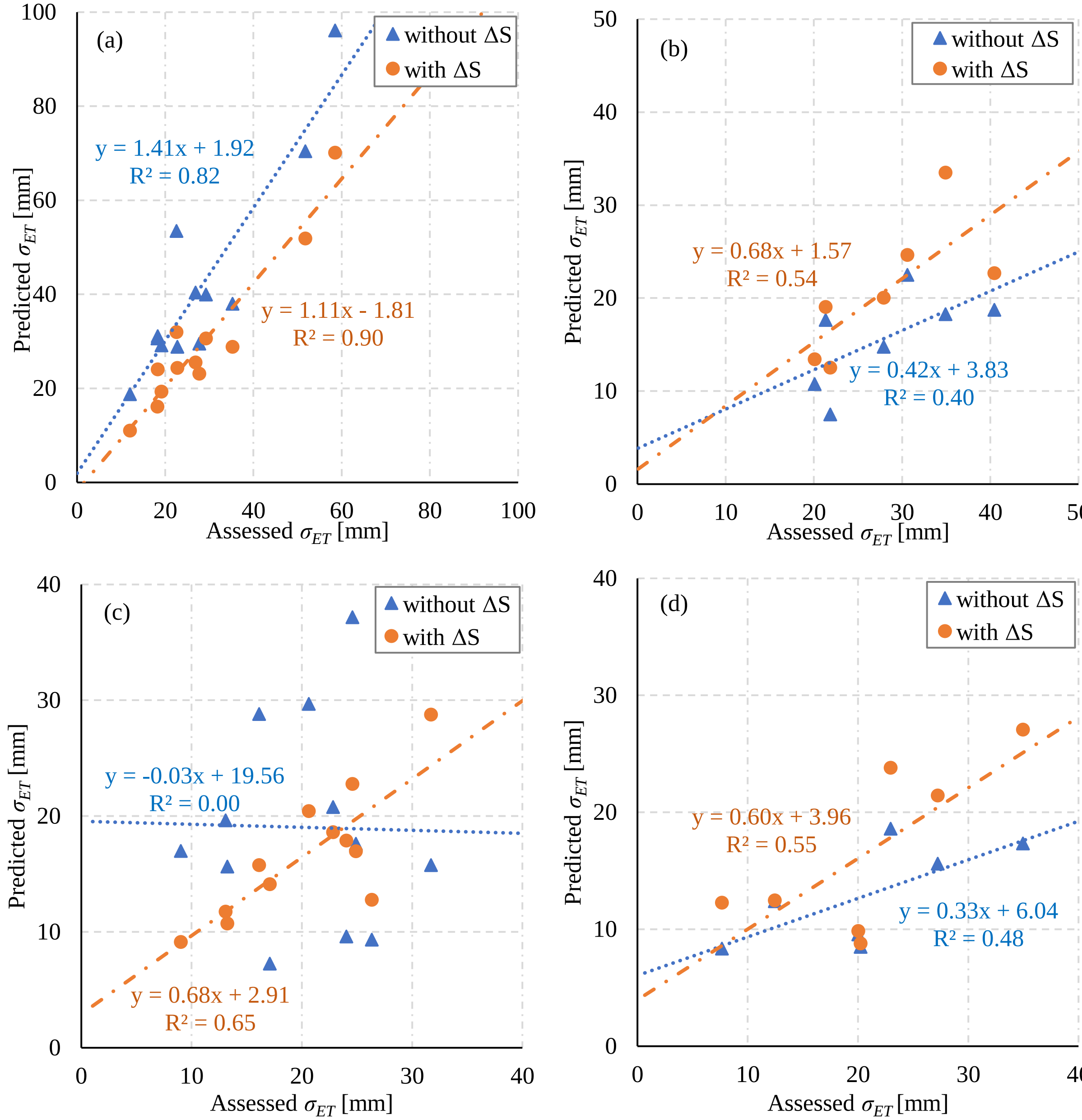


Figure 3.3. Effect of terrestrial storage change on ET inter-annual variance in a) arid climates and b) humid climates, and effect of terrestrial storage change on ET intra-annual variance in c) arid climates and d) humid climates.

The method developed by *Zeng and Cai* [2015] incorporates *ΔS* in the estimation of ET temporal variance (Eqn.(3.4) and (6)). Using this method, the r-squared of the intra-year estimate increases to 0.65. Thus, terrestrial storage change shifts from being a "buffering" effect at the inter-annual scale to a main factor at the intra-annual scale in arid climates. It is noted that the r-squared for ET intra-annual variance is smaller than the inter-annual variance due to increased contributions from other factors (e.g., seasonality) at a fine time scale.

In Figure 3.3d, Eqn.(3.7) (without considering ΔS) yields a larger regression slope (0.33 vs. 0.60) than Eqn.(3.6) (with considering ΔS) for the seven humid basins with $\bar{\phi} < 1$. Also, the r-squared improves from 0.48 to 0.55, which implies that ET intra-annual variance would be under-estimated without considering ΔS in humid climates, and a similar situation is observed at the inter-annual scale.

3.3.5 Geographic pattern of ET intra-annual variance

As there are more factors for ET variance at the intra-annual scale, catchments are categorized into five groups by clustering based on the magnitude of each source, which results in a geographic pattern as shown in Figure 3.5. *PET* dominates ET variance in the two most humid and cold basins (i.e., the Northern Dvina and the Pechora). ΔS dominates ET variance in three arid basins in Middle Asia (i.e., the Aral, Don and Ural). ΔS in these basins is driven by *PET* and exhibits much larger variation than P. The value of the weighting function $w_{PET,\Delta S}$ is relatively small due to the arid climate. As a result, ΔS becomes the main source of ET variance. Four basins in the Indian Monsoon region (i.e., Amur, Yellow, Pearl, and Yangtze) have ET intra-annual variance mainly attributed to *PET* and *P* seasonality. Warm seasons temporally coincide with rainfall seasons in these basins, and they also have relatively small storage variation, since to a large extent the ET water supply meets the ET energy demand during the vegetation growing season. The basins where ET variance is controlled by *P* and ΔS are located in low latitude arid regions in general. Note that although the Mekong basin is in the Indian Monsoon region, it has a sub-tropical climate, and the ET intra-annual variance is dominated by *P* and ΔS given that the energy-water supply does not show significant seasonal phasing.

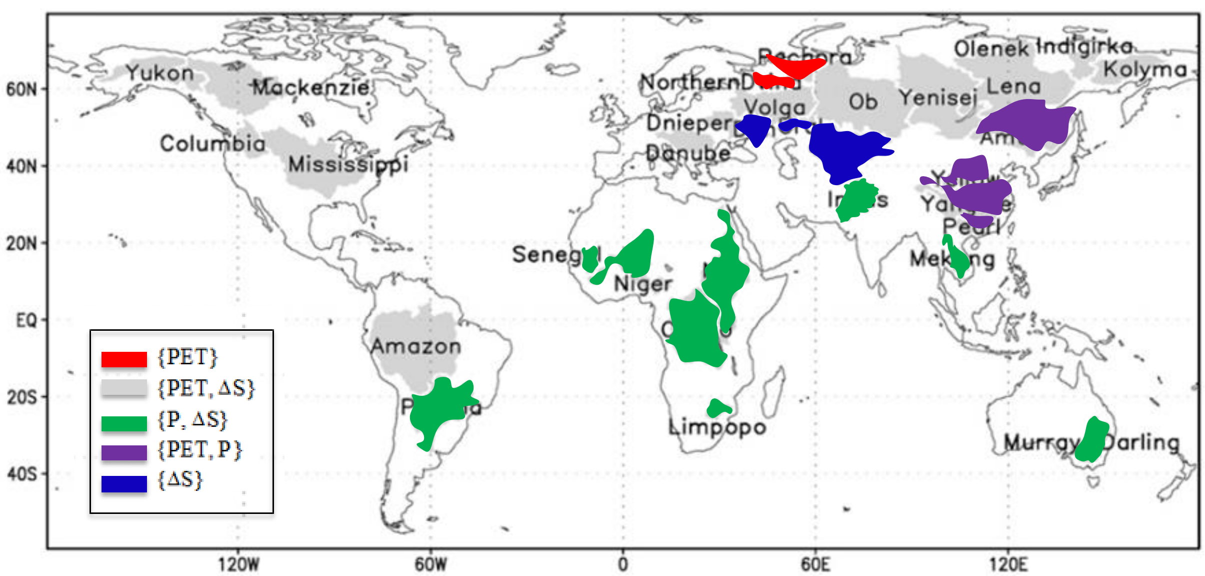


Figure 3.4. Geographic zonation based on the sources of ET intra-annual variance. The main control on ET intra-annual variance in denoted in the braces.

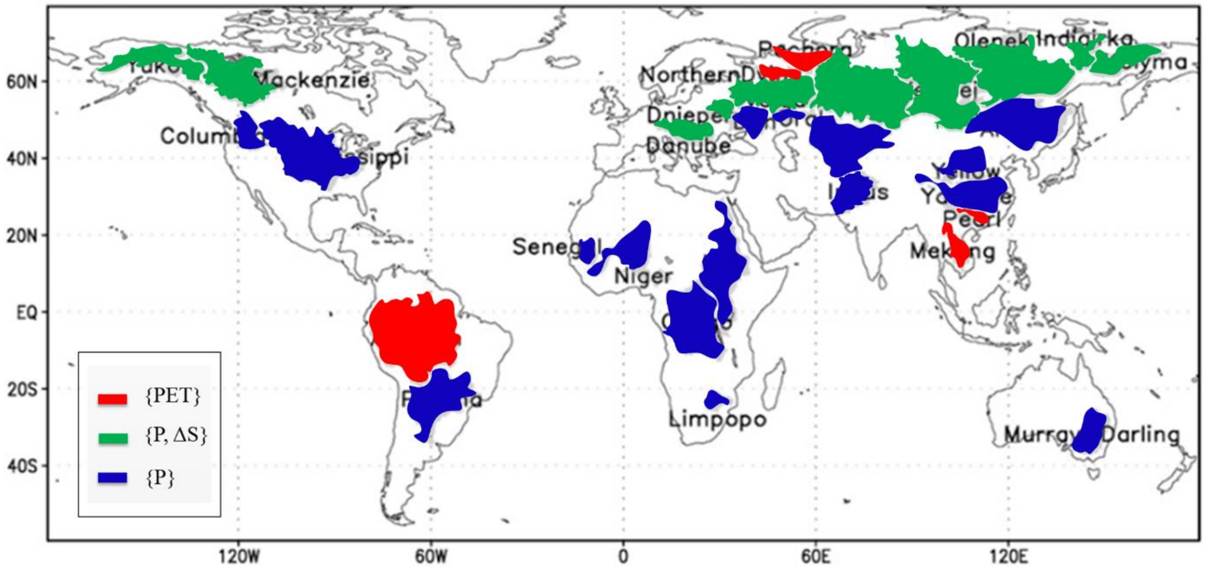


Figure 3.5. Geographic zonation based on the sources of ET inter-annual variance. The main control on ET inter-annual variance in denoted in the braces.

We also conducted the analysis of ET variance classification at the inter-annual scale as shown in Figure 3.6. The ET inter-annual variance yields three patterns dominated by P, PET, and P & ΔS, respectively. In addition to the idea of climatic control on ET inter-annual variance, the classification result shows that the catchment storage changes as additional control to ET inter-annual variance mainly in moderate basins ($2/3<\bar{\phi}<3/2$, and annual P and PET are less than 1000 mm). Catchment storage cannot buffer climatic fluctuations at the annual scale under very arid or humid climate, since the storage capacity is limited compared to the annual P or PET at the annual scale.

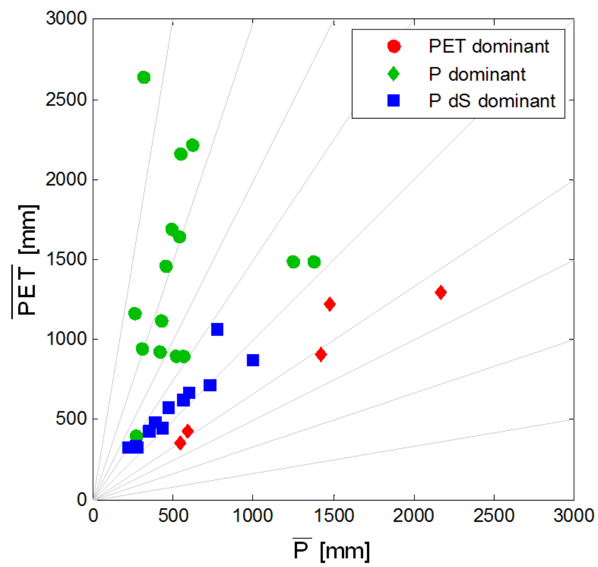


Figure 3.6. Main factors controlling ET inter-annual variance in $\overline{P}$ and $\overline{PET}$ plane

3.4 Discussion and conclusions

This study explicitly quantifies the effect of storage change on ET variance at both inter- and intra-annual scale. The impact of terrestrial storage on both the average and variance of catchment water balance has been recognized in many studies by introducing terrestrial storage-related factors, such as plant-available water capacity [*Zhang et al.*, 2001] and soil moisture storage capacity [*Sankarasubramanian and Vogel*, 2002]. Those studies consider catchments' responses to *P* variance only and show the "damping" effect of terrestrial storage. Our study considers catchments' responses to both *P* and *PET* variance and thus represents more comprehensive factors on ET variance. That is, ET variance would be dampened by terrestrial storage in arid climates, but strengthened by it in humid climates. The different impacts highlight that without considering the effect of terrestrial storage, ET variance would be possibly over- and under-estimated in arid and humid climates, respectively. This is also a possible reason for the systematic bias in the simulated ET variance from climate models that do not accurately represent terrestrial storage [*Mueller et al.*, 2011].

For cold regions, it is found that frozen rainfall does not significantly change the long-term average ET since ET tends to be energy-limited throughout the time [*Williams et al.* [2012]. Our study confirms that frozen rainfall may not be significant for ET inter-annual variance, however the melting of frozen rainfall strengthens the intra-annual ET variance. Even in those humid and cold basins (e.g., Pechora and Volga) where *PET* is a limiting factor at the inter-annual scale, ET is controlled by *P* during warm seasons at the intra-annual scale. As a result, the terrestrial storage in the form of snow pack and frozen soil moisture would be important to sustain warm season ET in the humid and cold regions [*Dunn et al.*, 2007; *Lettenmaier and Milly*, 2009].

In the Indian Monsoon regions, ET intra-annual variance is mainly controlled by the seasonal pattern of *P* and *PET*, which is significantly affected by the variation of monsoon systems [*Fasullo and Webster*, 2003; *Hoyos and Webster*, 2007]. Current in-phase *P* and *PET* pattern is suitable for vegetation and crop ET, and no significant storage capacity is needed to buffer the imbalance between *P* and *PET*. Furthermore, the phasing patterns between *P* and *PET* at inter- and intra-annual scale represent distinct characteristics of local climatic characteristics. At the annual scale, *P* and *PET* are out-of-phase for all 32 basins as shown by the negative $\mathrm{cov}(\mathrm{P}, \mathrm{PET})$ in Figure 3.1a, which indicates that dry years come with hot years. While at the monthly scale, some basins (e.g., Amazon and Columbia) have an out-of-phase *P* and *PET* pattern as shown in Figure 3.2a. Both in-phase and out-of-phase *P* and *PET* represent a climatic pattern that drives local hydrologic processes.

This study re-examines ET variance with an emphasis on the impact of terrestrial storage change and finds that ΔS damps/strengthens ET variance in arid/humid climates, respectively. Although we can calculate these terms from water balance time series, the terrestrial storage change data are not always available. A more profound issue is how to derive those covariance terms. With the advances in monitoring components of the hydrologic cycle, it is ready to calculate the covariance terms from independent observation data sources. Moreover, recent studies have shown that terrestrial storage can be inferred from runoff observation by distinguishing the liquid/non-liquid form of storage change [*Riegger and Tourian*, 2014].

Most of basins in the dataset used by this study have in-phase *P* and *PET* seasonality, yet ET variance under completely out-of-phase *PET* and *P* seasonality is worth studying since such catchments simultaneously experience water-limited and energy-limited conditions within a year [*Rana and Katerji*, 2000]. For example, *Ryu et al.* [2008] found that ET variance is much less than the climate variance in the Mediterranean climate zone in California. *Williams et al.* [2012] found that the storage cannot buffer the strong seasonality of Mediterranean climate and is insufficient to fully carry over precipitation from a wet season to a dry season. Under this situation, the impact of anthropogenic-related terrestrial storage change (such as irrigation) will have a major contribution to ET variance [*Bridget R. Scanlon et al.*, 2012; *Wada and Bierkens*, 2014]. Further research is needed to characterize the anthropogenic impacts, such as land use change irrigation, water storage and groundwater pumping [*Weiskel et al.*, 2014; *Vogel et al.*, 2015], on ET variance for regions with out-of-phase *PET* and *P* seasonality. On the other, future research is also needed to understand how intra- and inter-annual ET variance change affects water resources management under changing climate, especially to match crop ET requirement [*Cai et al.*, 2015].

ET inter- and intra-annual variances are assessed separately in this study, yet there remains a fundamental question about how ET variance is linked across scales [*Zanardo et al.*, 2012]. Studies have reported the impact on ET long-term mean from climatic variance at the inter-annual [*Li*, 2014], seasonal [*N.J. Potter and L. Zhang*, 2009; *Feng et al.*, 2012] and storm event scales [*Potter and Zhang*, 2007]. Those studies highlight the increasing significance of terrestrial storage change on ET variance at even smaller time scales. Further studies should address whether there exists any underlying mechanism governing the variance across the scales and explore how small-scale variance propagates to large-scale variance.

CHAPTER 4
ET TEMPORAL AND SPATIAL PATTERNS CHANGE IN THE HIGH PLAINS

The global basins in Chapter 3 have a large spatial coverage where human activities are relatively small compared to climatic forcings. In this chapter, we focus on basins in the High Plains where extensive groundwater irrigation has significantly altered the spatial and temporal patterns of ET. Having a similar climate, analysis on these sub-basins illustrate how farmers' irrigation behavior heterogeneity propagates to the heterogeneity of ET signatures.

4.1 Introduction

Crop production depends on massive groundwater-fed irrigation in many places around the world. Groundwater provides relatively stable water sources, especially in arid and semi-arid regions; however, over-exploitation of groundwater resources has been recognized as a major concern in sustainable regional development [*Konikow*, 2011; *Aeschbach-Hertig and Gleeson*, 2012; *Bridget R. Scanlon et al.*, 2012; *Famiglietti*, 2014]. The Republican River Basin, located in the Northern High Plains and comprising parts of Nebraska, Colorado, and Kansas, provides an outstanding example of how intensive groundwater irrigation effects propagate to hydrologic change, ecological deterioration and water rights conflicts [*McGuire*, 2014]. Since the 1960s, the widespread adoption of central pivot irrigation systems has gradually caused groundwater depletion and reduced stream flow in the Republican River. In 1998, the downstream state Kansas sued upstream states Nebraska and Colorado in the Supreme Court for violating surface water rights due to the over pumping of groundwater. At the global scale, many other regions, such as California's Central Valley [*B. R. Scanlon et al.*, 2012], the North China Plain [*Liu et al.*, 2008], India [*Rodell et al.*, 2009] and the Middle East [*Joodaki et al.*, 2014], have experienced similar situations of irrigation development and aquifer depletion. As seen in the above examples, human interferences are impacting hydrologic systems; water resource management will increasingly require scientific understanding of human impacts on hydrologic systems as future water needs change due to socioeconomic development and climate change.

Many studies have been conducted to understand groundwater depletion [*Strassberg et al.*, 2009; *B. R. Scanlon et al.*, 2012; *McGuire*, 2014; *Haacker et al.*, 2015] and streamflow changes [*Szilagyi*, 2001; *Burt et al.*, 2002; *Zeng and Cai*, 2014] in the High Plains. However, the impacts of irrigation on the temporal and spatial pattern of evapotranspiration (ET) change in the region are not well documented. ET represents the major water consumption in the region, especially for agricultural water use, and some studies have argued that water resources management is essentially "ET management" [*Foster and Garduño*, 2004]. Meanwhile, ET connects land surface energy and water budgets, driven by both climatic and anthropogenic forcing. A comprehensive assessment of patterns of ET change is essential, not only to understand the human interferences to the hydrologic cycle retrospectively, but also to design sustainable water resources management for this region.

Limitations in ET observation and simulation prevent a comprehensive understanding of human induced change over a large temporal span and spatial extent, such as the High Plains. Current ET remote sensing products are available only after the 1980s, while irrigation has been a practice for hundreds of years [*Mutiibwa and Irmak*, 2013]; thus a complete picture of the "pre-development" condition is not available for a baseline comparison. Furthermore, land surface models simulate irrigation as a result of soil moisture conditions [*Ozdogan et al.*, 2010; *Lawston et al.*, 2015], either neglecting anthropogenic forcings (e.g., pumping volume) or using estimates subject to significant bias and uncertainty [*Rossman and Zlotnik*, 2013; *Demissie et al.*, 2015].

This study provides a quantitative framework to assess ET temporal and spatial changes using long-term climate data and water table observations [*Haacker et al.*, 2015] in the High Plains. To understand human interferences to ET processes and diagnose the climatic and anthropogenic factors to ET changes, the framework is established in the context of a coupled-nature-human-system (CNHS) [*Liu et al.*, 2007]. Questions to address include: 1) What are the changes of ET temporal characteristics (i.e., mean and seasonal variance) due to groundwater-fed irrigation in the High Plains? 2) What spatial changes have irrigation practices imposed upon the natural hydroclimatic gradient in the region? 3) How are the ET change patterns related to farmers' behaviors in balancing crop production profit and risk aversion, and is there any general mechanism that governs the coupling of natural and human systems in a CNHS, and/or emerging phenomena at the system level that reveal the interactions of the two? Answers to these questions are also expected to provide a predictive understanding to what may happen in other regions that have experienced similar challenges.

4.2 Data and methods

Monthly climate data from 1940 to 2010, including precipitation (P), maximum temperature and minimum temperature, were obtained from PRISM Climate Group with a spatial resolution of 30 arcsec (~800 m) [*Daly et al.*, 2008]. Potential ET (PET) is calculated from the Hargreaves temperature-based method [*Hargreaves and Samani*, 1982]. The assessment with temperature based PET is consistent with another PET dataset calculated from a modified Penman scheme [*L. Mahrt and Michael Ek*, 1984] in the North American Land Data Assimilation System. With annual P ranging from 328 mm to 830 mm and annual PET from 995 mm to 1538 mm, the High Plains have an arid or semi-arid climate, where the aridity index (i.e., *PET/P*) varies between 1.4 and 4.7. The water table change is interpolated from groundwater well measurements over the Ogallala Aquifer by ordinary kriging at a resolution of 250 m, following methods detailed in *Haacker et al.* [2015]. On average, about 12000 measurements in each season are used for spatial interpolation. In addition, elevations from 1984 stream locations are incorporated to filter out the interpolated water table values that are above the terrain surface. Since most of the regional aquifer is unconfined, the water table change is multiplied by specific yield [*Gutentag*, 1984; *McGuire et al.*, 2012] to obtain the aquifer water storage change (ΔS), which is calculated for both growing season (i.e., May to October when pumping wells are active) and non-growing season in this study. Furthermore, to account for the hydroclimatic and anthropogenic heterogeneity, the USGS hydrologic unit code (HUC) 1:250000-scale Hydrologic Units (HUC250k) map [*Seaber et al.*, 1987] is used to delineate the High Plains overlaying with the Ogallala Aquifer into 120 sub-basins. The PRISM climate data and USGS groundwater storage change are aggregated to the sub-basin level based on HUC250k boundary.

The annual mean ET is calculated using the modified Budyko curve by incorporating the change of terrestrial water storage [*Han et al.*, 2011; *Zeng and Cai*, 2015]:

$$\frac{ET}{P-\Delta S} = 1 + \frac{PET}{P-\Delta S} - \left[1 + \left(\frac{PET}{P-\Delta S}\right)^{\varpi}\right]^{1/\varpi} \qquad \text{Eqn.(4.1)}$$

For sub-basins with negligible groundwater pumping (i.e., $\Delta S = 0$), Eqn.(4.5.1) is the same as the original Budyko curve [*Fu*, 1981]. The aquifer storage depletion (i.e., $\Delta S < 0$) increases total water availability (P-ΔS), which is partitioned between ET and runoff. In this study, the empirical parameter ϖ is fixed at 2.6 based on our previous study [*Zeng and Cai*, 2014].

The ET variance is calculated based on the Evapotranspiration Temporal VARiance Decomposition (ETVARD) framework [*Zeng and Cai*, 2015; 2016]:

$$\sigma_{ET}^2 = w_P\sigma_P^2 + w_{PET}\sigma_{PET}^2 + w_{\Delta S}\sigma_{\Delta S}^2 + w_{P,PET}cov_{P,PET} + w_{P,\Delta S}cov_{P,\Delta S} + w_{PET,\Delta S}cov_{PET,\Delta S} \qquad \text{Eqn.(4.2)}$$

where the ET temporal variance (σ_{ET}^2, at seasonal scale in this study) is decomposed into components from climatic fluctuations (i.e., σ_P^2 and σ_{PET}^2), seasonality between water and energy supply (i.e., $cov_{P,PET}$), groundwater storage variation (i.e., $\sigma_{\Delta S}^2$), and the responses of groundwater storage change to climate (i.e., $cov_{P,\Delta S}$ and $cov_{PET,\Delta S}$). The weighting factors (*w*), quantifying the contribution from each source to σ_{ET}^2, are derived from the sensitivity of ET to land surface moisture or energy constraints based on the long-term climatic condition (i.e., the aridity index, *PET/P*. When farmers respond to low rainfall with groundwater-fed irrigation, the activity is captured via strong correlation between *P* and ΔS. At the same time, the arid and semi-arid climate condition yields a high value of weight $w_{P,\Delta S}$, which also reflects the irrigation impact via terrestrial storage change. Therefore, Eqn.(4.2) can be used to quantify the contribution attributed to both climatic and storage change factors (especially via irrigation in this study). The function forms for the six weight factors and detailed discussion for σ_{ET}^2 under the various climate and storage change conditions are documented in *Zeng and Cai* [2015].

ET temporal variability exhibits different magnitudes and contains different components depending on the time scale (e.g., annual, seasonal and monthly) [*Zeng and Cai*, 2016]. This study focuses on ET temporal variability at the seasonal scale (i.e., growing and non-growing season) for two reasons. First, in the crop grow season, aquifer water storage change is the main source of terrestrial water storage change; while other forms of water storage change such as soil moisture and snow cover are negligible. In addition, direct observations on soil moisture are not available for any large spatial extent and temporal span, and soil moisture from land surface model simulations may suffer from significant bias and uncertainty as irrigation is not well represented in existing models. Second, agricultural pumping is a seasonal event, and its impacts on ET patterns are with a seasonal scale.
The 70-year seasonal time series of *P, PET* and ΔS are divided into two periods (i.e., pre-1975 and post-1975, with the same length for both periods) to calculate ET mean by Eqn.(4.1) and temporal variance by Eqn.(4.2), respectively. It is noted that large scale groundwater-fed irrigation development in this region occurred around 1950, especially in the South High Plains [*McGuire*, 2009; *Bridget R. Scanlon et al.*, 2012; *Haacker et al.*, 2015]. Although the two periods (pre-1975 and post-1975) are not accurately referred to "pre-development" and "development" conditions, the degree of intensive and extensive groundwater pumping after 1975 provides a comparative case of human interferences on ET patterns.

4.3 Results

4.3.1 Spatial and temporal change in ET mean and variance

Most of the sub-basins yield a greater mean ET post-1975 than pre-1975, as shown in Figure 4.1a. On average, ET increases by 74.1 mm from 399.6 mm pre-1975 to 473.7 mm post-1975, and the increase is more than 100 mm in some sub-basins. An increase in *P* is observed in most sub-basins as shown in Figure 4.3a, but the magnitude of *P* increase is only 35.0 mm on average. Therefore, irrigation from groundwater pumping contributes to more than half of the ET increase.

Spatially, the average ET exhibits a clear east-to-west gradient pre-1975 as captured by the left map of Figure 4.1c. ET is relatively large in sub-basins in the east and small in the west, following the *P* gradient. Under significant irrigation practices post-1975, the east-to-west gradient in ET is no longer obvious in the Central High Plains (CHP) and the Southern High Plains (SHP). The increase of mean annual ET is consistent with the groundwater depletion map. As shown in the right map of Figure 4.1c, most areas of the CHP and the north part of the SHP show a significant ET increase. These regions also experienced the most significant decline in water table [*Bridget R. Scanlon et al.*, 2012; *McGuire*, 2014; *Haacker et al.*, 2015]. In the Northern High Plains (NHP), ET increases are mainly located in the Republican River Basin and Lower Platte River Basin, while ET in the Upper Platte River Basin and

Sand Hills region has little change since agricultural development in these regions is not intensive due to the sandy soils. Therefore, the ET east-to-west gradient still remains in the NHP without significant anthropogenic interferences.

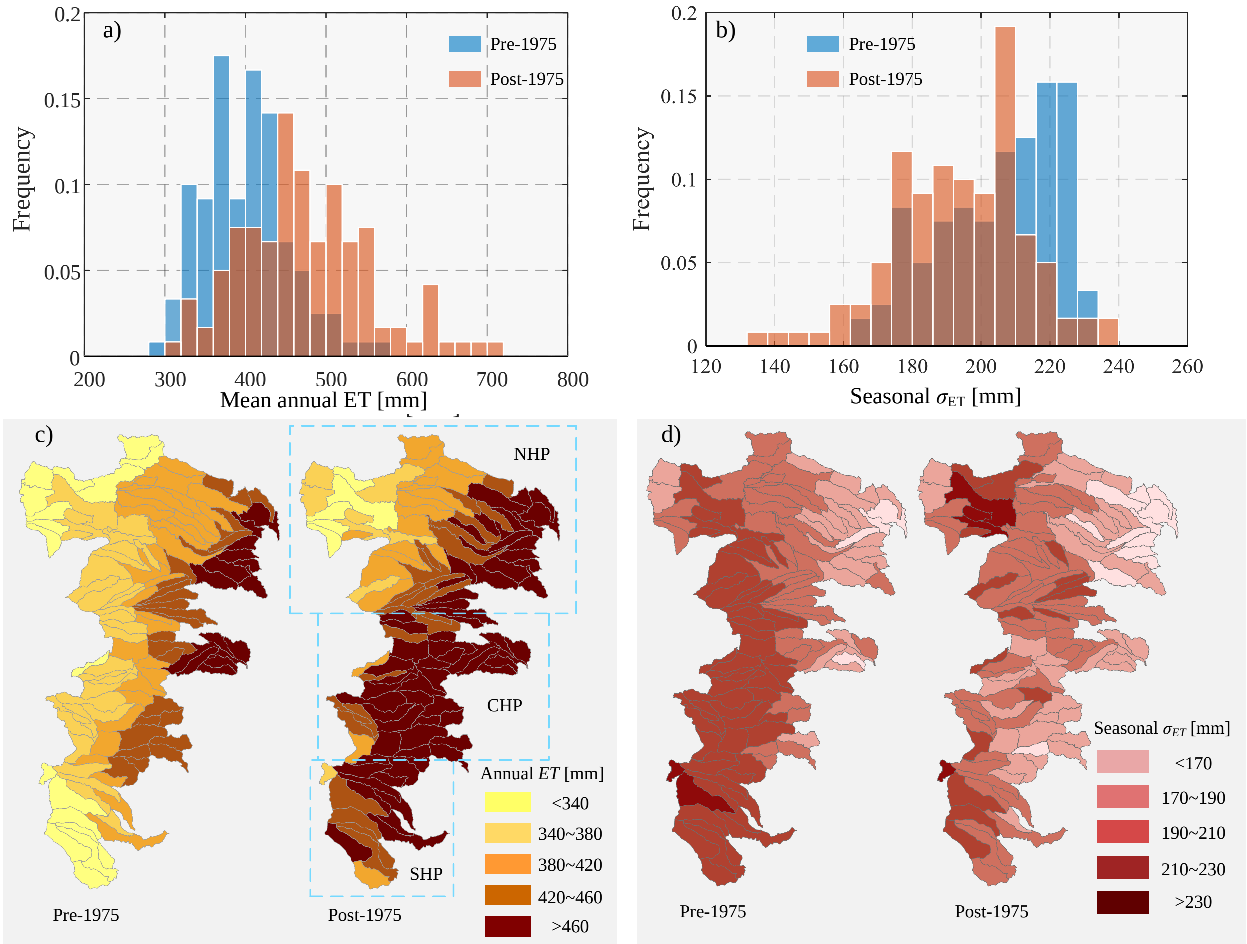


Figure 4.1. Histogram of a) mean and b) variance of ET pre-1975 and post-1975; the spatial pattern of ET c) mean and d) variance in 120 sub-basins over the High Plains.

The seasonal σ_{ET} decreases from 205.3 mm pre-1975 to 193.6 mm post-1975 on average, and the change can be attributed to climate and storage components according to ETVARD by Eqn.(4.2). On one hand, some sub-basins experience damping in climatic fluctuation, which contributes to the decrease of σ_{ET} by 20.7 mm on average, as shown in Figure 4.4a. The damping in climatic seasonal variability is mainly due to a more stable P in the region, since PET variability remains almost unchanged. On the other hand, the storage components in σ_{ET} increases by 14.2 mm on average as shown in Figure 4.4b, especially in sub-basins in CHP and SHP with significant amount of groundwater pumping. Therefore, the decline in the total σ_{ET} results from the combined effects of less variance in P and larger variance in the storage components due to irrigation pumping. In terms of absolute values, the climatic components and storage components account for 59% and 41% change in σ_{ET}, respectively.

The ET seasonal variance also followed an east-to-west gradient pre-1975 as illustrated in the left map of Figure 4.1d, where the arid sub-basins in the west exhibit higher ET variability than those relatively humid sub-basins in the east. However, more erratic spatial heterogeneity is seen in σ_{ET} post-1975 shown in the right map of Figure 4.1d. Despite a common decrease in σ_{ET} climatic components, the increase of σ_{ET} storage components at some locations causes the spatial heterogeneity, especially in the Republican River Basin in the NHP and some sub-basins in the CHP and SHP. The east-to-west gradient in σ_{ET} only remains in the north part of the NHP, where groundwater pumping is negligible.

Change in ET mean and variance can be represented by the coefficient of variation (*CV*), a dimensionless indicator for relative variability. To be consistent in the temporal scale, ET seasonal variance is normalized by the seasonal average (i.e., half of the annual average). Similar to the mean and variance, the *CV* of seasonal ET pre-1975 in the left of Figure 4.2a displays an apparent east-to-west gradient with lower *CV* in the east sub-basins and higher *CV* in the west sub-basins. However, the spatial gradient is not preserved post-1975, except for in the north part of the NHP. There is a consistent decrease in *CV* in most sub-basins in the south parts of the NHP, CHP, and SHP. Although some sub-basins have increased σ_{ET} due to irrigation, the increase in mean ET is more significant. Therefore, irrigation dampens ET variability in term of *CV*, while as shown above, irrigation increases the absolute value of variability (σ_{ET}). Some sub-basins in the CHP and east part of the NHP have *CV*s lower than 0.6, which is not observed pre-1975.

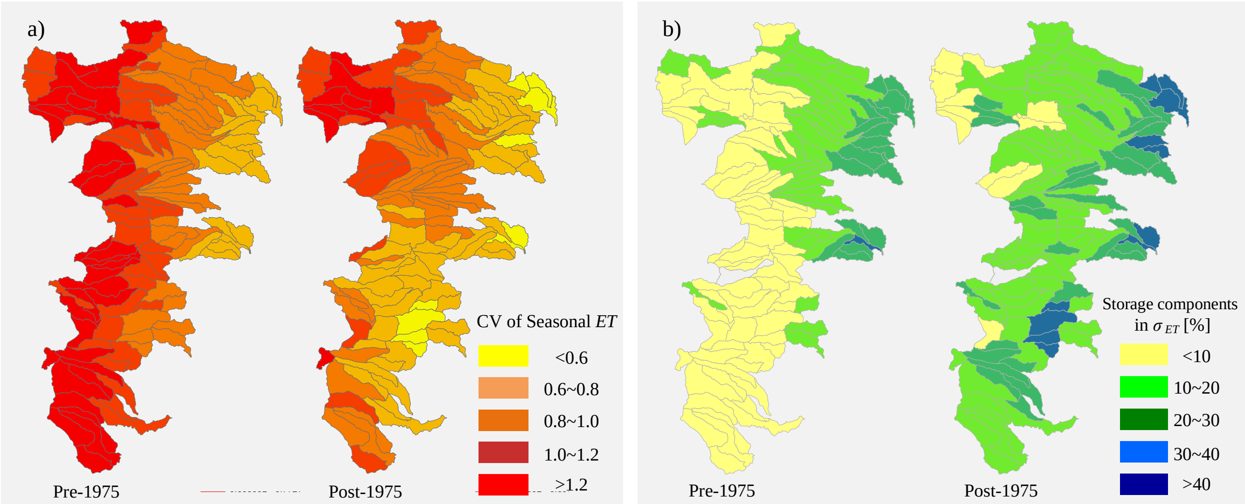


Figure 4.2. a) The coefficient of variation (*CV*) of seasonal ET and b) the percentage of storage components in σ_{ET}.

4.3.2 The sources of σ_{ET} change

As discussed above, changes in climate and groundwater-fed irrigation are the two main sources of σ_{ET}. Figure 4.2b illustrates the percentage of storage components (i.e., $w_{\Delta S}\sigma_{\Delta S}^2 + w_{P,\Delta S}cov_{P,\Delta S} + w_{PET,\Delta S}cov_{PET,\Delta S}$) in total σ_{ET}^2. Note that since the components from storage change can be negative, the percentage is calculated using the absolute values. Before 1975, the storage components account for less than 10% of σ_{ET} in most sub-basins of the CHP and SHP. The east part of the NHP has a higher portion (about 20%) with the storage components than the west part (less than 10%). After 1975, there is a significant increase in the storage components of σ_{ET}: Most sub-basins have more than 10% and some have even more than 40%. Sub-basins with contribution from storage components higher than 20% are distributed unevenly, showing some spatial heterogeneity (the left map of Figure 4.2b). The groundwater fluctuation ($\sigma_{\Delta S}$) and pumping response to rainfall deficit ($cov_{P,\Delta S}$) increase the contribution of storage components to σ_{ET}. The spatial heterogeneity caused by irrigation is also reflected by storage components in σ_{ET} in the right map in Figure 4.2b. Thus, compared to the climatic components, the anthropogenic induced storage components play a notable role in shaping the ET variability in this region.

4.4 Discussion

4.4.1 The overlapping of natural gradients and anthropogenic-induced heterogeneity

The spatial characteristics in the High Plains under natural condition exhibit a clear east-to-west gradient. Due to the rising terrain towards the west, *P* decreases from east to west [*Daly et al.*, 2008]. The regional water table under the "pre-development" condition (or the near natural condition) also follows the east-to-west gradient [*McGuire*, 2014; *Haacker et al.*, 2015]. The north-to-south gradient of *PET* (following the temperature gradient) does not manifest in the hydroclimatic spatial pattern in the arid and semi-arid climate, since the land surface processes in this region are constrained by *P*. The ET characteristics, including the mean, seasonal variance, seasonal *CV* and storage components, also show an apparent east-to-west gradient pre-1975. Natural vegetation or rain-fed crops depend on soil moisture to buffer the climatic fluctuation and have limited accessibility to groundwater. Therefore, the ET in the region pre-1975 was mainly affected by the climate, and its spatial pattern followed the east-to-west climatic gradient. With substantial groundwater-fed irrigation development, localized groundwater depletion propagates to ET pattern changes and introduces anthropogenic spatial heterogeneity over the natural east-to-west gradient, as shown by the post-1975 maps of Figures 1 and 2. As anthropogenic-induced storage components in σ_{ET} became significant, the hydrologic system in the High Plains went beyond a "natural" system, to be driven by both natural (climatic, geomorphic) and anthropogenic (e.g. land use and water withdrawals) factors. The High Plains has been shifted from a natural system to a CNHS with extensive irrigation development in the region.

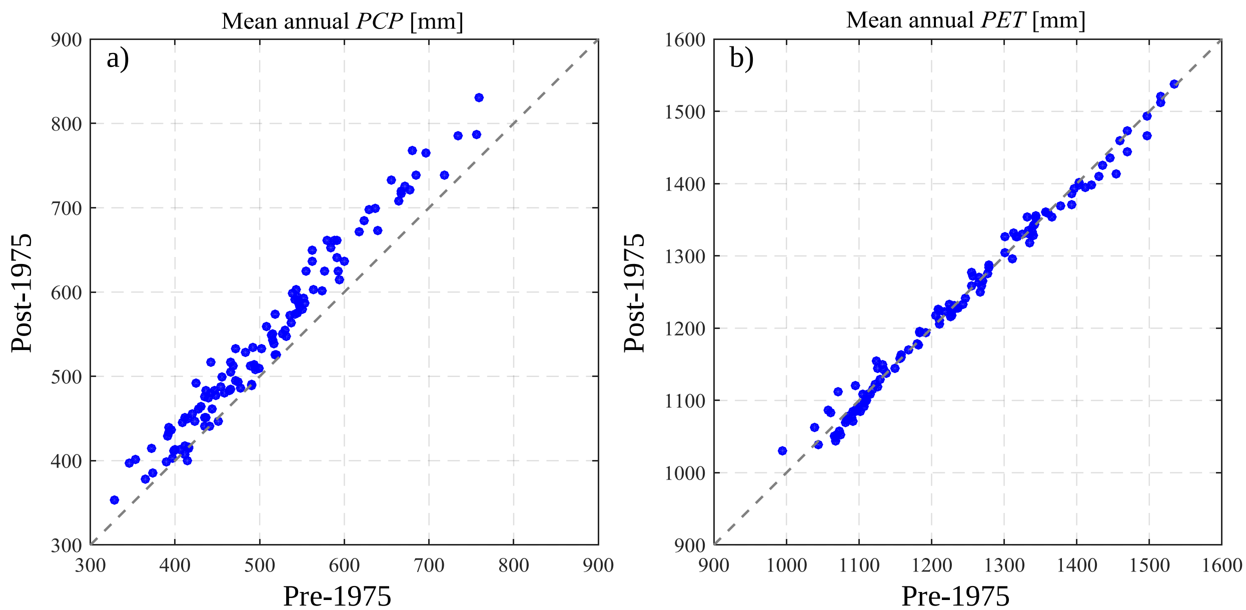


Figure 4.3. Mean annual a) precipitation (*P*) and b) potential evaporation (*PET*) of the 120 sub-basins in the High Plains pre-1975 and post-1975. Most sub-basins have an increased *P* post-1975, while *PET* generally remains unchanged.

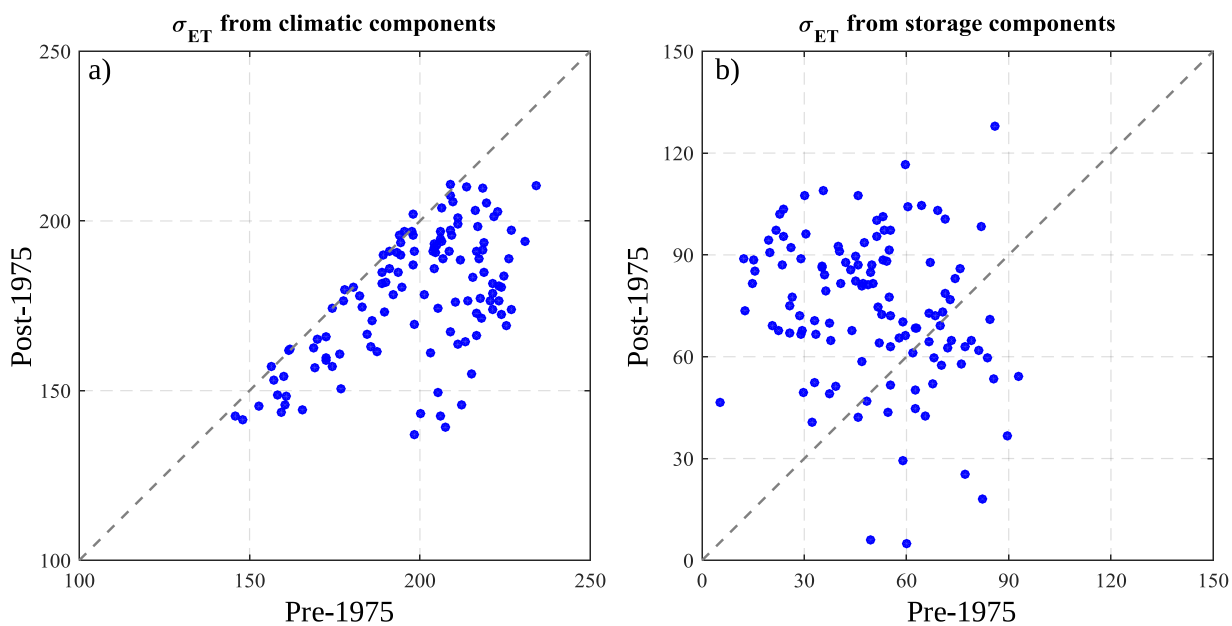


Figure 4.4. σ_{ET} sources from a) climatic components (i.e., $w_P\sigma_P^2 + w_{PET}\sigma_{PET}^2 + w_{P,PET}cov_{P,PET}$) and b) storage change components (i.e., $w_{\Delta S}\sigma_{\Delta S}^2 + w_{P,\Delta S}cov_{P,\Delta S} + w_{PET,\Delta S}cov_{PET,\Delta S}$) pre-1975 and post-1975.

The anthropogenic-induced ET heterogeneity in the High Plains highlights an indispensable issue in understanding hydrologic systems in the context of a CNHS, where the natural east-to-west gradient and anthropogenic-induced heterogeneity jointly shape the spatial pattern of ET. Our understanding of natural dynamics and capability to predict hydroclimatic changes have been fundamentally improved by observations across different spatial scales and spatially-distributed simulation models. The advances in data acquisition and model improvement enable earth scientists to explore the complexity of natural processes, which further supports and strengthens scientific-based decision-making on water resources development. However, the spatial heterogeneity on the human dimension is less well established in either the data or modeling aspects [*Xu et al.*, 2014]. Traditional water resources planning and management models generally adopt a simplified top-down and homogenous institution and overlook the heterogeneity in human behaviors [*Yang et al.*, 2009]. As evidenced in the ET pattern change in the High Plains, anthropogenic-induced ET components show even stronger heterogeneity than the natural gradient due to spatially diversified irrigation practices resulting from human behaviors. An individual farmer's irrigation decisions are affected by spatially distributed environmental conditions [*Noël and Cai*, 2017] and the farmer's response to weather forecasts [*Hejazi et al.*, 2014]. Further, *Foster et al.* [2014] found that irrigation behavior in Texas High Plains region exhibits complex nonlinear responses to changes in groundwater availability and well yield. Below the state level, groundwater-management authorities, such as Natural Resources Districts with the State of Nebraska, monitor groundwater usage and regulate farmers' pumping volume by setting pumping permits. At the state level, water right conflicts are settled by the Republican River Compact Administration, which allocates the water rights of the Republican River among Colorado, Kansas and Nebraska [*Draper*, 2007]. These cross-scale anthropogenic complexities require a new modeling paradigm to incorporate institutionally-sound and behaviorally-realistic decision mechanisms to support the modelling of CNHS [*Vogel et al.*, 2015]. Emerging efforts, such as of socio-hydrology [*Sivapalan et al.*, 2012] and hydro-geomorphology [*Vogel*, 2011], and the "bottom-up" approach in agent-based modelling [*Ng et al.*, 2011; *Noël and Cai*, 2017], provide promising guidelines towards strengthening the human dimension in CNHS.

4.4.2 Correspondence between ET pattern and crop production

ET processes occur within the context of crop production management for both irrigated and rain-fed agriculture. Thus, the pattern of ET change can help unveil crop production variability. Figure 4.5a shows the evolution of ET pattern for the 120 sub-basins for the periods pre- and post-1975. The pairwise points (pre- and post-1975) in each sub-basin generally shift from upper-left (high *CV* and low mean) to lower-right (low *CV* and high mean) as substantial groundwater pumping has developed over the region. Therefore, irrigation alters two aspects of crop water consumption by 1) increasing the mean crop water consumption (in order to mitigate the deficit in rainfall) and 2) damping the variation of crop water consumption (in order to buffer the climatic fluctuation). We may make an analogy between crop water consumption and crop yield by showing a similar pattern observed for croyield as shown in Figure 4.5b, the rain-fed and irrigated corn yield in Nebraska based on data from the United States Department of Agriculture National Agricultural Statistics Service. The rain-fed corn yield has high *CV* and low mean, and the irrigated corn yield exhibits low *CV* and high mean, corresponding to high mean and low variance of ET with irrigated corn and low mean and high variance of ET with rain-fed corn.

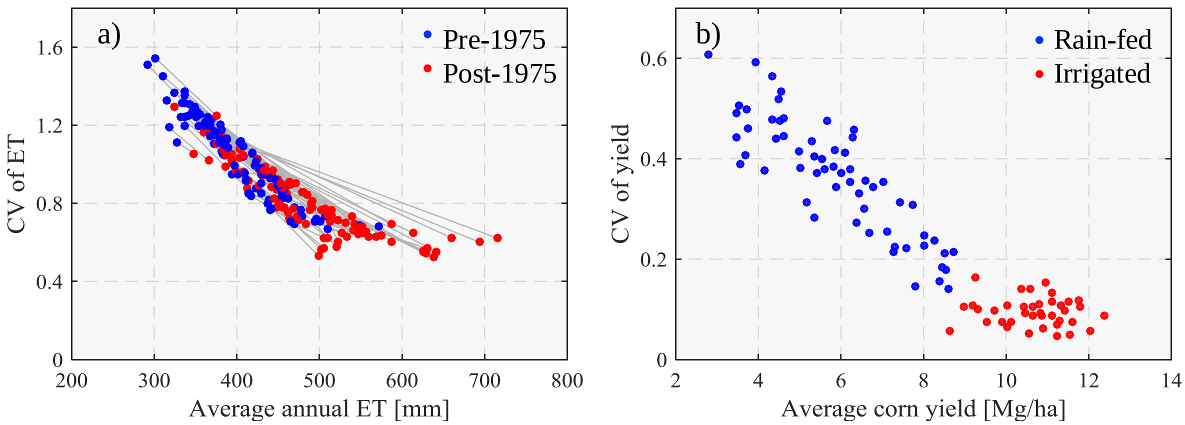


Figure 4.5. a) Annual average ET vs. *CV* for the 120 sub-basins in the High Plains, pre-1975 and post-1975; b) average vs. *CV* of rain-fed and irrigated corn yields in Nebraska.

The correspondence between ET pattern (a natural process variable) and crop yield (an anthropogenic variable) manifests as an emerging phenomenon of CNHS in the High Plains, as the nature and human components are coupled closely. Regarding the crop yield as farmers' benefit and crop yield *CV* as the risk, the transition from high-*CV*-and-low-mean to low-*CV*-and-high-mean in Figure 4.5b shows that irrigation has turned the crop system into a win-win state (higher profit and lower risk than those from rain-fed crops). Associated with such a state change of the human system, ET pattern change (Figure 4.5a) presents a hydroclimatic signal of the natural system. However, the ET change has been accompanied by some environmental consequences, i.e., water table drawdown and stream depletion in the region [*Konikow and Kendy*, 2005; *Bridget R. Scanlon et al.*, 2012; *Haacker et al.*, 2015]. Eventually the CHNS faces a tradeoff between agricultural profit and environmental sustainability, and upon reaching a tipping point, the tradeoff will cause state shifts of both nature and human systems. The 2002 Supreme Court final settlement on the Republican River Basin water rights conflict between downstream Kansas and upstream Colorado and Nebraska ended with more limited pumping permits for farmers in the upstream states, especially for those who have irrigated land along the river (e.g., the pumping permit reduces from 20 inches in the 1980s to 13.5 inches at present [*Kuwayama and Brozović*, 2013]). The regulation changes have limited the groundwater pumping for irrigation, prevented water depletion, and restored the streamflow to some extent [*Smith et al.*, 2011].

Therefore, Figure 4.5 provides a case of linkage among climatic fluctuation, engineering and socioeconomic measures, crop yield, and water table change in the context of CHNS. Farmers need to hedge their income risk against the climatic variability if groundwater pumping permit is further restricted, and institutional change might be needed (e.g., on crop insurance) to protect farmers' income. The relationship between crop yield or ET mean and *CV* provides some basic information for agricultural insurance policies designed to buffer against natural fluctuations [*Schurle*, 1996; *Glauber*, 2004].

4.5 Conclusion

Extensive groundwater pumping for irrigation has caused groundwater storage and streamflow depletion in the High Plains. To address such a concern of sustainability, this study assesses changes in ET spatial-temporal patterns in the High Plains region using climate and water table observations starting from 1940 and ETVARD, a tool to compute the inter-period ET variance. By comparing the ET patterns pre-1975 and post-1975, we find that, on average, groundwater pumping contributes about 39.1 mm to the total 74.1 mm increase in ET, while precipitation contributes about 35.0 mm. The decrease in seasonal σ_{ET} is mainly due to the decline in climatic components (20.7 mm) and offset by the increase from storage components (14.2 mm). In terms of magnitude, groundwater-fed irrigation accounts for 41% of changes in σ_{ET} after 1975. The substantial magnitude of groundwater irrigation makes the anthropogenic components in σ_{ET} comparable to the climatic factors. Before 1975, the ET spatial pattern exhibits a clear east-to-west gradient following the natural (e.g., terrain and P) gradient. After 1975, anthropogenic-induced heterogeneity overrides the natural east-to-west gradient due to localized groundwater pumping, which is affected by the various hierarchical institutional factors such as regional water rights, regulation by local natural resources authorities, and individual farmer's preferences. This study does not intend to explicitly assess the dynamics and feedbacks between nature and human components. This study focuses on the system-level signatures as the results of interactions between nature and human components. The coupling between nature and human activities is captured by the groundwater table changes. The groundwater table drawdown and fluctuation are driven by climate and farmers' response to climate, with the latter might be further constrained by groundwater depletion and social-economic factors.

We further show a statistical correspondence between the mean and *CV* of ET and that of crop yield, illustrating an analogy between ET change and crop water management. Such a correspondence manifests the coupling of the hydroclimatic and anthropogenic components in the High Plains; the tradeoff of agriculture production and environmental consequences affects the stability of the CNHS, and appropriate policies may shift the system to a more sustainable state.

CHAPTER 5

ASSESSMENT OF MULTIPLE-SOURCE AND MULTIPLE-VARIABLE EVAPOTRANSPIRATION VARIANCE OBSERVATIONS

This and the following chapter presents methods and analysis to understand ET variance by finding the congruence among theory, observation, and simulation. The theoretic ETVARD framework derived in Chapter 2 is cross-validated using ET observations and land surface model simulations. This chapter focuses on using ETVARD to constrain multi-source and multi-variable hydroclimatic observations for the contiguous United States.

5.1. Introduction

Numerous efforts have been made in hydrologic observations and simulations to advance the understanding of *ET*. At the observation side, the efforts include remote-sensing signal retrieval [*Zhang et al.*, 2010; *Mu et al.*, 2011], fluxtower network development, and data assimilation [*Pan and Wood*, 2006; *Munier et al.*, 2015; *Rodell et al.*, 2015]. Meanwhile, the land surface modelling community has developed many numerical models that include *ET* simulation, with different process representations, parameterizations, data requirements and model structures, such as the Global Land-Atmosphere Coupling Experiment (GLACE) [*Koster et al.*, 2004] and Land Data Assimilation System (LDAS) [*Rodell et al.*, 2004]. Hydrologic observations and numerical simulations play complementary and inter-dependent roles in advancing our knowledge about the various hydrological processes and systems. Hydroclimatic observations provide inputs and validation references for numerical models; meanwhile models generate data with continuous space and time dimensions, which are often used for interpolating point-scale observation [*Jung et al.*, 2009], observation network design, and conceptual validation [*Pan et al.*, 2011]. Moreover, observations also serve as the source for hydrologic concept development and hypothesis testing, such as the Budyko hypothesis on long-term *ET* [*Budyko*, 1974], the complementary relationship between actual and potential *ET* [*Brutsaert and Stricker*, 1979], and the evaporative fraction between latent heat flux and available energy [*Shuttleworth et al.*, 1989]. In turn, theoretical developments are used for observation network design and model configuration and improvement **[***Gulden et al.*, 2007; *Leung et al.*, 2015**]**. This paper assesses hydrologic data, model and theory congruence with a particular focus on *ET* in the continuous US (CONUS).

Although advances have been made in monitoring and simulating *ET* over several decades, there is a pressing need to systematically evaluate observation and model consistency and enhance their complementary outputs for hydrologic knowledge discovery [*Shuttleworth*, 2007; *Sivapalan et al.*, 2011;

Montanari et al., 2013]. Hydrologists nowadays often face a paradoxical situation: large amounts of data exist yet data uncertainty is inadequately assessed. Although grounded *in situ* measurements provide a high-resolution observation, it remains a challenge to up-scale the limited observation samples to capture the heterogeneity in a large spatial domain. Remote-sensing based *ET* products provide observation in large regions with increasing resolution, but the data retrieval processes from images (e.g., via inverse-modelling) inevitably add uncertainties to the final outputs. Therefore, when the modeling community addresses the sensitivity of model performance to forcing data or parameters [*Montanari and Di Baldassarre*, 2013; *Badgley et al.*, 2015; *Xia et al.*, 2015b], conclusions made on model evaluation [*Cai et al.*, 2014; *Swenson and Lawrence*, 2015; *Xia et al.*, 2015a; *Xia et al.*, 2016] are essentially conditioned on the quality of reference observation data. Mistakes, such as accepting a wrong model or rejecting a good model, can be made due to unreliable reference data. On the other hand, model results have been increasingly used for estimating some hydrologic variable values complementary to observation data. For example, *ET* and soil moisture simulated by land surface models are used, together with observation data, for estimating groundwater storage depletion and assessing water budget closure [*Pan et al.*, 2011]. Data scientists are concerned primarily with situations where inaccurate model results contaminate good data signal [*Beven and Freer*, 2001]. When observation, model, or both are wrong for particular studies, a good agreement between model results and observation data does not lead to a correct understanding of any hydrologic process. Under all these cases, a hydrologic theory that represents falsifiable conceptualizations of the real world is needed to justify the observation and model and diagnose the biases or errors involved in either observation or model, or both. In this paper, we adopt the Evapotranspiration Temporal VARiance Decomposition (ETVARD) framework *Zeng and Cai* [2015] as a diagnostic tool and formulate hypothetical insights about the quality of multiple *ET* products based on the various observations. This paper will focus on the assessment of observations, and the study will be extended to include model assessment in Chapter 6.

5.2 Methodology and Data Sources

Compared with the considerable progresses in data and model development, theory development and its application for model and data assessment are limited in hydrology; developing new theories and making better use of existing theories to underpin current models and data are urgently needed [*Kirchner*, 2006; *Beven*, 2012; *Clark et al.*, 2016]. In the literature, statistical methods, including recent data mining techniques [*Schnier and Cai*, 2014; *Xu et al.*, 2014], provide tools to identify the controlling factors of *ET* variance. For example, *Syed et al.* [2004] applied principal component analysis (PCA) to investigating process controls on the hydroclimatic cycle based on hydrologic data from observations and model results. However, data-driven approaches in general provide a black or grey (correlational) box without explicit physical (causal) insights. Data-driven approaches also suffer from extrapolation, since the resulting relationship is based on what has already occurred and may not correctly reflect what will occur in the future, which may lay outside the range of historical observations, especially under the situation of non-stationary hydroclimatic states due to climate change and human impacts [*Milly et al.*, 2008]. In general, physically-based methods have the advantage over data-driven methods in terms of the explanatory power and extrapolation based on predictive insights.

In addition, statistical methods, such as triple collocation [*Stoffelen*, 1998; *Pan et al.*, 2015] are used to identify the uncertainty from hydrologic observations. Although *ET* remote-sensing products are retrieved based on different methods, they generally share some inputs from the same satellite data, resulting in different degrees of correlation among these *ET* remote-sensing products. Therefore, different *ET* products are not independent from each other and do not satisfy the independence required by statistical methods. To avoid the correlation issue, *ET* observation should be validated together with other independently measured hydroclimatic variables (such as *P*, *PET*, *Q* and *S*) by examining their compatibility via a constraining hydrological principle. For example, water balance is usually used as a closure constraint for multi-variable observations [*Sheffield et al.*, 2009; *Gao et al.*, 2010a] and the Budyko water-energy relationship is applied to assessing the *ET* average and inter-annual variability in the International Satellite Land Surface Climatology Project Initiative [*Koster et al.*, 2006]. However, the constraining relationships used in previous studies are usually suitable for long-term averages of relevant hydroclimatic variables, assuming that the long-term watershed system storage remains stable. This assumption is invalid for assessing variability at a relatively short time scale (annual or monthly) and for watersheds with systematic terrestrial storage change over a long-term period. In the present study on the congruence among theory, data, and model, we adopt ETVARD as a diagnostic tool, as introduced in the following.

5.2.1 ET temporal variance decomposition

Introducing terrestrial water storage change as an variable in watershed water balance, *Zeng and Cai* [2015] extended the Budyko relationship and, based on which, an equation to decompose ET temporal variance into multiple contributing components as shown below:

$$\sigma_{ET}^2 = w_P\sigma_P^2 + w_{PET}\sigma_{PET}^2 + w_{\Delta S}\sigma_{\Delta S}^2 + w_{P,PET}cov_{P,PET} + w_{P,\Delta S}cov_{P,\Delta S} + w_{PET,\Delta S}cov_{PET,\Delta S} \qquad \text{Eqn.(5.1)}$$

where σ represents the standard deviation; cov represents the covariance; w represents the weighting factors, which quantify the contribution from different variance/covariance sources to *ET* variance. The weighting factors, as shown in Figure 5.3, can be calculated from the aridity index ($\bar{\phi} = \overline{PET}/\bar{P}$), Budyko equation $F(\bar{\phi})$ [*H Yang et al.*, 2008] and its first-order derivative $F'(\bar{\phi})$, which is detailed in *Zeng and Cai* [2016] and *Zeng and Cai* [2015]. By Eqn.(5.1), ET variance is contributed by long-term climatic condition (through the weighting factors), climatic fluctuations (σ_P^2 and σ_{PET}^2) and phasing ($cov_{P,PET}$), hydrologic storage variability ($\sigma_{\Delta S}^2$) and its response to climate ($cov_{P,\Delta S}$ and $cov_{PET,\Delta S}$). ETVARD provides an analytic way to decompose ET variance into climatic and hydrologic components and offers an independent estimate of ET variance based on meteorological and catchment storage data.

We can further aggregate the ET variance components based on their sources into two categories: One represents the contribution to ET variance from the variability of climatic forcing (σ_{ETF}^2) and the other from hydrologic storage (σ_{ETS}^2), i.e.

$$\sigma_{ETF}^2 = w_P\sigma_P^2 + w_{PET}\sigma_{PET}^2 + w_{P,PET}cov_{P,PET} \qquad \text{Eqn.(5.2)}$$

$$\sigma_{ETS}^2 = w_{\Delta S}\sigma_{\Delta S}^2 + w_{P,\Delta S}cov_{P,\Delta S} + w_{PET,\Delta S}cov_{PET,\Delta S} \qquad \text{Eqn.(5.3)}$$

Eqn. (5.1) becomes

$$\sigma_{ET}^2 = \sigma_{ETF}^2 + \sigma_{ETS}^2 \qquad \text{Eqn.(5.4)}$$

The time scale of ET variance depends on the time scale of the various variance/covariance terms. This study addresses ET variance at the monthly scale, while the analysis on ET variance at both annual and monthly scale can be found in *Zeng and Cai* [2016].

ETVARD relates the ET temporal variance (σ_{ET}^2) to the variance and covariance of climatic variables (e.g., *P*, *PET*) and hydrological system variables (e.g. *S*, which can be changed by climate, land use and water use). By assessing $\sigma_{ET}^2, \sigma_{ETF}^2$ and σ_{ETS}^2 by grid (1° by 1° according to the GRACE grids) in the CONUS, the spatial patterns of the ET temporal variance are calculated. The assessments based on ETVARD will be used as a reference, and those from multiple ET products are compared to the reference, by which the possible bias and uncertainty involved in each of the ET products and their spatial patterns will be discussed.

5.2.2 Multi-source multi-variable hydroclimatic observations

This study uses monthly meteorological forcing data (*P* and *PET*) obtained from the North American Land Data Assimilation System Phase 2 (NLDAS-2) [*Mitchell et al.*, 2004; *Xia et al.*, 2012b]. *P* in NLDAS-2 is a product of gauge-only NOAA Climate Prediction Center, which conducted orographic adjustment of daily precipitation based on the PRISM climatology. The non-precipitation land-surface forcing fields for NLDAS-2 are derived from the analysis fields of the NCEP North American Regional Reanalysis (NARR) and further vertically adjusted to account for the vertical difference between the NARR and NLDAS fields of terrain height. *PET* is calculated from modified Penman scheme [*L Mahrt and Michael Ek*, 1984] from the land-surface forcing fields for NLDAS-2. The same forcing data fields are also used to drive the NLDAS-2 land surface models, which will be discussed in detail in Chapter 6.

P and *PET* from NLDAS-2 forcing data sets have a spatial resolution of 0.125° by 0.125° and cover the period from 1979 to 2015.

The terrestrial water storage (TWS) measured by the twin GRACE satellites is based on the distance change between the two satellites due to gravity field variation [*Tapley et al.*, 2004]. GRACE satellites primarily capture the mass change caused by TWS since other temporal changes of mass are negligible. GRACE based TWS includes the sum of storage in various media such as aquifer, soil profile, snow/glacier and surface reservoir/lake. The GRACE satellites provide a unique measurement of TWS with a large spatial coverage [*Lettenmaier and Famiglietti*, 2006] and have been widely applied for hydrologic studies such as groundwater depletion assessment [*Famiglietti et al.*, 2011], water budget closure estimation [*Pan et al.*, 2011] and land surface model improvement [*Gulden et al.*, 2007] and evaluation [*Cai et al.*, 2014; *Xia et al.*, 2017]. The GRACE TWS data set used in this study, with a spatial resolution at 1° by 1°, provides a monthly time series from January of 2003 to June of 2013. The monthly terrestrial storage change (*ΔS*) is calculated as the difference from the monthly GRACE TWS time series [*Landerer and Swenson*, 2012] based on the CSR RL5.0 release from the Center for Space Research at the University of Texas at Austin.

5.2.3 Multiple ET products based on observations

Three *ET* products, two based on remote-sensing observation and one based on FLUXNET are used in this study. The remote-sensing product by *Qiuhong Tang et al.* [2009] calculates *ET* as the combination of bare soil evaporation and vegetation transpiration based on the constant daily evaporative fraction assumption [*Shuttleworth et al.*, 1989]. Bare soil evaporation is estimated from surface radiation budget and soil temperature, and vegetation transpiration is calculated using the complementary relationship to bridge the actual *ET* with potential evaporation calculated by Priestley-Taylor scheme. The data set covers the extent of the CONUS from 2001 to 2008 at a spatial resolution of 0.05° by 0.05° and is denoted as "ET_{RS-UW}" in this study. The data and more details about the methodology can be find at Evaporation Estimation Using Remote Sensing at University of Washington. This *ET* product has been used to assess watershed water budget [*Gao et al.*, 2010a] and *ET* interannual variability [*Cheng et al.*, 2011].

Another *ET* product by *Mu et al.* [2011] calculates *ET* from vegetation transpiration and soil evaporation based on the Penman-Monteith scheme presented in *Mu et al.* [2007]. Vegetation evaporation is further separated into wet canopy surface evaporation and dry canopy vegetation transpiration, and the rates are regulated by aerodynamics resistance and surface resistance. Soil evaporation is divided into saturated soil potential evaporation and moist soil evaporation that is constrained by soil moisture stress. The monthly version global *ET* from 2000 to 2009 at 0.5° by 0.5° spatial resolution is obtained from MOD16 Global Terrestrial Evapotranspiration Data Set and denoted as

“$ET_{RS-MOD16}$” in this study. The *ET* product has been applied for many studies such as drought assessment [*Mu et al.*, 2013] and land surface model evaluation [*Cai et al.*, 2014].

The third *ET* product developed by *Jung et al.* [2009] is different from the remote-sensing products in terms of both data sources and retrieval algorithms. It is essentially the spatial up-scaling of point measurements from eddy covariance flux tower. This approach uses model tree ensemble, a machining learning technique, to up-scale current global network of eddy covariance towers (FLUXNET) and evaluates results from the “virtual reality” produced by Lund-Potsdam-Jena managed Land biosphere model simulation. This *ET* estimate has been applied for *ET* trend analysis [*Jung et al.*, 2010] and land surface model improvement [*Bonan et al.*, 2011] and evaluation [*Cai et al.*, 2014]. The global monthly *ET* estimate from 1982 to 2008 at 0.5° by 0.5° spatial resolution is denoted as “$ET_{FLUX-MTE}$” in this study.

Table 5.1. Data sources

Variables	Source	Spatial resolution	Temporal coverage
P	NLDAS-2	0.125° by 0.125°	1979-2015
PET	NLDAS-2	0.125° by 0.125°	1979-2015
ET	Tang et al. [2009]	0.05° by 0.05°	2001-2008
	Mu et al. [2011]	0.5° by 0.5°	2000-2009
	Jung et al. [2009]	0.5° by 0.5°	1982-2008
ΔS	GRACE	1° by 1°	2013.01-2013.06

Since NLDAS-2 meteorological forcing, GRACE TWS, and the *ET* products have different spatial resolutions and temporal coverage, these datasets are processed to calculate *ET* variance using the following procedures: First, the time series from all data sets are spatially aggregated and matched at the 1° by 1° GRACE grids to for the Contiguous United States domain of latitude between 25°N and 53°N and longitude 67°W and 125°W. At the 1° by 1° spatial resolution, we assume the terrestrial water storage change caused by groundwater lateral flow is negligible. Second, the weighting factors in Eqn.(5.1) are calculated from long-term average climate condition based on NLDAS-2 *P* and *PET* from 1979-2015. Third, climatic variabilities (i.e., σ_P, σ_P and $cov_{P,PET}$ in Eqn.(5.2)) are calculated from monthly time series during 1979-2015 and the variabilities associated with storage change (i.e., $\sigma_{\Delta S}$, $cov_{P,\Delta S}$ and $cov_{PET,\Delta S}$ in Eqn.(5.3)) are calculated from monthly time series during the period of January of 2003 to June of 2013 during which GRACE is available. *ET* variance from direct observations are calculated from their temporal coverages. We assume the variance and covariance terms are statistically stationary during the relatively short period.

Figure 5.1 shows the schematic diagram to calculate σ_{ET}^2 through multiple paths using observations, numerical models and ETVARD. σ_{ETF}^2, the climatic forcing components of σ_{ET}^2 (Eqn.5.2) is calculated using *P* and *PET* data from NLDAD-2. σ_{ETS}^2, the hydrologic components of σ_{ET}^2 (Eqn.5.3) is calculated using *ΔS* estimated from GRACE, denoted as $\sigma_{ETS-GRACE}^2$. It might be noted that compared to existing studies by data-driven methods such as the principle component analysis [*Syed et al.*, 2004], ETVARD calculates *ET* variance based on the climatic forcings, terrestrial storage change, and their correlations (Eqn.5.1) without direct *ET* observation inputs.

5.2.4 Experiment design

We design Experiment 1 by setting σ_{ETVARD}^2=σ_{ETF}^2+$\sigma_{ETS-GRACE}^2$ using *P* and *PET* observations from NLDAD-2 and GRACE-based *ΔS*. Through this experiment, we will investigate the climatic and hydrologic contributions in σ_{ET}^2 and their spatial patterns for the CONUS. We interpret the σ_{ET}^2 components by ETVARD and assess the relative role of climatic variables and hydrologic system variables (including human interference).

We design Experiment 2 by comparing σ_{ETVARD}^2 calculated from Experiment 1 to σ_{ET}^2 from the three observation-based ET products ($\sigma_{RS-UW,}^2$ $\sigma_{RS-MOD16}^2$ and $\sigma_{ET-FLUX}^2$). Through this experiment, we compare the three estimates to σ_{ETVARD}^2 at particular locations of interest and their spatial distribution across the CONUS.

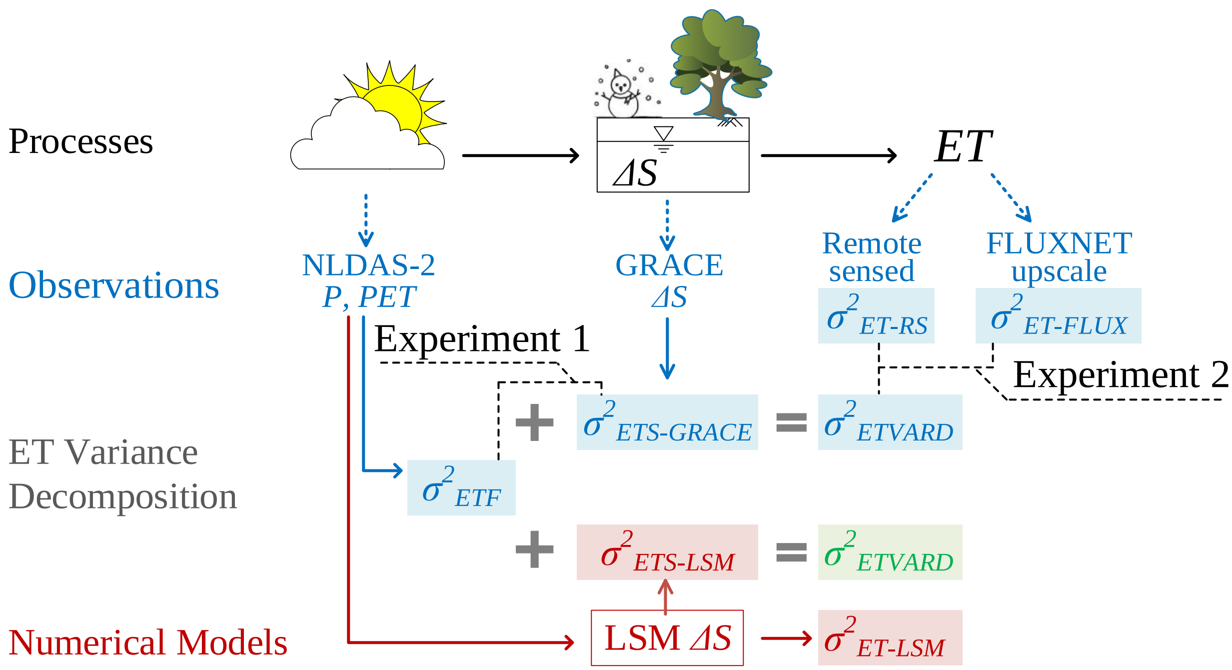


Figure 5.1. Schematics of hydrologic processes along with various ET variance estimates and its components from observation, simulation and ETVARD approaches. Variables that are represented in black for true but unknown values, in blue for quantities from direct or indirect observation data, in red for quantities by numerical models and in green for "hybrid" quantities from both observation and model results. Blue dash lines represent for data acquisition which is inevitably subject to observation and/or processing error. Solid lines represent for information propagation indirection. Black dash lines illustrate the pairs of quantities assessed in each experiment in the two companion papers. Note that $\boldsymbol{\sigma^2_{ETS-LSM}}$ denotes $\boldsymbol{\sigma^2_{ETS}}$ that is estimated with *ΔS* simulated by LSMs in Chapter 6.

5.3 Results

5.3.1. Spatial patterns of climatic and hydrologic variability

The climatic and hydrologic variance/covariance at monthly scale are shown in Figure 5.2. σ_P^2 exhibits significant heterogeneity across CONUS, as displayed in Figure 5.2a. σ_P^2 is very high in coastal regions of Washington, Oregon and California, north shore of Gulf of Mexico and the region between Lake Huron and Lake Ontario. σ_P^2 is high across the High Plains and low in the Mountain States. Scattered locations in north Idaho and Montana also have high σ_P^2. σ_{PET}^2 shows contrast pattern in the east and west and is roughly divided by 100°W longitude line, as shown in Figure 5.2b. The humid east has relatively stable *PET*; while the arid west experiences large fluctuations in *PET*. The coastal regions of Washington and Oregon and scattered spots in the Mountain States have low σ^2_{PET}. $cov_{P,PET}$ indicates the phasing between water and energy supply as shown in Figure 5.2c. Positive $cov_{P,PET}$ means the warm months coincides with the wet months, and vice versa. $cov_{P,PET}$ is high in the Midwest where the warm season comes together with rainfall season, providing a good climate condition for agricultural production. The west coastal region has negative $cov_{P,PET}$due to its Mediterranean climate. The central part of the South along Louisiana, Mississippi and Alabama also have a slightly out-of-phase *P~PET* pattern due to the conjunctive effects of subtropical jet stream from the north and moisture from the gulf of Mexico.

$\sigma^2_{\Delta S}$ calculated from GRACE data shows noteworthy fluctuations in the north part of the Pacific Northwest, the Mississippi Embayment Aquifer and the Midwest, as shown in Figure 5.2d. Since the terrestrial storage change measured by GRACE includes storage change in the aquifer, soil, snow/glacier and river, the $\sigma^2_{\Delta S}$ at monthly scale may be caused by different local processes. Note the very large value of $\sigma^2_{\Delta S}$ along the downstream of Mississippi River is caused by flood storage, which is routed from upstream of Mississippi and caused by lateral flow across the cells. $cov_{P,\Delta S}$ represents the storage change caused by catchments' response to *P* fluctuation (e.g., aquifer replenishment due to excessive rainfall, groundwater pumping for irrigation during drought or snow/glacier accumulation). Figure 5.2e shows that $cov_{P,\Delta S}$ is high in the north part of the western coast and moderate in the Mississippi Valley, meaning that *ΔS* increases during rainfall season. $cov_{P,\Delta S}$ is slightly negative in the High Plains, indicating that catchment storage decreases even during the rainfall season. Since the *PET* is in-phase with *P* in these regions, the *ΔS* decrease during rainfall season is caused by groundwater pumping to meeting the high evaporative demand of crops in this region. The high ET flux during high *PET* months consumes water from both *P* and *ΔS*, which is also evidenced by the negative $cov_{PET,\Delta S}$ in these region, as shown in Figure 5.2f. In addition, $cov_{PET,\Delta S}$ also characterizes the energy-induced storage change by the snow and glacier thaw/melting processes, especially in the west mountainous regions. It is noted that the cells long the coast reflect the blend GRACE signal of mass change caused by both terrestrial and sea water, which may underestimate the magnitude of *ΔS* in these cells.

5.3.2. Spatial patterns of weighting factors

The climatic and hydrologic provides the sources for ET variance, however, they do not equally contribute to ET variance. Their contribution to ET variance is further restrained by the long-term climate

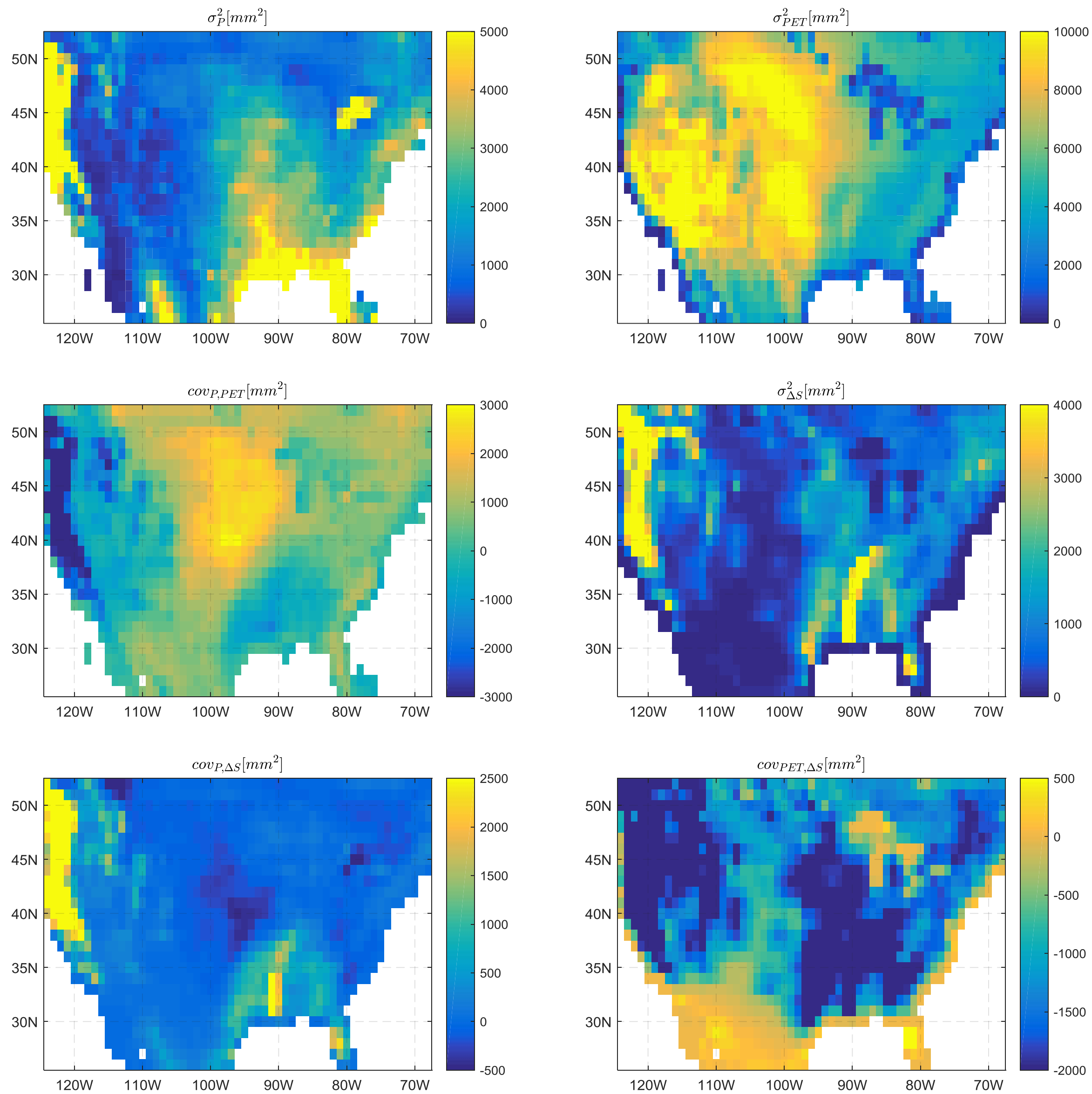


Figure 5.2. Climatic and hydrologic variability and their covariances

condition represented by Eqn.(5.1). The weighting factors essentially extend the Budyko hypothesis on hydroclimatic control ET from the long-term average to variability. For example, in a very arid condition when the hydrologic processes are constrained by water supply, even although *PET* monthly variance may be big, its contribution to ET variance would be small. The analytical range of the weighting factors are evaluated in *Zeng and Cai* [2015]: w_P, w_{PET} and $w_{\Delta S}$ range between 0 and 1; $w_{P,PET}$ lies between 0 and 0.35; $w_{P,\Delta S}$ ranges between -2 and 0; and $w_{PET,\Delta S}$ ranges between -0.35 and 0.

Note that w_P and $w_{\Delta S}$ are the same and their spatial pattern is shown in Figure 5.3a and S2d. These two weighting factors shows apparent east-west gradient from the humid east to arid humid, indicating that fluctuations in atmospheric and catchment water supply would increasing contribute to ET variance as climate becomes drier. The western coast and the Central Valley in California have low w_P and $w_{\Delta S}$. w_{PET} is low for majority of CONUS except in the coastal region of Washington and Oregon and the Great Lakes, as show in Figure 5.3b, since w_{PET} rapidly as climate becomes dryer. $w_{P,PET}$ determines the contribution from *P~PET* phasing to ET variance, as shown in Figure 5.3c. Since $w_{P,PET}$ is positive, the in-phase seasonality (i.e., positive $cov_{P,PET}$) will enhance ET variance, and vice versa for out-of-phase seasonality. The in-phase *P~PET* provides a favorable condition for ET, allowing both *P* and *PET* variability added to ET variance. $w_{P,PET}$ is large mainly in the East, western coast and the Central Valley in California.

$w_{P,\Delta S}$ is negative in CONUS and its magnitude is large in the arid western regions, as shown in Figure 5.3e. In these water limited regions, a negative $cov_{P,\Delta S}$ (representing storage release during rainfall periods) will effectively enhance ET variance. From the perspective of ecohydrology, it is critical for vegetation to manage *ΔS* (in soil profile or shallow aquifer) corresponding to P fluctuation to satisfy the evaporative demand; on the other hand, vegetation has limited capability to regulate *ΔS*, thus the ET variance would not exactly follow the *PET* fluctuation. Furthermore, as irrigated agriculture is expanding in water limited region, the conjunctive management of surface-ground water for irrigation (as a way to change *ΔS*) may substantially reform the ET variance. Similarly, a positive $cov_{P,\Delta S}$ (representing storage recover during rainfall periods) will effectively dampen ET variance in the arid region, since the recharge into *ΔS* results in less water available for ET. $w_{PET,\Delta S}$ is opposite to $w_{P,PET}$ by noticing they differ in sign in Eqn.(5.3), as shown in Figure 5.3f.

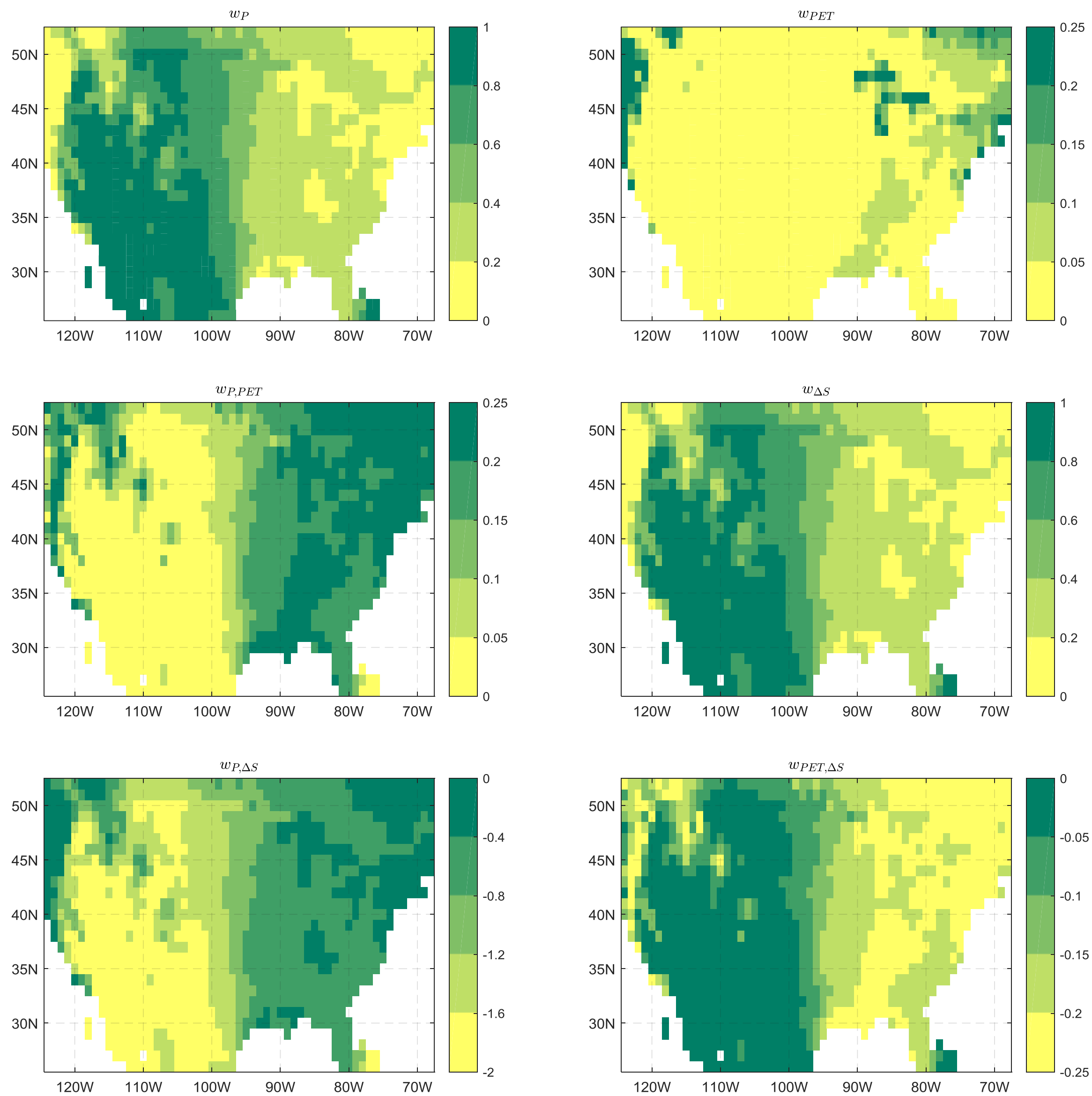


Figure 5.3. Weight factors determining in the contribution of climatic and hydrologic variability to σ_{ET}^2.

5.3.3 Experiment 1: σ_{ET}^2 components in the CONUS

Figure 5.4 (a, b and c) display the magnitudes and spatial distribution of σ_{ET}^2 components from climatic variables *P*, *PET*, and their phase, respectively. As can be seen from Figure 5.4a, the contribution from *P* ($w_P\sigma_P^2$) is more than 2000 mm^2 in California and southern Florida. In the High Plains, $w_P\sigma_P^2$ is also overall significant (around 1500 mm^2) and decreases gradually from south to north. $w_P\sigma_P^2$ is small (less than 500 mm^2) in the Mountain States and negligible above the Great Lakes (which is due to the fact that ET is from the water surface, and the fluctuation in *P* does not affect the ET variance much). The contribution from PET variability ($w_{PET}\sigma_{PET}^2$) is relatively small compared to $w_P\sigma_P^2$ and exhibits a sharp contrast along the east-west direction, as shown in Figure 5.4b. The Mountain States (west of 97th Meridian West) have negligible $w_{PET}\sigma_{PET}^2$; however, the coastal regions of Washington and Oregon states, parts of the California Central Valley, and the Great Lakes have significant amount of contribution from PET variability (more than 500 mm^2). In these regions, ET is limited by energy supply, and fluctuation in PET dominates the ET variance. In addition, the northeastern region has a visible $w_{PET}\sigma_{PET}^2$ component (between 200 and 400 mm^2). Figure 5.4c shows that the in-phase of P~PET enhances ET variance in the Corn Belt; while the out-of-phase P~PET reduces ET variance in the coastal regions of Washington, Oregon and California Central Valley due to the Mediterranean climate in those regions.

Figure 5.4d-f display the magnitudes and spatial distribution of the contribution to σ_{ET}^2 from the variance of ΔS, the covariance of P and ΔS, and the covariance of PET and ΔS, respectively. $w_{\Delta S}\sigma_{\Delta S}^2$ is more than 1000 mm^2 in the Pacific Northwest and California and more than 500 mm^2 in the south part of the High Plains and Mississippi Embayment region. As can be seen from Figure 5.4e, the interaction between P and ΔS ($w_{P,\Delta S}cov_{P,\Delta S}$) significantly reduces σ_{ET}^2 in the western coast especially in California (2000 mm^2) but slightly enhances σ_{ET}^2 (around 500 mm^2) in the North High Plains and part of the East. Although $cov_{PET,\Delta S}$ is significant in the West (Figure 5.2f), its contribution to ET variance concentrates to a limited region in California due to a low weighting factor, as shown in Figure 5.4f. The South and the Appalachian Mountains also have fairly significant $w_{PET,\Delta S}cov_{PET,\Delta S}$ component (more than 300 mm^2).

Adding the climatic components together by Eqn.(5.2), σ_{ETF}^2, the overall ET variance from the climate variables, is shown in Figure 5.5a. Generally, the distribution of σ_{ETF}^2 follows that of precipitation in most places. In the Corn Belt, the in-phase of P~PET provides a favorable condition for crop water consumption, yielding a relatively large σ_{ETF}^2 (more than 1500 mm^2). The coastal regions in Washington and Oregon have relatively mild σ_{ETF}^2, since the out-of-phase P~PET in those regions dampens ET variance. The Appalachian Mountains have low σ_{ETF}^2 (less than 1000 mm^2); the Mountain States have the lowest σ_{ETF}^2 (less than 500 mm^2) in magnitude.

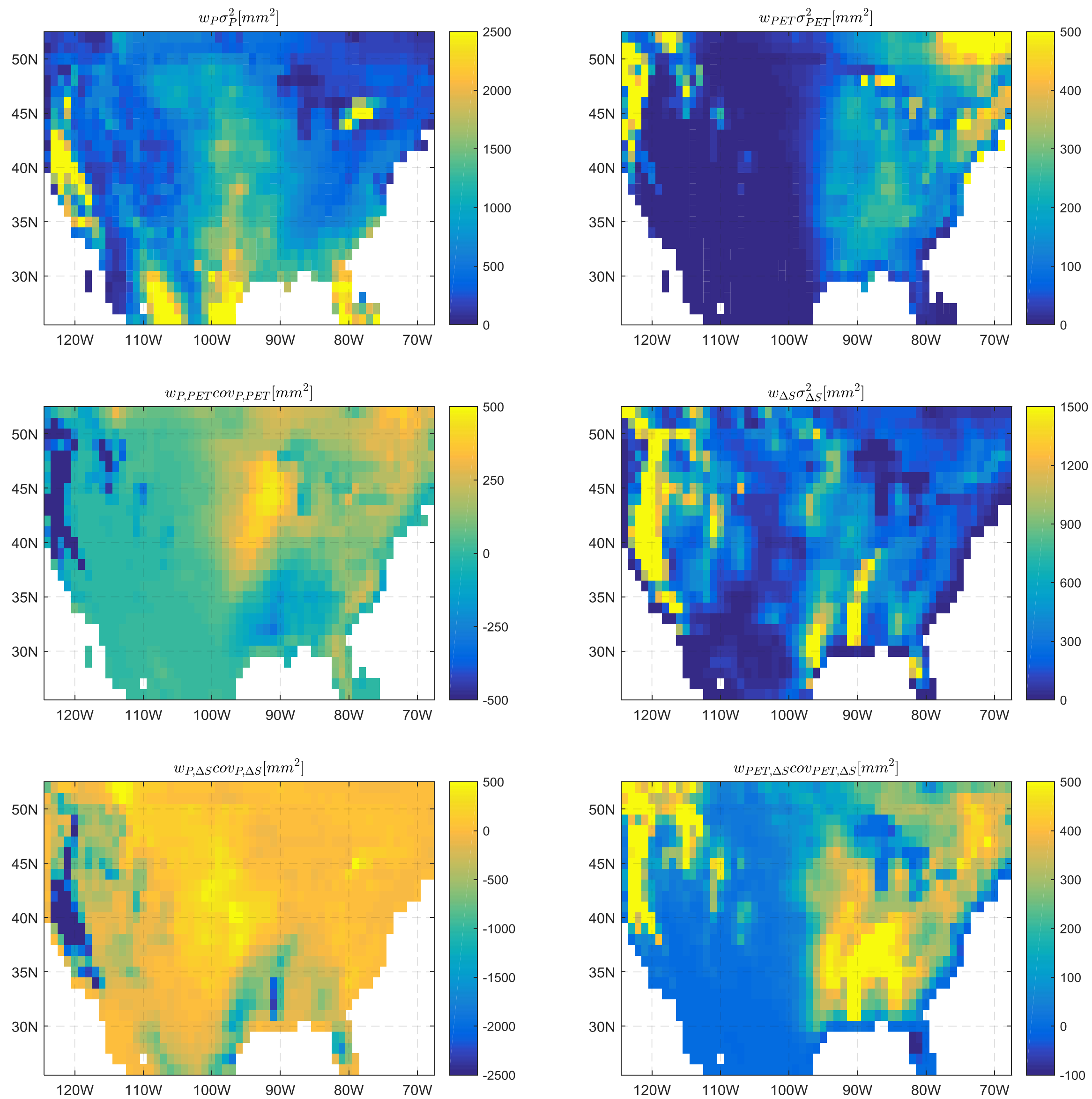


Figure 5.4. Individual components of σ_{ET}^2 derived from ETVARD

The aggregated hydrologic system components of ET variance σ_{ETS}^2 by Eqn.(5.3) is shown in Figure 5.5b. Given that σ_{ETS}^2 denotes the effect of catchments' responses to climate and human interference, it can be negative (i.e., a dampening effect) or positive (i.e., an enhancing effect). In general, the magnitudes of σ_{ETS}^2 is smaller than those of σ_{ETF}^2, indicating the major impact of climatic variance in general. However, σ_{ETS}^2 enhances ET variance (more than 1000 mm^2) over the High Plains and Mississippi downstream and reduces ET variance (more than 500 mm^2) in California Central Valley. These regions with strong σ_{ETS}^2 components overlap with major aquifers that have been depleted for irrigation [*Konikow*, 2015], Anthropogenically induced storage change either enhances or dampens ET variance depending on the local climate and has higher signals in ET variance than the natural vegetation. The Cascade Range and northern part of the Rocky Mountains also have positive σ^2_{ETS}, mainly because the snow accumulating and melting processes provide a temporal redistribution of water from cold to warm seasons.

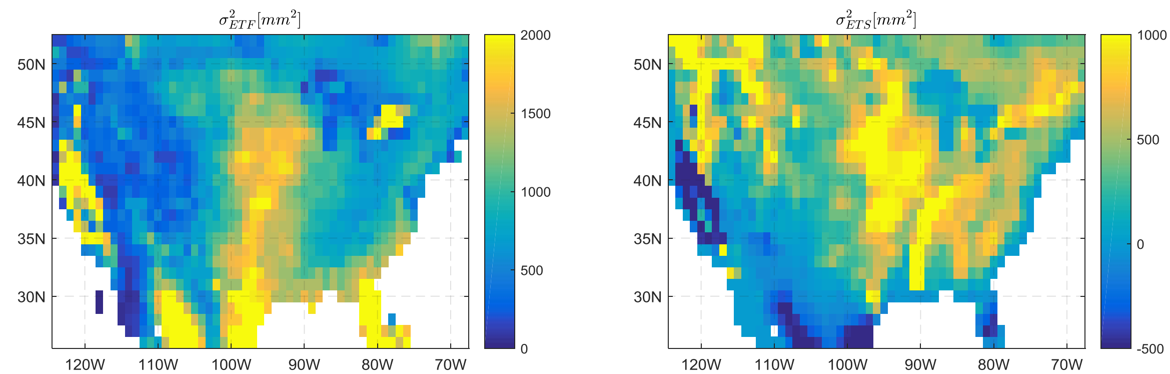


Figure 5.5. Climatic and hydrologic components of σ^2_{ET}

5.3.4 Experiment 2: Multi-source σ^2_{ET} comparison

Since the total ET variance is all positive, the following analysis on ET variance from multiple observations is assessed in terms of ET standard deviation. σ_{ETVARD} sums up the climaitic and hydrologic components from Experiment 1 (Eqn.1 or Eqn.4). In Figure 5.6a, σ_{ETVARD} ranges between 0 mm to 60 mm; the maximum σ_{ETVARD} occurs acrossing the High Plains and dereases toward west with the minimum located along the east of Sierra Nevada Mountains. Florida also has noticeable σ_{ETVARD} (above 40 mm); the western coastal region and the Appalachian-Northeast line also have moderate σ_{ETVARD} (around 30 mm). As shown in Figure 5.6b, the remote-sensing σ_{RS-UW} exhibits similar spatial zonation to σ_{ETVARD}, with the peak value in the Midwest and the coastal region of the North Pacific. σ_{RS-UW} is in general larger than 40 mm on other parts of the East and less than 30 mm in the Mountain States, with the minimum located along the east of Sierra Nevada Mountains. Remote-sensing based $\sigma_{RS-MOD16}$ shows a contrast west-east spatial pattern (Figure 5.6c.) The $\sigma_{RS-MOD16}$ in western CONUS is mostly below 20mm, which is smaller than that from σ_{ETVARD} or σ_{RS-UW} (above 20 mm in the region.) However, the northern pacific coast is exceptional with $\sigma_{RS-MOD16}$ around 30 mm. $\sigma_{RS-MOD16}$ is about 10 mm smaller than σ_{ETVARD} or σ_{RS-UW} in the western CONUS. The peak values (larger than 50 mm) of $\sigma_{RS-MOD16}$ are located along the downstream of the

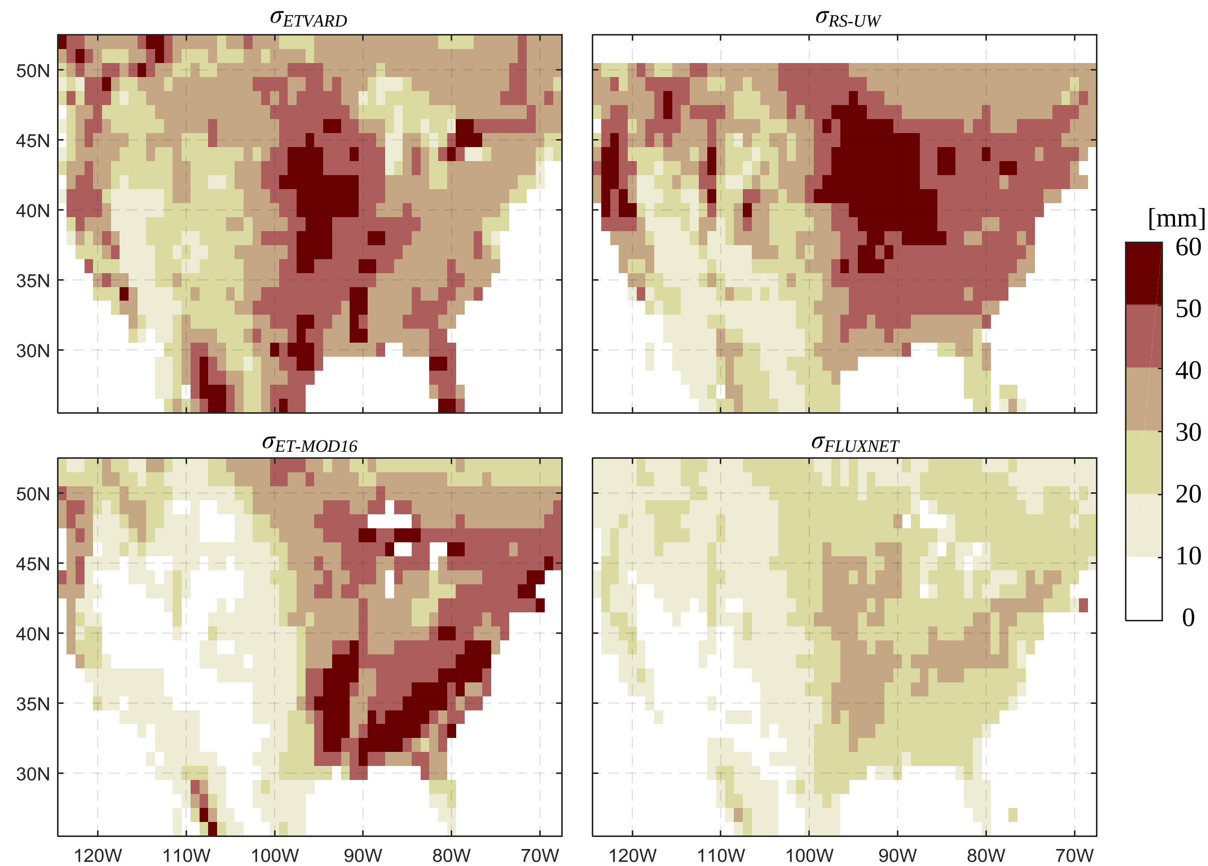


Figure 5.6. Spatial pattern of total ET variance by four observation-based estimates

Mississippi River and the Southeast. The Midwest and Northeast has moderate $\sigma_{RS-MOD16}$ between 30mm and 50mm. The FLUXNET up-scaling estimate $\sigma_{FLUX-MTE}$ is shown in Figure 5.6d. Apparently, $\sigma_{FLUX-MTE}$, ranging s between 0 to 40 mm, is smaller than the other three estimations. The spatial distribution of maximun $\sigma_{FLUX-MTE}$ is similar to that of σ_{ETVARD}, extending from the Midwest to the south part of the High Plains. The Appalachian-Northeast line also has substential $\sigma_{FLUX-MTE}$; the western CONUS has $\sigma_{FLUX-MTE}$ generally below 20mm, which shows a similar range to that of $\sigma_{RS-MOD16}$.

It is not surprising to see the discrepency of spatil patterns of ET variance from these four estimates, but it is difficult to draw the conclusion on which product is more reliable than others, since the ture value is not known. In general, σ_{ETVARD} and σ_{RS-UW}, the two independent estimates yield similar spatial patterns and magnitudes. $\sigma_{FLUX-MTE}$ can be underestimated, compared to other three products. This is probbaly because the flux tower sites are too sparse to capture the heterogeneity of ET for a large region. Errors in σ_{ETVARD} may exist at coastal grids where GRACE-estimated ΔS contains signals of sea water.

Another purpose of Experiment 2 is to assess the compatiblity of a set of multi-variable (i.e., P, PET, ET and ΔS) observations under the theoretical ETVARD framework, by assessing the residual between σ_{ETVARD} and the other three estiamtes. The frequency histograms of the residuals between σ_{ETVARD} and σ_{RS-UW}, $\sigma_{RS-MOD16}$ or $\sigma_{FLUX-MTE}$ are plotted in Figure 5.7. As can be seen in Figure 5.7a, the residual between σ_{RS-UW} and σ_{ETVARD} fits a Gaussian distribution with a mean of 0.52 mm and standard deviation of 11 mm. The small residual (i.e., σ_{RS-UW} - σ_{ETVARD}) indicates that this set of multi-variable hydroclimatic observations (i.e., NLDAS-2 P and PET, GRACE-estimated ΔS, and ET_{RS-UW}) that are used to calculate σ_{ETVARD} are statistically unbiased under the general laws embedded in ETVARD.

The residual between $\sigma_{FLUX-MTE}$ and σ_{ETVARD} as shown in Figure 5.7b yields a Gaussian distribution with mean of -15 mm and standard deviation of 8.1 mm. The relatively small residual standard deviation indicates $ET_{FLUX-MTE}$ may have relatively smaller uncertainty than the other two ET products, while the large residual mean indicates that $\sigma_{FLUX-MTE}$ is probably underestimated. The residual between $\sigma_{RS-MOD16}$ and σ_{ETVARD} as plotted in Figure 5.7c yields a slightly bi-modal distribution, and a Gaussian fit results in a mean of -7.8 mm and standard deviation of 15 mm, the largest uncertainty among the three observation based ET products, when σ_{ETVARD} is used as a reference.

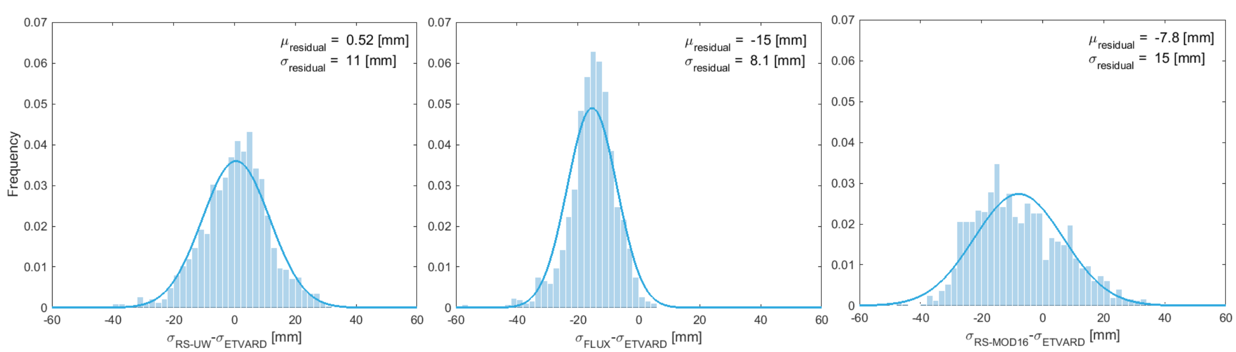


Figure 5.7. Residual total ET variance from different sets of hydrologic variables under ETVARD

5.4 Discussion

5.4.1 The clustering of most important components of *ET* variance

Zeng and Cai [2015] qualitatively divided the $(\bar{P}, \overline{PET})$ plane into several zones with various controlling factors on σ_{ET}^2 based on the weighting factors. Here, we take the largest absolute value of the six components in each grid (Figure 5.6), and identify that as the most important controlling component of ET variance. Those identifications are plotted in the $(\bar{P}, \overline{PET})$ plane as shown in Figure 5.8. It confirms that in the CONUS P and PET are the major controls of σ_{ET}^2 in arid ($\phi > 1$) and humid regions ($\phi < 1$), respectively. The major components associated with ΔS are located in the lower-left region, where the water and energy fluxes have relatively small values (approximately, $\bar{P}$<1000 mm, and $\bar{P} + \overline{PET}$<2200 mm). The empirical threshold exists since the catchment storage has relatively limited capacity to buffer the water and energy fluctuations.

Exceptionally, several major components associated with storage (ΔS, and $P\&\Delta S$) are far beyond the thresholds as shown in Figure 5.8. These points represent the major components of California or the areas along the lower reaches of the Mississippi River. The deviation of the largest components in these regions is mainly due to the water storage change by agricultural water uses, which have significantly larger capability to use storage (e.g., pumping groundwater or surface water storage) than natural vegetation. This confirms that when human water use significantly affects the ET process, and the largest components of σ_{ET}^2 are deviated from those in natural catchments. Thus, the $(\bar{P}, \overline{PET})$ plane provides a visual diagnostic tool to detect human interferences on σ_{ET}^2.

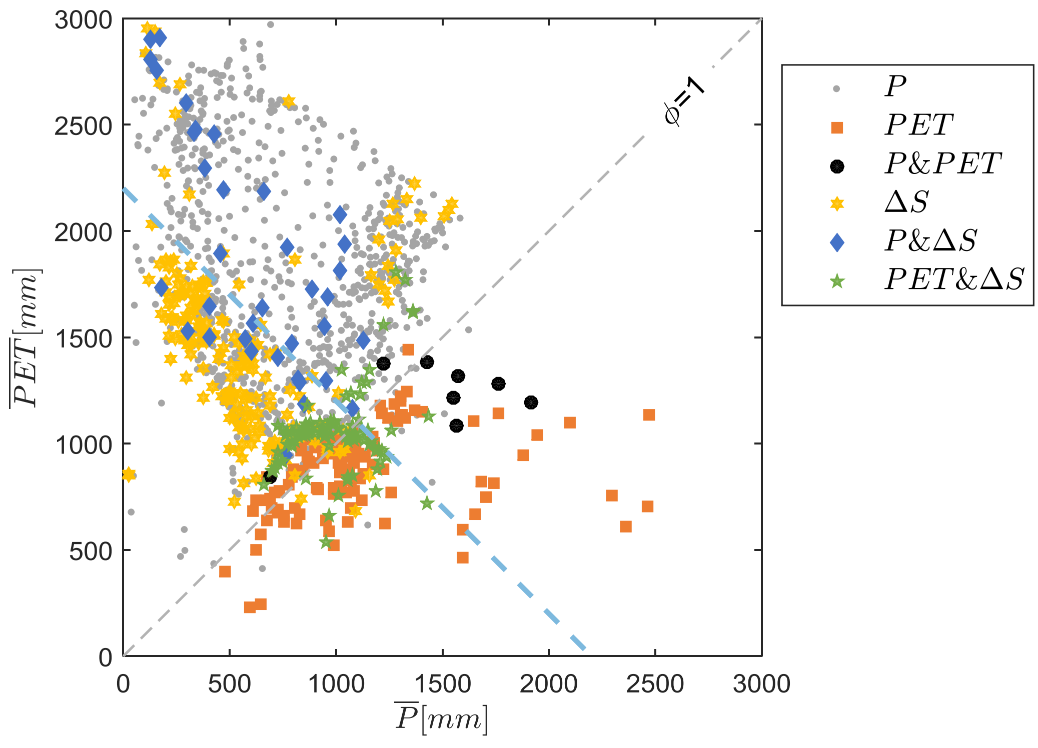


Figure 5.8. Largest σ_{ET}^2 component in each grid in the CONUS in the $(\bar{P}, \overline{PET})$ plane

5.4.2 Implication of σ_{ET}^2 components for model development

The climatic components in Eqn.(5.2) and storage components in Eqn.(5.3) of σ_{ET}^2 provide valuable information for hydrologic model development in terms of increasing the accuracy of model inputs and the improvement of model structures. For regions where σ_{ET}^2 climatic components are significant (as shown in Figures 5.4a-c and Figure 5.5a), more reliable model input fluxes (i.e., *P* and *PET*) would improve the model performance. For example, σ_{ET}^2 in the western CONUS is not significantly affected by *PET* (Figure 5.4b). Therefore, the hydroclimatic processes and models in this region may not need to be sensitive to the fluctuations in *PET*. On the other hand, improving the model structure to better capture how hydrologic state variable *S* (e.g., snow, soil moisture and groundwater) responds to climate is important in regions where σ_{ET}^2 storage components are significant (Figures 5.4d-f and Figure 5.5b). For example, $w_{P,\Delta S}cov_{P,\Delta S}$ represents catchments' response (both natural and anthropogenic) to *P*, such as groundwater recharge and pumping. Agricultural irrigation enhances the σ_{ET}^2 in the High Plains and dampens the σ_{ET}^2 in California (Figure 5.4e.) Therefore, farmers' irrigation behavior should be reasonably represented in the models developed for these regions. $w_{PET,\Delta S}cov_{PET,\Delta S}$ represents catchments' response to *PET*, such as snow melting and vegetation water demand. Figure 5.4f indicates that the snow dynamics in north pacific coast and vegetation dynamics in Eastern CONUS are important processes controlling the σ_{ET}^2 in these regions, respectively. Using ETVARD as a tool for multiple land surface model inter-comparison and diagnosis will be discussed in Chapter 6.

5.5 Conclusions

This paper reconciles multi-source, multi-variable hydrologic observations along with a theoretical *ET* variance assessment framework (ETVARD). The overall *ET* variance is categorized into one part with climatic variables (σ_{ETS}^2) and the other part with hydrologic system variables (σ_{ETF}^2). Based on σ_{ET}^2 derived from ETVARD (Experiment 1), we characterize the spatial distribution of σ_{ET}^2 and its climatic and hydrologic components over the CONUS. Although the contribution to σ_{ET}^2 from the climatic variables is larger than that from the hydrologic system variables in most of the regions of the CONUS, we identify some regions such as California and the lower reach of Mississippi River, where terrestrial water storage and components related to terrestrial storage change significantly change the σ_{ET}^2. In those regions, groundwater pumping for irrigation (e.g., in California) and water withdrawal from surface water (e.g., lower reach of Mississippi River) have led to systematic change of the terrestrial storage. Based on the comparison of three observation-based *ET* products using ETVARD as a reference, we propose diagnostic hypotheses in terms of the possible bias and uncertainty involving the various *ET* products: ET_{RS-UW} captures the high σ_{ET}^2 signals in the Midwest, with negligible "bias" and moderate uncertainty over the CONUS; $ET_{FLUX-MTE}$ systematically underestimates σ_{ET}^2 over the CONUS but with the lowest level of uncertainty; $ET_{RS-MOD16}$ has medium bias with highest level of uncertainty, and the spatial distribution of high σ_{ET}^2 signal from $ET_{RS-MOD16}$is different from other estimates. Note that these hypotheses each assume that the reference σ_{ET}^2 value derived from ETVARD is accurate, which may be also uncertain. This reference value itself depends on the quality of the multiple data sources that are used to estimate the climatic and hydrologic variables involved in ETVARD (*P* and *PET* from NLDAS-2 and *ΔS* from GRACE), including errors that can be caused by the aggregation processes of the data sources with different spatial and temporal resolutions. Nevertheless, the experiments presented in this paper demonstrate how to use a hydrologic theory (i.e., ETVARD in this case) to compare multi-source, multi-variable hydroclimatic observations. As the *ET* products based on observations have often been used for hydroloigc model development and water resources management, inter-comparison of the various *ET* products is necessary to reconcile any inconsistency or uncertainty. Given such discrepency among various *ET* products, model calibration and validation are conditioned on the quality of the observation data used. The insights on hydroclimatic observations provided in this paper will be extended by comparing and diagnosing *ET* variance simulated by multiple land surface models in Chapter 6.

CHAPTER 6

EVALUATION AND DIAGNOSIS OF ET VARIANCE FROM MULTIPLE LAND SURFACE MODELS

Following the observation assessment in Chapter 5, this chapter focuses on using ETVARD as a diagnostic tool to benchmark multiple land surface models. The insights from ETVARD and information embedded in observations will help to pinpoint the processes in land surface models that need improvement.

6.1. Introduction

Improved understanding of hydroclimatic processes, increasing computational power, and expanding data repository enable us to depict dynamic hydrologic processes and systems using physically more realistic numerical models, namely land surface models (LSMs). Such models are increasingly used for scientific understanding and decision-making support, and there are growing needs for systematic approaches for model evaluation and improvement [*Clark et al.*, 2015b]. The modelling community attempts to reduce model errors from several sources, including model input [*Cosgrove et al.*, 2003; *Badgley et al.*, 2015; *Herold et al.*, 2016], model structure [*Gulden et al.*, 2007; *Xu and Valocchi*, 2015] and model parameter [*Orth et al.*, 2016]. In addition, hydrologic modelers must consider the impacts of the quality of reference observations, which are usually not "accurate" though they are used as a benchmark for model outputs. For instance, *Tiedeman and Green* [2013] found that omitting observation error could either increase or decrease the parameter variance, depending on the correlation between observation errors and parameter sensitivities; *Montanari and Di Baldassarre* [2013] explored how an appropriate selection of model complexity would help reduce the effect of reference observation uncertainty. Observation errors not only affect the model results directly through model inputs [*Cosgrove et al.*, 2003] and data assimilation procedures [*Beven and Freer*, 2001], but also complicate model validation [*Hejazi et al., 2009*]. Due to potential error with the reference observation, a small discrepancy between model output and reference observation may not necessarily mean that the model is acceptable; meanwhile, a poor fit to a set of noisy observation data does not provide a sufficient reason to reject a model. Thus, the efforts in reducing the discrepancy between model outputs and observations in model calibration exercises may fail to improve the model, if the reference observations involve systematic errors. As shown by previous studies, a perfectly calibrated model may convert the reference observation error into error within a set of over-confident parameters [*Hejazi and Cai*, 2009].

Model evaluation can be further complicated when multiple inconsistent reference observations are available. For example, *Cai et al.* [2014] reported a reasonably good agreement in evapotranspiration (*ET*) annual mean estimates between simulations from LSMs in Phase 2 of the North American Land Data Assimilation System (NLDAS-2) and two remote-sensing *ET* products [*Jung et al.*, 2009; *Mu et al.*, 2011]. However, *Xia et al.* [2016] found that the same LSMs failed in generating the *ET* seasonal cycle observed from gridded FLUXNET observations [*Jung et al.*, 2009].

We may argue that hydrologic theories play a necessary role in bridging the gap between models and observations. Although observations, models, and theories are not perfect, each contains complementary information about the real world. Observations can capture the full range of hydrologic dynamics driven by climatic, biophysical and anthropogenic forcings [*Rodell et al.*, 2015]; models can predict the hydrologic responses to either stationary or nonstationary forcings and explore the feasible space of a hydrologic variable [*Kumar*, 2011]; theories are used to synthesize our understanding of hydrologic phenomena and expand hydrologic knowledge [*Kirchner*, 2006; *Clark et al.*, 2016]. This paper presents a model evaluation framework based on an observation-model-theory triplet (Figure 6.1), through which we examine both the congruence and discrepancy among observations, models and theories and provide guidelines for model improvement based on effective use of hydrologic observations and hydrologic theories.

Following the assessment of multiple observations in Chapter 5, in this paper we expand the context of observation-model-theory triplet (Figure 6.1) for LSM assessment over the contiguous US (CONUS) with respect to their estimates of *ET* monthly variance (σ_{ET}). In this context, the Evapotranspiration Temporal VARiance Decomposition (ETVARD) framework is used as a diagnostic tool that is based on a general theory [*Zeng and Cai*, 2015; 2016] and is independent from any particular LSM structures.

σ_{ET} from the LSMs will be compared against multiple observations. Furthermore, possible deficits in model structures of the LSMs will be diagnosed using the same set of inputs. As concluded in in Chapter 5, the inconsistency among the four observation-based *ET* variance (i.e., σ_{ET}) estimates show certain spatial patterns and cannot be simply treated as "white noise". The uncertainties in observation-based *ET* products, such as eddy flux tower and remote sensing products have also been identified by previous studies [*Jung et al.*, 2009; *Q. Tang et al.*, 2009; *Mu et al.*, 2011; *Landerer and Swenson*, 2012]. However, according to our knowledge, in the literature few have systematically investigated the impacts of the errors of multiple *ET* reference observations on LSM inter-comparison.

In the rest of this paper we will conduct the LSM assessment via three pre-designed experiments, which will cover: 1) the cross-evaluation of four LSMs (MOSIAC, NOAH and VIC from NLDAS-2 project [*Mitchell et al.*, 2004; *Xia et al.*, 2012a; *Xia et al.*, 2015a] and NOAH-MP [*Cai et al.*, 2014]) subject to multi-source reference observations; 2) comparison of the LSMs to ETVARD (as a benchmark) and the diagnosis of the possible deficits in each of the LSMs; 3) comparison of the LSMs' simulation of the terrestrial water storage to GRACE-estimated storage and their effects on the hydrologic system components of σ_{ET} through ETVARD, in which σ_{ET} is split into two parts (one from climatic variables and the other from hydrologic variables such as especially the terrestrial water storage change.) Finally based on the results from the experiments, diagnostic hypotheses will be provided regarding the evaluations and improvements of the LSMs.

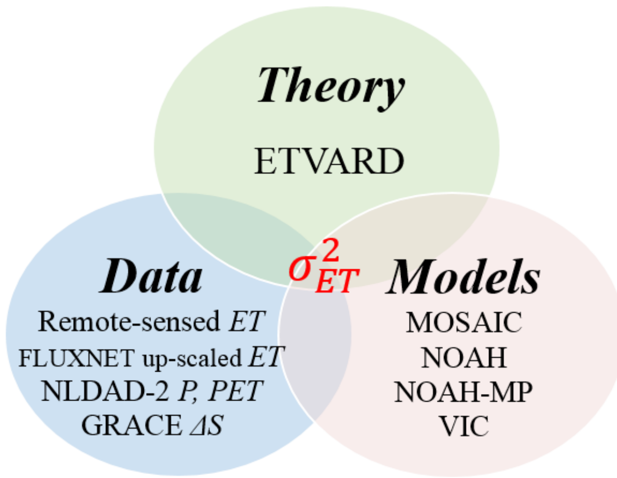


Figure 6.1. The congruence among hydrologic theories, multi-source multi-variable hydroclimatic observation data and multiple numerical models represent our organized understanding of hydrologic processes.

6.2. Methods

It has been argued that the model evaluation process should be "diagnostic", i.e., to obtain knowledge that can be used to either validate or reject the hypotheses underlying the model conceptualization and structure, which will eventually lead to improved models and advanced theories [*Gupta et al.*, 2008]. Hydrologic responses simulated by a model can rarely capture the full spectrum of hydrologic dynamics and/or hydrologic variability [*Kumar*, 2015], which however can be reflected by observations. Especially, current data acquisition has gone beyond what some existing LSMs can take as inputs. New variables, such as terrestrial water storage [*Long et al.*, 2015], are now available at the global scale. *Xia et al.* [2017] evaluated the monthly terrestrial water storage anomaly and the individual water storage components from three LSMs (i.e., CLM 4.0, Noah-MP and CLSM-F2.5, all including a groundwater component) against GRACE. However, the change of terrestrial water storage, which is widely caused by human interferences, remains as an outstanding issue with hydrologic models in general since few models have a reasonable depiction of the human dimension and its interactions with hydrologic processes [*Vogel et al.*, 2015].

Recent research efforts have been made to develop frameworks for LSMs inter-comparison, diagnosis and benchmarking in land surface modelling communities, for example, the Framework for Understanding Structural Errors (FUSE) [*Clark et al.*, 2008], the Joint UK Land Environment Simulator (JULES) [*Best et al.*, 2011], the Structure for Unifying Multiple Modeling Alternatives (SUMMA) [*Clark et al.*, 2015a; *Clark et al.*, 2015c] and the PALS Land Surface Model Benchmarking Evaluation Project (PLUMBER) [*Best et al.*, 2015]. These comprehensive frameworks examine the simulation of relevant hydroclimatic and land surface processes (e.g., *ET*, infiltration, streamflow, etc.) through inter-comparison of the various model configurations and process representations (e.g., VIC calculates *ET* from soil evaporation, canopy evaporation and vegetation transpiration [*Gao et al.*, 2010b]; NOAH calculates *ET* from snow sublimation, bare soil evaporation, canopy water evaporation and vegetation transpiration [*Niu et al.*, 2011]). In this paper, we do not provide a comprehensive LSM inter-comparison framework like those listed above. Our assessment focuses on *ET* variance, which is related to other hydroclimatic variables and processes. ETVARD is used as an analytical, diagnostic tool that is independent from any of the four LSMs. The diagnostic framework is based on the observation-model-theory triplet, with ETVARD as the theoretical component.

6.2.1 Model evaluation and diagnosis driven by hypothesis test

Some researchers [*Clark et al.*, 2011; *Beven*, 2012; *Clark et al.*, 2016] argued that a hydrologic model could be reviewed as a set of connected variables and coupled hypotheses associated to physical processes in either empirical or theoretical forms**.** For one particular process, there are usually several alternative hypotheses (e.g., infiltration can be described by Richard's equation or Green-Ampt method; potential *ET* can be calculated by Penman-Monteith equation or temperature based methods). In this context, a subset of hypotheses can be used for model construction and others for model evaluation and diagnosis. This is similar to the model calibration/validation procedures, where a portion of data is used to tune model parameters and the rest for model validation. An LSM represents a set of hypotheses that are posed for the various hydrologic processes such as runoff generation, infiltration, *ET*, etc. Since these processes are inter-connected, the test of these hypotheses should be conducted in a systematic framework. For example, *Koster and Suarez* [1999] related *ET* variance (σ_{ET}) to precipitation variance based on the Budyko theory and evaluated *ET* simulations from LSMs at the river basin scale.

Inter-comparison of multiple LSMs with different model structures (or hypotheses) and even different data inputs will need an independent reference or benchmark that is based on general theory. ETVARD plays such a role in this study. ETVARD relates σ_{ET} to the variance and covariance of precipitation (*P*), potential *ET* (*PET*) and terrestrial water storage change (ΔS) at the watershed scale

[*Zeng and Cai*, 2015; 2016], all of which are available from multi-source observations or assessments. As elaborated by Eqn.(5.1) in Chapter 5, ETVARD decomposes σ_{ET} into climatic and hydrologic components (σ_{ETF} andσ_{ETS}), and by comparing these two components to those derived from the LSMs will allow us to identify some specific processes (climatic vs. hydrologic) simulated by a LSM that may involve some discrepancy, for example, the terrestrial storage change associated with groundwater pumping for irrigation and/or water supply, which needs to be added or refined by existing hydrologic models [*Vogel et al.*, 2015].

In addition, the spatial patterns of σ_{ET} components identified by ETVARD (from Experiment 1 in Chapter 5) will be compared to the patterns generated by the LSMs, which will indicate particular locations where possible discrepancies in simulating *ET* processes exist with an LSM.

It should be noted that there are several other approaches towards reconciling models with data and making more effective use of data for model improvement. Especially, data assimilation has been widely used for assimilating real-time observation of model state variables to adjust model outputs dynamically [*Cosgrove et al.*, 2003]. Numerous efforts have also been made using machine learning algorithms. For example, such algorithms have been used to construct error models to correct the epistemic error of spatially-distributed physically-based models [*Xu et al.*, 2014]. However, it remains challenging with data driven methods on how to best utilize the ever expanding repository of hydrologic observation for model diagnosis, such as attributing model error to specific model processes [*Xu et al.*, 2017a]. The approaches based on theoretical hydrologic relationships therefore can have advantage over data-driven approaches for model diagnosis due to their explanatory power[*Xu et al.*, 2017b], which is analogy to the issue of "white box" vs. "black or gray box".

6.2.2 Experiment design

The experiments to be used in this paper follow the two experiments in Chapter 5 with three additional experiments designed for LSM assessment, as shown in Figure 6.2. Under the three experiments, the four LSMs (MOSIAC, NOAH, VIC NOAH-MP). These models use the same climate forcing and land cover parameters at the same temporal and spatial scales, which allows the inter-comparison to focus on model structure. This study uses the monthly scale model inputs (i.e., *P* and *PET*) and outputs (i.e., *ET* and *ΔS*), which are available at NOAA/NCEP/EMC NLDAS ftp servers. An LSM calculates terrestrial *ET* from soil, canopy, snow and vegetation, depending on the processes formulated in the LSM. Terrestrial water storage change (*ΔS)* includes the changes of soil moisture, snow and aquifer storage. To compare a LSM to ETVARD, the LSM results obtained at the resolution of 0.125° by 0.125° (the common scale used by all the LSMs) to 1° by 1°, the resolution of GRACE data. Each of the LSMs simulates *PET* with different methods but using the same meteorological foricngs; while ETVARD uses the *PET* that is also calculated using NLDAS-2 forcing data [*L. Mahrt and Michael Ek*, 1984].

As shown in Figure 6.2, Experiment 3 compares the *ET* variance at the monthly scale from the four LSMs (denoted as σ^2_{ET-LSM}) to that from four observation-based *ET* products

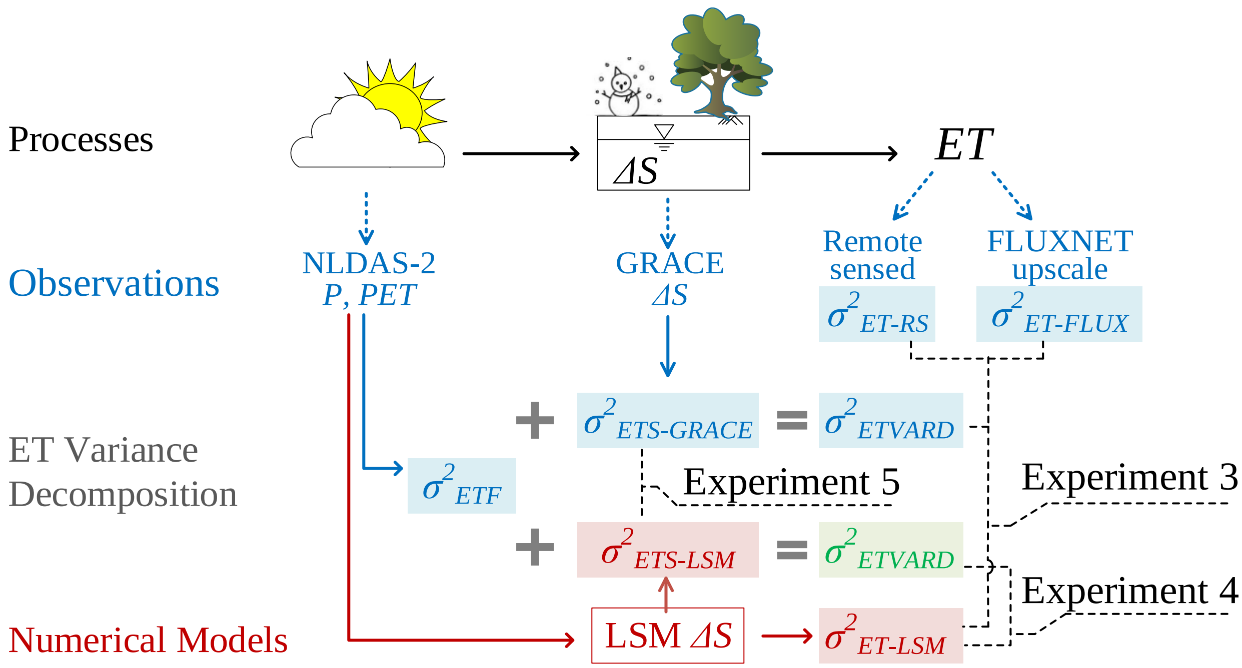


Figure 6.2. Schematics of hydrologic processes as a simple system, along with various ET variance estimates and its components from observation, simulation and ETVARD approaches. Variables that are represented in black for true but unknown values, in blue for quantities from direct or indirect observation data, in red for quantities by numerical models and in green for "hybrid" quantities from both observation and model results. Blue dash lines represent for data acquisition which is inevitably subject to observation and/or processing error. Solid lines represent for information propagation indirection. Black dash lines illustrate the pairs of quantities assessed in each experiment in the two companion papers.

(Experiment 2 in Chapter 5). By calculating the difference of σ^2_{ET} between each LSM simulation and each observation-based product, we will obtain a matrix showing the comparisons of four LSMs and four observations. This experiment is designed to show how the conclusion of model evaluation varies with the references.

In Experiment 4, σ^2_{ET} calculated from ETVARD takes the same climate forcings (i.e., *P* and *PET*) as those used in the LSMs and the terrestrial storage *ΔS* from each of the four LSMs. Thus in this experiment, σ^2_{ET-LSM} and σ^2_{ETVARD} are based on the same climatic and hydrologic inputs. This experiment isolates the effect on σ_{ET} estimates associated with governing processes fromthat associated the input data.

Therefore, the comparison will focus on the difference caused by the model structure (i.e., physical process representation) of an LSM and the analytical form (Eqn.5.1 in Chapter 5) of ETVARD.

Experiment 5 is particularly designed to assess the impact on σ_{ET}^2 from the terrestrial water storage change (*ΔS*) estimates based on two sources: GRACE-based observation and LSM-simulation, i.e**.,** the only variable of interest in Experiment 5 is *ΔS*. With the same climatic forcings (i.e., *P* and *PET*) for ETVARD and LSMs, we focus on the comparison of the hydrologic component σ_{ETS}^2 (Eqn.5.3 in Chapter 5). σ_{ETS}^2 includes water storage change variability, the correlation between *P* and *ΔS* (e.g., soil moisture replenish, aquifer recharge to rainfall excess and/or pumping and water withdrawal in dry days) and the correlation between *PET* and *ΔS* (e.g., snow melting and thaw). Note that σ_{ETS}^2 represents the water storage related components in σ_{ET}^2, therefore can be both negative or positive. Since GRACE observation includes *ΔS* from groundwater, which however is generally not simulated by operational LSMs [*Xia et al.*, 2017]. Through Experiment 5 we expect to identify the locations for LSM improvement, where land surface processes actively interact with groundwater, by either natural processes (e.g., groundwater recharge/discharge) or human activities (e.g., groundwater pumping), or both.

6.3. Results

6.3.1 Inter-comparison of σ_{ET} among multiple reference observations, ETVARD and multiple LSMs (Experiment 3)

The monthly σ_{ET} from the four LSMs ranges from 0 to 60 mm as shown in Figure 6.3. The four LSMs commonly produce high σ_{ET} (above 40 mm) in Midwest and low σ_{ET} (below 20mm) in the western region of meridian 100°W. Meanwhile the four LSMs produce different levels of σ_{ET} in the northeastern region of CONUS, where σ_{ET} above 30 mm by MOSAIC and NOAH-MP and around 20 mm by NOAH and VIC. The LSM-simulated σ_{ET} values show significant differences along the West Coast compared to the four observation-based estimates (Figure 5.8 in Chapter 5). Compared to four observation-based estimates, which all yield noticeable σ_{ET} (larger than 30 mm) along the West Coast though varying by magnitude, the four LSMs results in low σ_{ET} (20 mm) along the West Coast. Compared to the result of ETVARD, the four LSMs consistently generate low σ_{ET} in the West Coast. A unique contributor to σ_{ET} along the West Coast is the Mediterranean climate. By ETVARD, the out-of-phase between the rainfall season and the warm season results in a negative climatic component (i.e., $w_{P,PET}cov_{P,PET}$) in σ_{ET} in this region, as shown in Experiment 1 and Figure 5.4c in Chapter 5. In addition, the contributions from terrestrial water storage change in this region are also significant. In California, the terrestrial water storage release during the dry season leads to a significant reduction in σ_{ET} via a negative $w_{P,\Delta S}cov_{P,\Delta S}$ component; while snow melting during the warm season enhances σ_{ET}

with a positive $w_{PET,\Delta S}cov_{PET,\Delta S}$ component in the coast region of Oregon and Washington. Thus the relatively low σ_{ET} from the four LSMs in the West Coast is probably due to the Mediterranean climate and/or the limited water storage representation in the models. More detailed results on terrestrial storage change effects should be referred to Experiment 5, which includes the impact of GRCACE-estimated terrestrial storage change in the comparison.

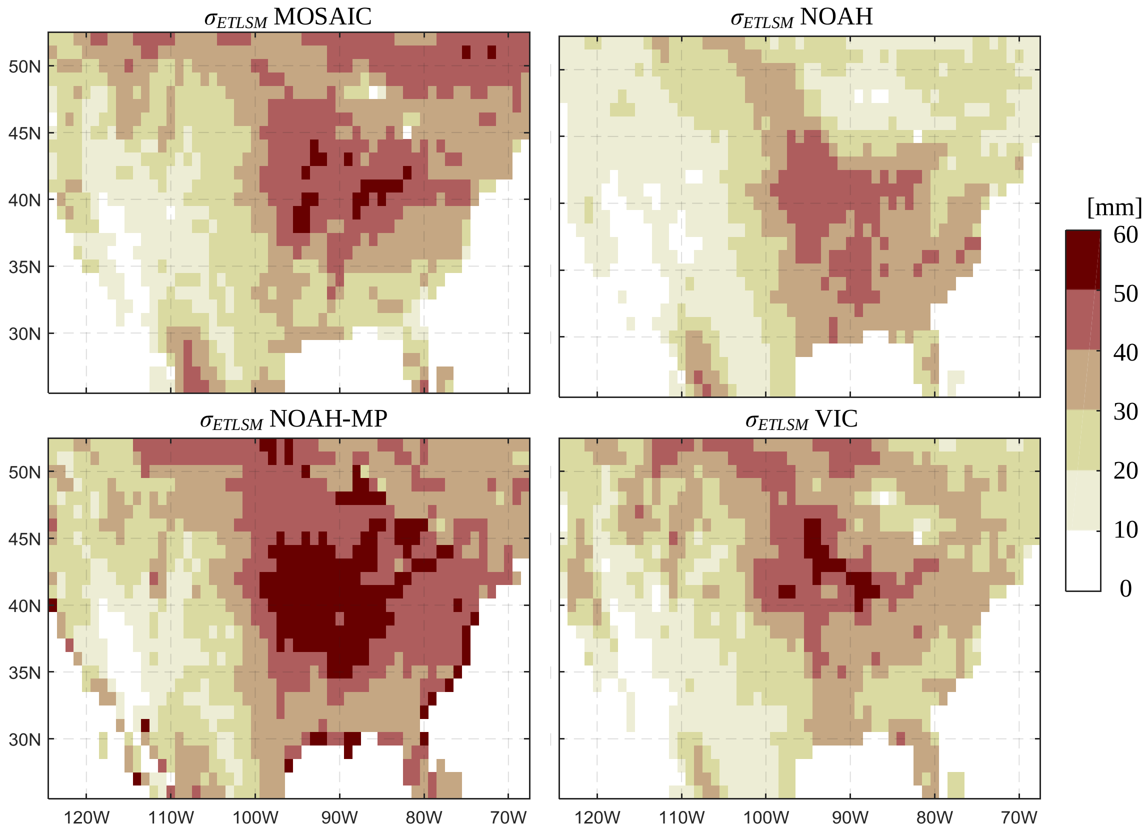


Figure 6.3. σ_{ET} (i.e., the red σ_{ETLSM} in Figure 6.2) simulated by the four LSMs (MOSAIC, NOAH, NOAH-MP and VIC), which are driven by the same forcing data sets and executed at the same temporal and spatial resolution.

The average residual (σ_{ETres}) between the σ_{ET} from a LSM and that from an observation based product is calculated as the mean absolute difference between the two over all grids in the CONUS, that is, $\sigma_{ETres} = \frac{1}{n}\sum_{i=1}^{n}|\sigma_{ETLSM} - \sigma_{ETObs}|$. The pair-wise inter-comparisons are shown in Table 6.1. σ_{ET} calculated by ETVARD is also used as a reference together with the observations. By each column of Table 1, one observation is used as the reference, and the model with the smallest residual is picked as the "best model". For example, when σ_{ET} from ETVARD is treated as reference, MOSIAC model has the smallest residual (i.e., 8.41 mm) among the four models and is therefore chosen as the "best model". It is surprising to find that each of the LSMs is identified once as the "best model" with the various references. This confirms that inter-comparison of the multiple LSMs is observation-dependent. Recognizing the possible limitations of using any single model for problem solution, ensemble-based approaches have been widely used to handle model uncertainties, in which the results from the various models are combined with given a certain set of priorities (often subjective) on the models.

Table 6.1. Pairwise σ_{ET} differences between land surface models and observation products. Column-wise comparison represents the average σ_{ET} residual when an observation-based σ_{ET} is used as reference, so the smallest absolute value in the column (in *italic*) indicates the best model.

[mm]	ETVARD	RS-UW	MOD16	FLUXNET
MOSIAC	***8.41***	7.72	10.96	12.26
NOAH	12.03	11.96	10.52	***6.21***
NOAH-MP	8.95	***6.58***	11.16	15.49
VIC	9.61	8.52	***10.26***	9.12
Best model	MOSIAC	NOAH-MP	VIC	NOAH

6.3.2 Model structure assessment by using ETVARD as a benchmark for LSMs (Experiment 4)

σ_{ET} by ETVARD with ΔS simulated by each of the four LSMs is shown in Figure 6.4. The four σ_{ET} estimates exhibit a clear contrast along the east-west direction near the meridian 100°W line. For all the cases, i.e., σ_{ET} from ETVARD using ΔS from all the LSMs is less than 20 mm in the west mountains and larger than 40 mm in the West Coast of California. The high σ_{ET} (about 50 mm) is generally located in some areas around the Midwest, while σ_{ET} with ΔS from NOAH-MP generates high σ_{ET} in the whole eastern part except for the areas along the Appalachian Mountains.

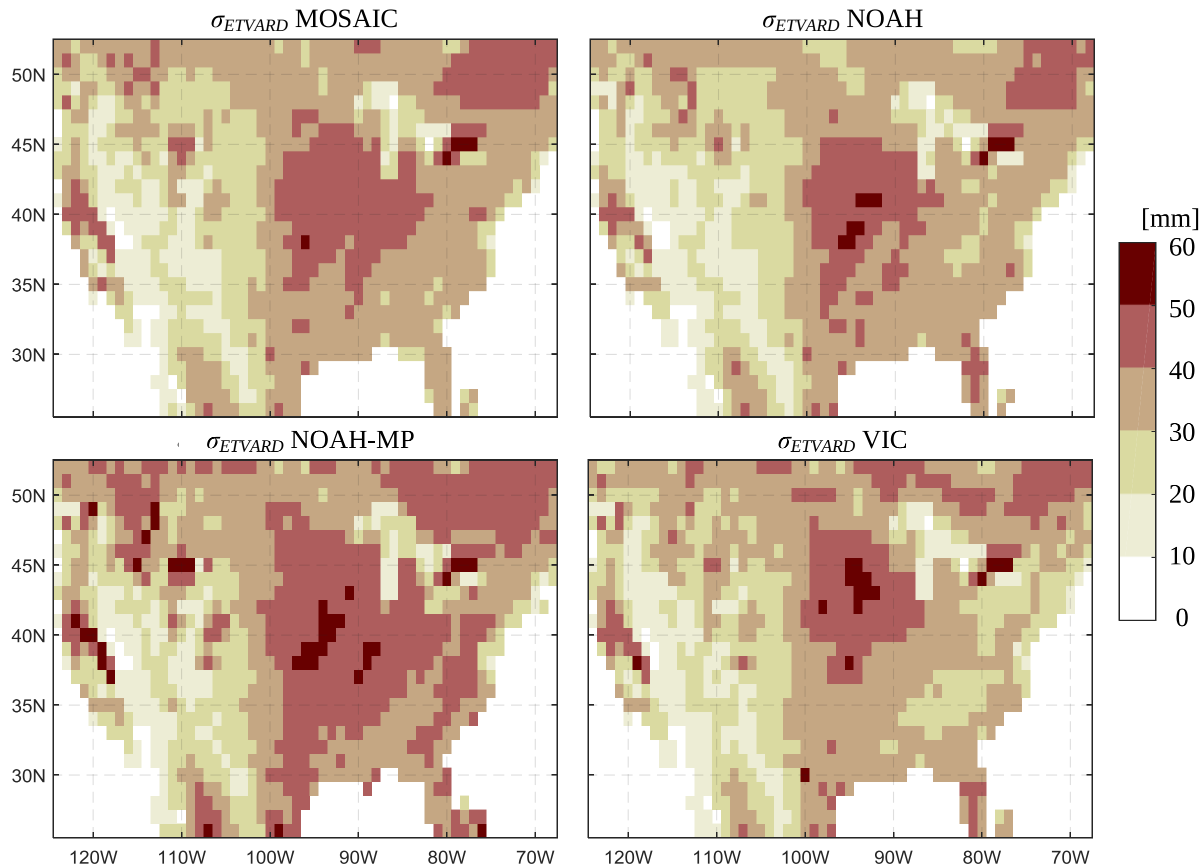


Figure 6.4. σ_{ET} calculated by ETVARD (i.e., the green σ_{ETLSM} in Figure 6.2) with terrestrial water storage change (*ΔS*, including soil moisture, snow and/or groundwater) simulated by different LSMs.

Figure 6.5 displays the differences in σ_{ET} from each of the four LSM results (i.e., $\sigma_{ETLSM} = f_{LSM}(P, PET, \Delta S_{LSM})$ in Figure 6.3, where f_{LSM} represents a LSM model function) and ETVARD with *ΔS* simulation from each of the four LSMs as input (i.e, i.e., $\sigma_{ETVARD} = f_{ETVARD}(P, PET, \Delta S_{LSM})$ in Figure 6.4, where f_{ETVARD} represents Eqn.(5.1) in Chapter 5. Note that the inputs $(P, PET, \Delta S_{LSM})$ to ETVARD and the LSMs are the same, Figure 6.5 isolates the impact on σ_{ET} from the input data and explicitly show the difference between an LSM and ETVARD caused by the physical process representation of σ_{ET} in LSM (i.e., f_{LSM}) and the analytical ETVARD (i.e., f_{ETVARD}). A common spatial pattern shared by the four LSMs is that σ_{ET} along the West Coast is significantly smaller (about 20 mm) than that from ETVARD. As discussed in Experiment 3, the most apparent σ_{ET} difference between LSMs results and the observation-based estimates (from Experiment 2 in Chapter 5) is located along the West Coast. We have suggested that the difference may be caused by inaccurate simulation of terrestrial water storage or by inadequate process representation under the Mediterranean climate. In this experiment that compares the model structures in terms of the differences in ET variance, we may further claim that the differences are mainly attributed to the model structures of the LSMs**.**

Although the σ_{ET} differences between LSMs and ETVARD are found with other regions, they are not consistently shared by the four LSMs. For instance, MOSAIC, NOAH-MP and VIC generally yield slightly higher σ_{ET} (less than 5 mm) than that from ETVARD in the Midwest and Northeast, while NOAH exhibits the pattern mainly in the Southeast. NOAH and NOAH-MP predict significant lower σ_{ET} (more than 20 mm) than that by ETVARD in the region around Idaho, where the covariance between ΔS and PET contributes considerably to σ_{ET} (in Figure 5.6f in Chapter 5). This implies the differences might be mainly associated with the snow processes or vegetation's responses to solar radiation in NOAH and NOAH-MP. In addition, NOAH and VIC show significantly lower σ_{ET} than that by ETVARD and observation-based σ_{ET} around the southern region along meridian 100°W, where P is the largest component in σ_{ET} as shown in Figure 5.6a in Chapter 5.

Although we do not claim that any of the estimates by LSMs, ETVARD or observation-based estimates is accurate, this experiment shows that in most of the regions in the CONUS, the estimates from ETVARD and observations are more similar compared to estimates from LSMs. Following the analysis of the contribution sources of σ_{ET} in Experiment 1, we can target some particular processes contributing to the disagreements for further studies. Moreover, taking the ETVARD as a benchmark, Experiments 1 and 4 can be used for identifying the processes controlling σ_{ET} and their spatial locations in the four LSMs. For example, Experiment 1 shows that the energy budget dominates σ_{ET} in the coast of Washington and Oregon. Therefore, the models in these regions should be examined in the energies related processes such as snow dynamics or vegetation water demand.

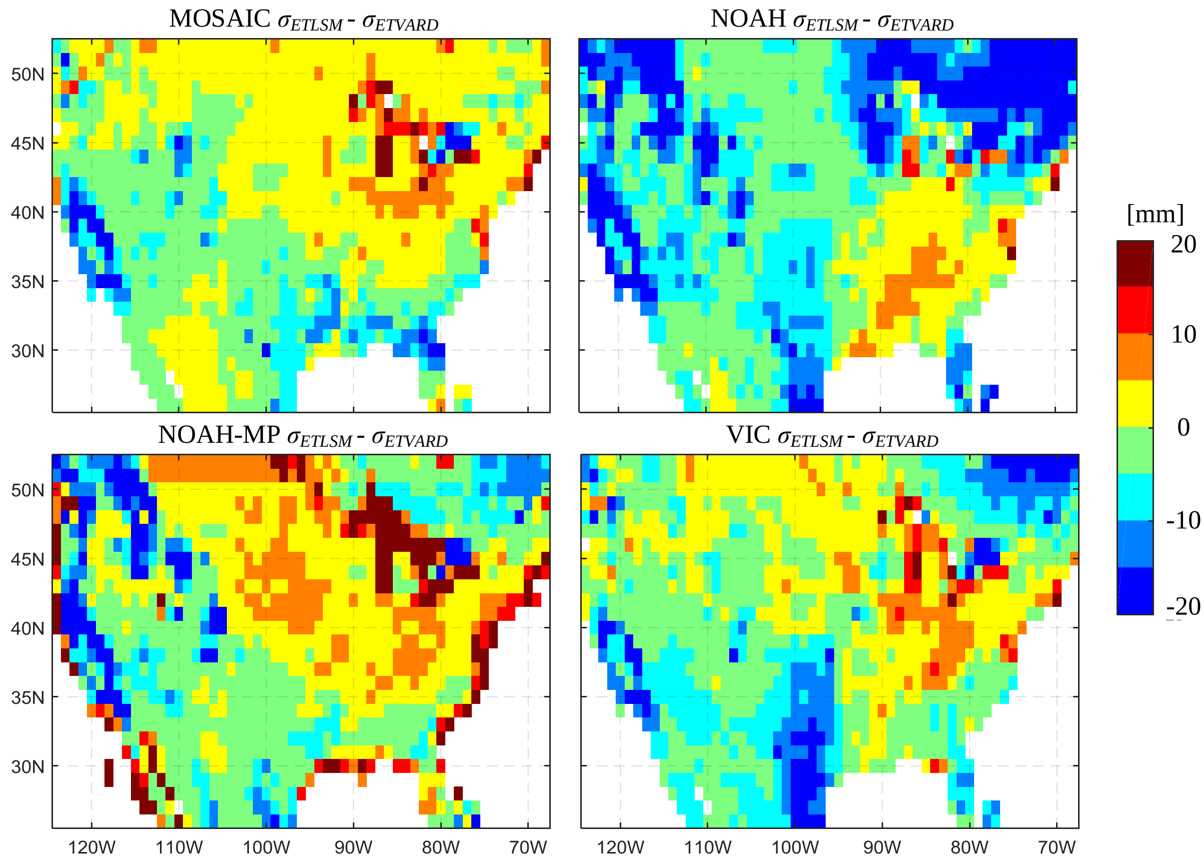


Figure 6.5. The σ_{ET} residual between ETVARD ($\sigma_{ET} = f_{ETVARD}(P, PET, \Delta S_{LSM})$) and LSM ($\sigma_{ET} = f_{LSM}(P, PET, \Delta S_{LSM})$). With the same input data, this residual shows the pair-wise discrepancy between benchmarking ETVARD and aggregated processes in LSM (i.e., f_{ETVARD} vs. f_{LSM}).

6.3.3 LSM diagnosis using hydrologic observations (Experiment 5)

The terrestrial storage component, σ^2_{ETS}, calculated from ΔS_{GRACE} and four LSM simulated ΔS_{LSM}, respectively, ranges from -800 to 1200 mm^2, as shown in Figure 6.6.6. The estimates from the four LSMs and GRACE are quite consistent in the South and the West, where *ΔS* buffers the *ET* fluctuation. In Idaho, all five σ^2_{ETS} estimates consistently indicate that *ΔS* enhances σ_{ET}, mainly due to the snow storage. NOAH-MP results exhibit the pattern in a slightly larger area than other models. Experiment 4 shows that the snow processes (variance of the storage) or vegetation's response to solar radiation (via the co-variance between *ΔS* and *PET)* in NOAH and NOAH-MP may be responsible for the difference between LSMs and ETVARD.

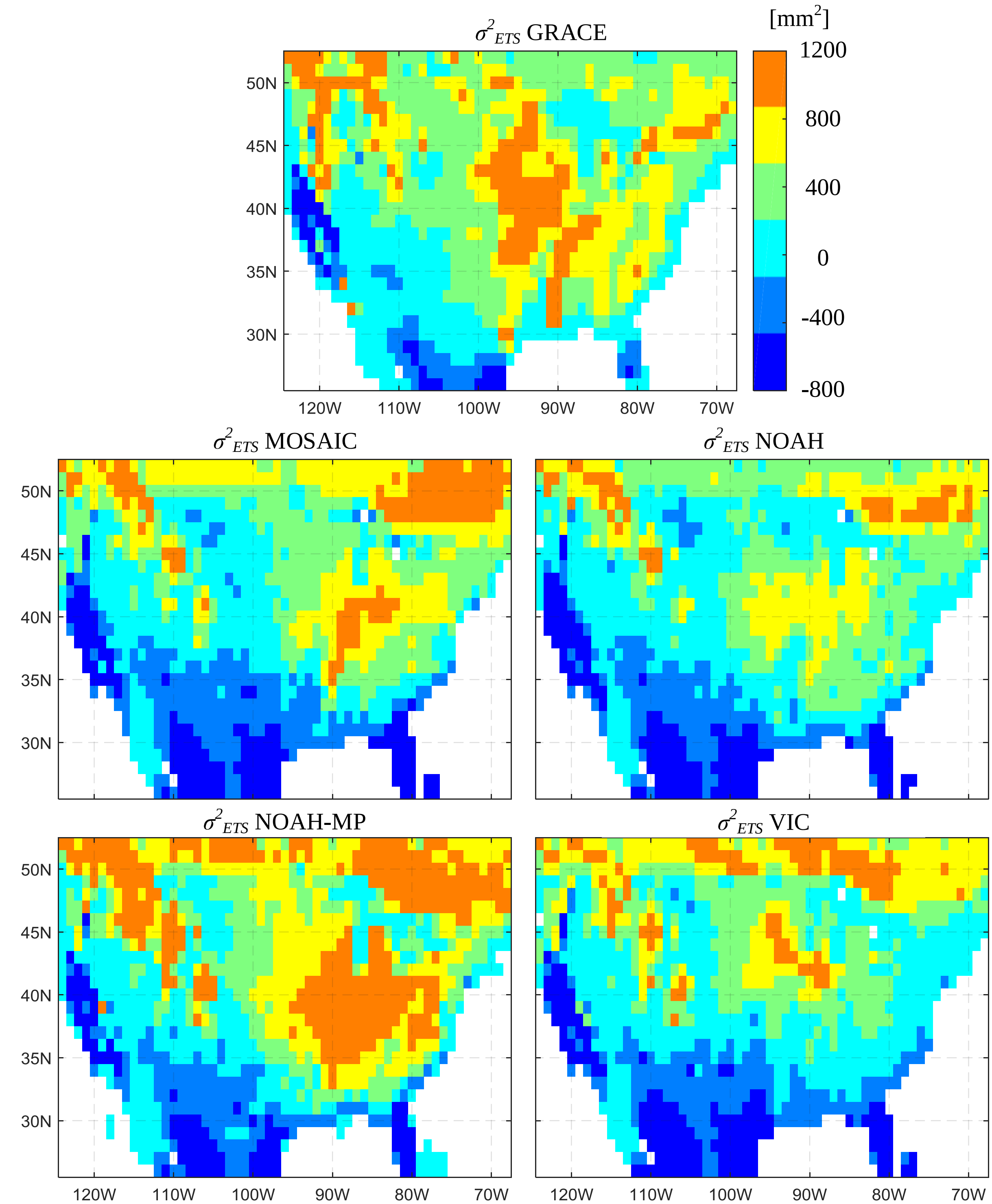


Figure 6.6. σ^2_{ETS}**, the terrestrial water storage change components in** σ^2_{ET}**, with** ***ΔS*** **from GRACE observation and the four LSMs simulations.**

Experiment 5 further finds that the vegetation's response to solar radiation (the covariance item, $w_{PET,\Delta S}cov_{PET,\Delta S}$) can be the primary reason for the difference.

The most apparent σ_{ETS}^2 difference between GRACE observation and LSMs simulation appears in the Midwest and the High Plains. $\sigma_{ETS-GRACE}^2$ shows that *ΔS* substantially enhances σ_{ET} in the Midwest and the northern and middle High Plains; while the four LSMs generate large σ_{ETS}^2 generally in the east of meridian 90°W and their spatial patterns are inconsistent. The significant impacts on *ET* from agricultural land use and groundwater based irrigation in these regions have been well-recognized by both remote-sensing estimates [*Strassberg et al.*, 2009; *Mutiibwa and Irmak*, 2013] and groundwater well measurements [*McGuire*, 2012; *Haacker et al.*, 2015]. However, an accurate representation of heavily managed agricultural land use still remains a challenge in LSM formulation LSMs generally have a relative shallow soil profile (e.g., 2m in VIC [*Liang et al.*, 1994]) which can be sufficient to characterize natural vegetation root water up-taking but cannot catch the effect of groundwater pumping which decreases water storage deep in the aquifer, and in turn cannot reflect the effect of accumulative depletion of aquifer storage [*Zeng and Cai*, 2014]. Although the NOAH-MP has a simple aquifer representation, the transient decline in groundwater level results in a large amount of storage change which is beyond the storage scope in LSMs. Thus beyond the issue of more accurate simulation of *ΔS*, Experiment 5 unveils how a better simulation of *ΔS*, especially in intensively managed agricultural land, would improve the simulation of *ET* and *ET* variance in LSMs.

The scatter plot of $\sigma_{ETS-GRACE}^2$ and $\sigma_{ETS-LSM}^2$ of the four LSMs in the CONUS is shown Figure 6.7. Overall, all LSMs yield a smaller σ_{ETS}^2 components than the GRACE-based σ_{ETS}^2 (the regression slopes are less than 1). Among all the LSMs, NOAH-MP gives closest σ_{ETS}^2 to GRACE-based estimate than other three LSMs, which is probably due to the aquifer module (though a simple one) in the NOAH-MP. It is noted that in regions where *ΔS* buffers σ_{ET} (i.e., the σ_{ETS}^2<0), the buffering effect by LSMs is consistently less than that reflected by GRACE observation. These regions are mainly located in the western mountainous regions where terrestrial water storage plays a more important role in σ_{ETS}^2 than other regions. Further study is needed to assess not only the accuracy of *ΔS* from GRACE but also the uncertainties involved in the LSMs especially in the process representations in those models associated with *ΔS* simulation.

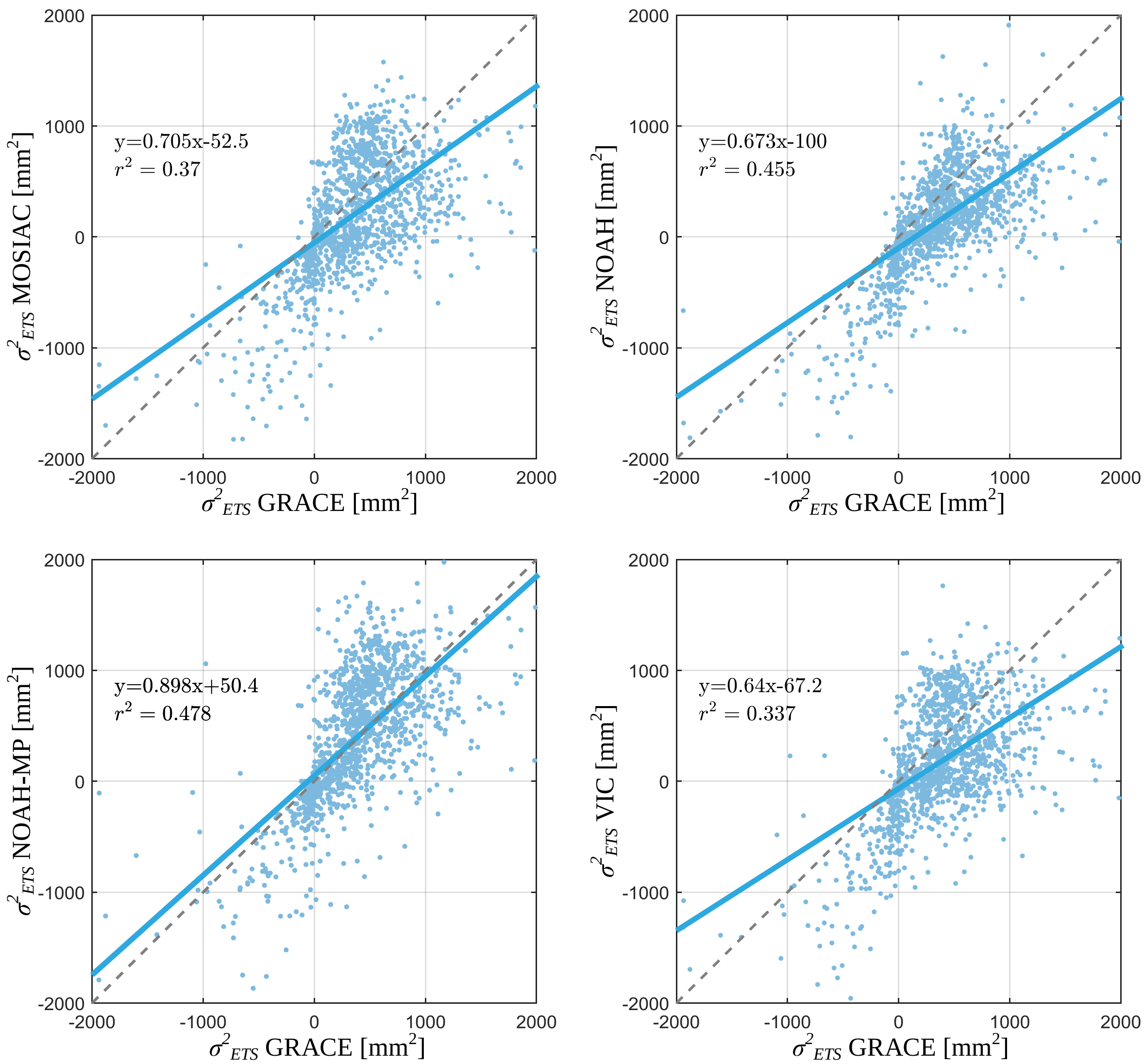


Figure 6.7. The scatter plot of σ^2_{ETS} from GRACE observation and the four LSM simulations. The positive σ^2_{ETS} indicates grids where σ_{ET} is enhanced by terrestrial water storage change, and negative σ^2_{ETS} indicates grids where σ_{ET} is dampened by terrestrial water storage change.

6.4. Discussions

6.4.1 Model evaluation as decision making with reference observation uncertainty

According to the terminology by *Best et al.* [2015], "model evaluation" means that model results are compared to observations to measure some error indices; "model comparison" involves calculating the error indices from multiple models with a common reference, and a model with smaller error indices is considered to be superior to other models. This paradigm implicitly assumes that there is a single observation that well captures the reality. However, for instance, Experiment 2 in Chapter 5 has shown that the four observation-based σ_{ET}^2 estimates are inconsistent and some or all of these observations are subject to bias and uncertainty. Given the inconsistency among multi-source observations, Experiment 3 shows that the selection of the "best model" among the four LSMs is observation-dependent, and a small discrepancy between an LSM and an observation does not necessarily imply that the model is accurate.

Setting model evaluation in a hypothesis test framework [*Vogel et al.*, 2013], traditional model evaluation is essentially based on the null hypothesis H_0, by which a model with a small error index compared to a reference observation is a good model with a certain significance level α, as shown in Figure 6.8a. The significance level α (e.g., 5%) implies that if H_0 is true, we have a probability of α to mistakenly reject a good model, termed as Type I error. H_0 reflects model developers' concern of model development failure. Given the uncertainty in a reference observation, however, we may also concern with an alternative hypothesis H_A , by which the model with small error indices compared to the observation is not a good model, as shown in Figure 6.8b. If H_A is positive and the observation used as reference is unreliable, there is another probability β that we may mistakenly accept a wrong model due to unreliable data. This is referred to a Type II error, which is usually ignored in traditional model evaluation but can have serious consequences if the wrong model is used for any operational applications such as real time drought monitoring [*Anderson et al.*, 2013].

In this study, the collision between models and observations shows that model evaluation should be regarded as "decision-making under uncertainty" to fully account for the reliability in reference data, as illustrated in Figure 6.8b. Compared to traditional model evaluation which is deterministic (implicitly ignoring observation data uncertainty), this paradigm emphasizes how observation data reliability affects the conclusion in model evaluation. Indices about data uncertainty (e.g., confidence intervals and error bars) in those observation-based estimates should be incorporated for model evaluation as stochastic decision-making, while quantifying the significance level (i.e., α and β) in both null and the alternative hypothesis.

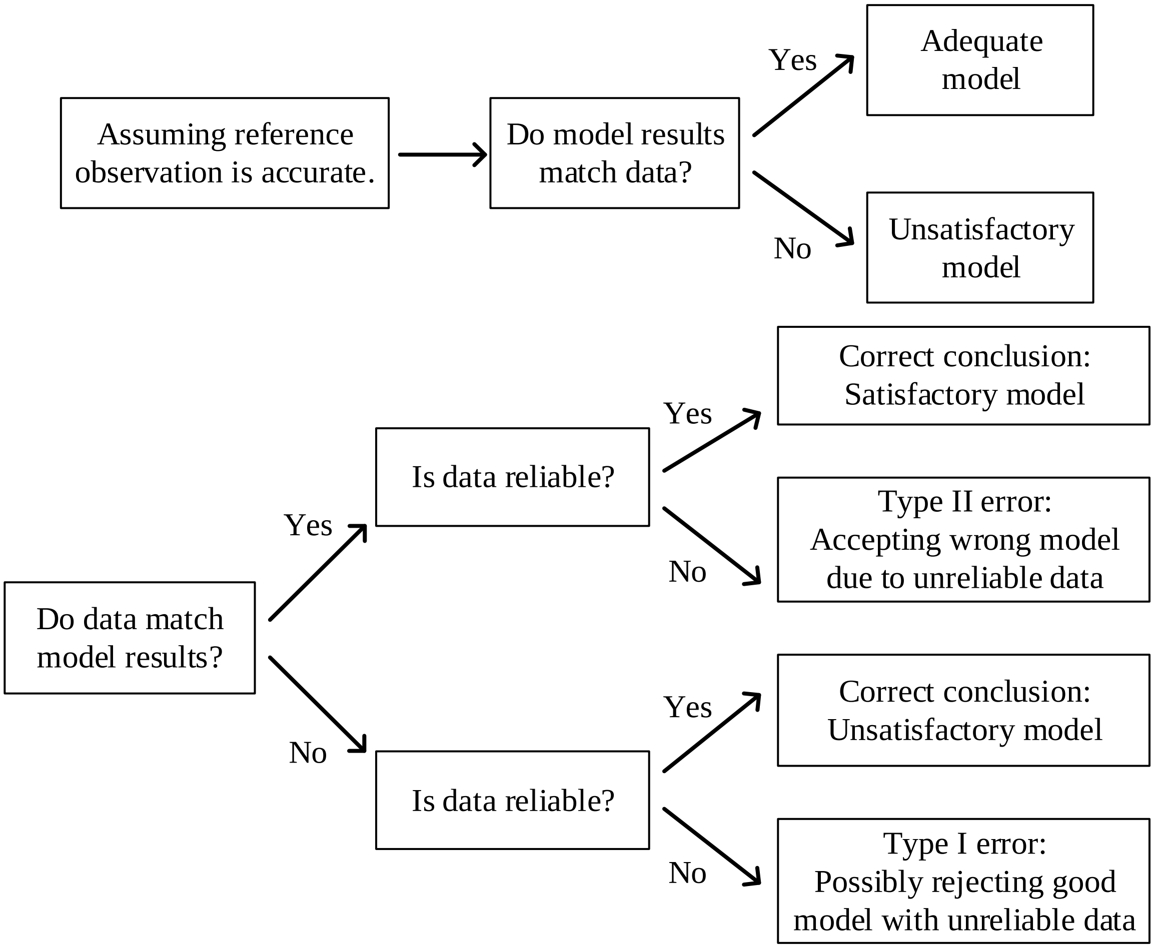


Figure 6.8. Decision diagram showing model evaluation a) assuming reliable reference observation and b) considering the effect of reference observation uncertainty.

Experiment 3 assesses σ_{ET} from multi-observations and multi-models, illustrating the possible fallacy in traditional model evaluation without carefully taking the reference observation uncertainty into account. This brings up a more profound issue in model-data interface that the match between observation and simulation does not necessarily correctly capture the reality due to the lack of a diagnosis tool based on a confirmed generic theory. Modelers continuously improve their models with more accurate forcing, better model structure and more realistic parameterizations to reduce the discrepancy between model result and observation data; meanwhile data scientists use more advanced data mining techniques and more physically-sound methods to retrieve target variables from raw sensor signals and justify the products against relevant theories, empirical relations, and/or model simulations. In this iteration between model update and observation improvement, we would hopefully reach a convergence between model and observation. However, this procedure may still lead to an unnoticeable fallacy that model and observation finally converge to a point that is far away from the reality.

This issue motivates the two companion papers to examine hydrologic knowledge as congruence among the observation, model, and theory, as shown in the triplet of Figure 6.1. Using *ET* variance as an example, we illustrate how multi-source multi-variable hydroclimatic observations, multiple LSMs and a theoretical ETVARD framework can serve complementarily to cross-diagnose each other through the five systematically designed experiments. We particular emphasize the role of ETVARD as an independent diagnosis tool in the observation-model-theory triplet. However, according to the Popperian falsification [*Popper*, 2005], we have to admit that even the congruence over the observation-model-theory triplet may not necessarily provide true knowledge that can be used to characterize σ_{ET}. In other words, although the congruence identified through our experiments σ_{ET} withstands falsification, it is only plausible and can be improved with better model, observation, or theory.

6.4.2 Limitations and future perspectives

The purpose of this study is to bring the theoretic ETVARD framework for the reconciliation between LSMs and observations and focuses on *ET* temporal variance at the month scale. This study does not aim at providing a comprehensive framework for LSM diagnosis, as did by others efforts [*Best et al.*, 2015; *Clark et al.*, 2015a]. However, we illustrate a meaningful framework in which ETVARD is used to disaggregate and diagnose σ_{ET} in LSMs while systematically adopting hydrologic observations that reflect some dynamics that may not be well captured by LSMs. We do not explicitly assess the impact of climatic forcings on σ_{ET}, given that the four LSMs underlying the three experiments use the same set of forcings from NLDAS-2 project). If another set of climatic forcings (*P* and *PET*) are available, this study can be extended to account how different climatic forcings impact σ_{ET}.

With growing amount of hydroclimatic observation data, LSMs have being improved. However, new theories and hypothesis are still needed to synthesize hydrologic knowledge through the observation-model-theory triplet. Researchers has recognized that the existing and even growing gap between models and theories is impeding the progress of hydrologic science [*Clark et al.*, 2016]. The ETVARD framework is our first attempt towards the congruence among the observation-model-theory triplet.

Another issue is that existing hydrologic relationships are generally obtained in natural watershed with minimal human interferences. Existing LSMs are essentially simulating the virgin hydrologic cycling without fully considering anthropogenic impacts. As human activities play an increasing role in transforming hydrologic processes, such as irrigation and baseflow [*Wang and Cai*, 2009], hydrologic models would be developed or improved to better capture the anthropogenic components at multiple temporal and spatial scales [*Vogel et al.*, 2015].

6.5. Conclusions

Following the multi-source, multi-variable observation assessment conducted in Chapter 5, this study further evaluates four LSMs (MOSAIC, NOAH, NOAH-MP and VIC) over the CONUS with respect to their estimates of monthly ET variance (σ_{ET}). In the context of an observation-model-theory triplet, the Evapotranspiration Temporal VARiance Decomposition framework is used as a diagnostic tool that is based on general theory and independent from any particular LSM structures. The LSMs are compared against multiple observations, as well as ETVARD. It is found that any of the four models compared can be the "best" one for a certain set of reference observations, which confirms our argument that inter-comparison of multi-models depends on the reference observation. Therefore, simply minimizing the residual between model and observation may result in rejecting a good model with unreliable observation (Type I error) or accepting a wrong model with unreliable observation (Type II error).

It is also found that σ_{ET} derived from ETVARD is consistently closer to observation-based estimates than the LSM simulations, especially in regions along the West Coast, Midwest and High Plains. The four LSMs might underestimate σ_{ET} along the West Coast due to the Mediterranean climate and human water use; the four LSMs might also underestimate the terrestrial storage contribution to ET variance in the High Plains compared to the ETVARD estimate and GRACE observation. This is probably due to the inappropriate representation of groundwater pumping and its impact on ET and other hydrologic processes in those LSMs. Furthermore, compared to GRACE-based estimates, the four LSMs do not capture the high σ_{ETS}^2 signal in the Midwest and High Plains. This is likely due to the limited representation of the hydrologic processes in the LSMs that control the terrestrial storage changes such as groundwater balance in aquifers and vegetation dynamics.

In Chapter 5 and 6, the ETVARD framework is applied toward reconciliation between hydrologic observations and LSM simulations with respect to monthly σ_{ET} for the CONUS. Via five systematically designed experiments, we diagnose the congruence in σ_{ET} among multi-source and multi-variable hydrologic observations, multiple LSMs, and ETVARD. Each experiment independently and complementarily provides information for the various assessments. Given possible errors and uncertainties in multiple models and multiple observations, the observation-model-theory triplet with a theoretical diagnostic tool is useful for cross-validating hydrologic theories, observations, and models. In particular, in this era with increasing multi-source and multi-variable hydrologic observations and improvement in various hydrologic models, we demonstrate the role of generic hydrologic theories (e.g., ETVARD in this study) as a bridge between models and observations and encourage stronger efforts along the line for the hydrologic community.

CHAPTER 7
A POWER LAW RELATIONSHIP BETWEEN ET MEAN AND VARIABILITY

This chapter syntheses the findings on ET variance in previous chapters and explores the linkage between the ET mean value and variance. An empirical statistical power law is found between ET mean value and monthly variance for various ecosystems. By incorporating the land use and vegetation structures, the ET power law relationship is examined for different ecosystem water use strategies, focusing on evaluating the trade-off between the mean and the variability of water consumption. This relationship provides insights to better understand and manage watersheds as coupled nature-human systems.

7.1. Introduction

As populations grow and technologies advance, societies increasingly find themselves wielding, intentionally or not, the power to impact the natural systems in which they live and depend. Resultant Coupled Human-Natural Systems [*Liu et al.*, 2007] (CHNS) are exceedingly complex and therefore present great difficulties for prediction, resulting in grave mismanagement of land and water resources, as seen in the cases of the shrinking Aral Sea [*Cai et al.*, 2003], hypoxia in the Gulf of Mexico [*Rabalais et al.*, 2001], and global depletion of groundwater resources [*Wada et al.*, 2010]. Anthropogenic and natural processes interact and feedback with one another across various scales of time and space, often producing unexpected emerging properties. Predicting the response of a CNHS is further clouded by changing and uncertain climatic forcing and human interferences. In particular, in a changing climatic future, governing principles for CNHS will be needed to reliably evaluate the sustainability and other consequences of policies and management practices.

Reliable precedence exists for the description of certain complex systems according to their statistical preferences. For example, crop yield [*Taylor et al.*, 1999; *Döring et al.*, 2015] and streamflow [*McMahon et al.*, 2007] have independently been demonstrated to obey a power law relation between variation and mean (in yield or flow respectively). Beyond these statistical findings, a more profound challenge is to track the origination of the emerging pattern through the inter-connections of these eco-hydrologic processes. Furthermore, with increased coupling among human and natural systems, it is unclear to what degree natural organization principles continue to govern human-managed landscapes. Driven by this knowledge gap, in this study we discover evidence of statistical preference in coupled human-hydrologic systems and discuss its utility to policy makers and managers.

7.2 Methods and data

We present an evapotranspiration (ET) modeling and statistics exercise adhered to by both naturally organized and intensively managed landscapes. ET is a major component of both the hydrologic cycle and terrestrial ecological systems. ET plays a key role in water balance, energy dispersion, and plant growth. Meanwhile, ET via agriculture is the largest source of anthropogenic water consumption and is heavily altered by humans as farmers and land owners pursue higher and more stable yields. Therefore, ET presents a link through which we henceforth evaluate the coupling between human and natural systems. In our previous work, we developed a ET Temporal VARiance Decomposition (ETVARD) [*Zeng and Cai*, 2016] framework incorporating both climatic variables and terrestrial water storage change. We therefore decompose ET variance into its contributions from climate (precipitation and potential evapotranspiration), climate phasing, and catchment response to climate, based on the coupling of water-energy cycle. The incorporation of watershed storage change into water balance and decomposition of ET variability provides means to account for the impact of human development (via the exploitation of the terrestrial water storage) on the ET process, thus enabling further exploration of ET statistical preference in the context of CHNS.

The field scale ET observation is obtained from AmeriFlux Level 2B datasets. The fine resolution flux-tower observation is aggregated into monthly scale, and sites with records longer than 48 months are used to calculate ET variance. The US 1°by1° grid scale ET mean and variance is calculated from climate and terrestrial storage change observations. The climate (i.e., precipitation and potential evaporation) observation is from North American Land Data Assimilation System (NLDAS) monthly forcing [*Mitchell et al.*, 2004] and the terrestrial storage change is from Gravity Recovery and Climate Experiment (GRACE) satellite based on the CSR RL5.0 release from the Center for Space Research at the University of Texas at Austin [*Landerer and Swenson*, 2012]. The ET variance from 2002-2015 is calculated based on ETVARD [*Zeng and Cai*, 2015]. The 32 regional scale basins ET from 1984–2006 is from multi-source hydrologic observation data assimilation [*Pan et al.*, 2011]. The ET variance in the High Plain is also calculated based on ETVARD, where the terrestrial storage change is from a spatial interpolation USGS groundwater monitoring wells[*Haacker et al.*, 2015] for pre-development (1940-1975) and managed periods (1975-2015).

The land use and land cover classification map for the US is from the Boston University's MODIS land cover product, which uses a 17-type IGBP (International Geosphere-Biosphere Programme) classification (http://www.bu.edu/lcsc/data-documentation/). The 0.125°by0.125° is aggregated into 1°by1° grid, and the land use type with largest counts within each grid is marked as main land use type.

7.3 Results and discussion

Based on multi-source and multi-scale hydro-climatic observations, we find that ET variance (σ^2) at monthly scale is proportional to the fractional power of the mean (μ) according to Taylor's Power Law (TPL) [*Taylor*, 1961]:

$$\sigma^2 = a\mu^b \qquad \text{Eqn.(7.1)}$$

Or equivalently as shown in Figure 7.1, the ET coefficient of variance (CV), defined as the standard deviation scaled by the mean, decreases with increasing mean ET (i.e., $CV = p\mu^q$). Although the three datasets used in Figure 7.1 are collected from independent observations, contain various land use and land cover types, and cover a wide spectrum of climatic and ecologic conditions, they ubiquitously converge to the TPL. An important feature is that the TPL, the only scale-invariant relationship between mean and variance, holds for the three datasets which range spatially from field (AmeriFLUX sites), grid (US 1°by1° grids) to regional (32 global basins) scale.

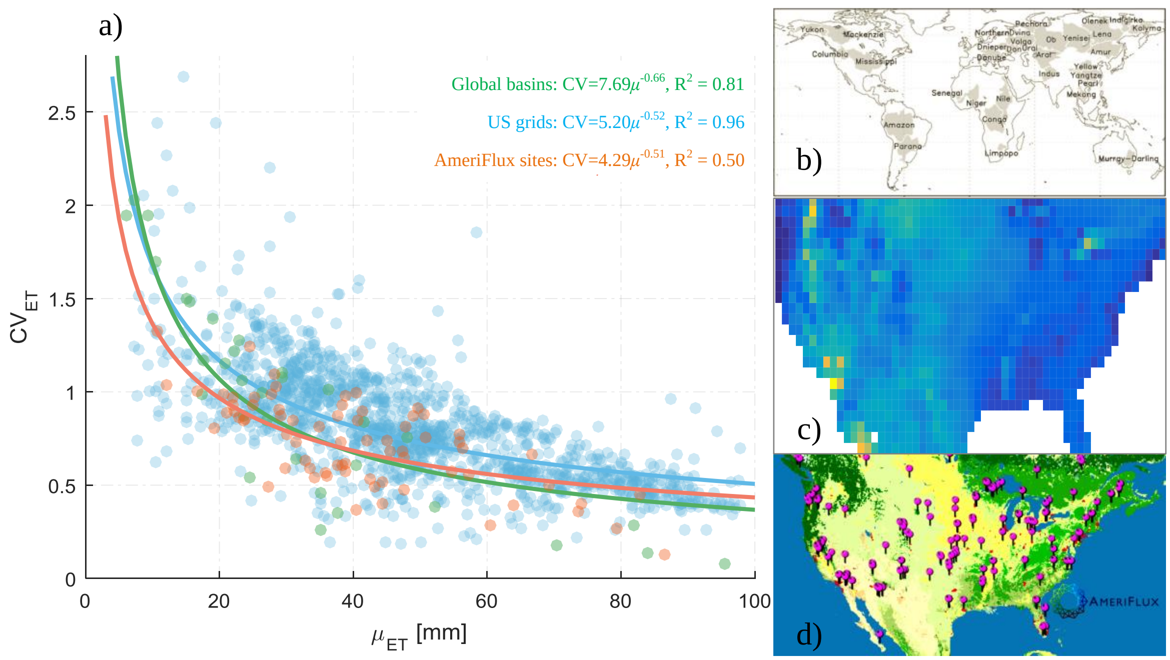
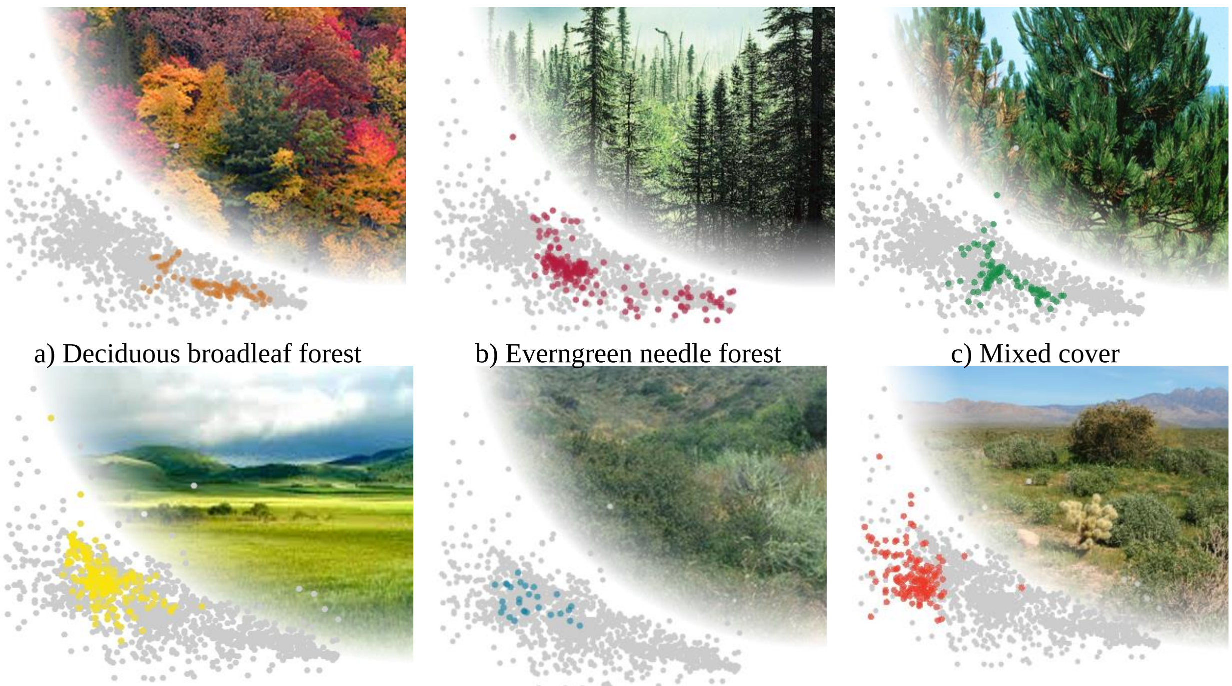


Figure 7.1. a) The power law relationship between monthly ET mean (μ_{ET}) and coefficient of variation (CV_{ET}) from observations at b) regional (32 global river basins), c) basin (1°by1° grids in US[*Mitchell et al.*, 2004]) and d) field (AmeriFlux sites) scale.

We further find that the ET power law describes a behavior of ecosystem self-organization. Note that precipitation, in contrast to ET, does not obey any sort of power law relating mean and variance, but rather, precipitation exhibits both ranges where variance decreases with mean and where variance increases with mean. Thus the ET power law cannot be explained by the statistics of precipitation (or climatic forcing in general) alone. We hypothesize that vegetation water use strategy may provide supplemental explanation for the observed power law. For the US watersheds data set, we classify the ET data points according to dominant land use and land cover (LULC) type, revealing distinct clustering of ecosystems in Figure 7.2. At the right end of the curve exist ecosystems with high and stable biomass production (high mean and low variability) such as evergreen and deciduous forests; conversely, at the left extreme of the curve exist ecosystems of relatively low and unstable biomass production (low mean and high variability) such as shrublands and grasslands. The discovered similar convergence of monthly ET variance and mean across scales suggests that ecosystems adhere to some common preferential ET behavior. Thus, the observed ecosystem clustering pattern along with the power law curve suggests that the coevolution/coexistence of vegetation and abiotic conditions establishes the ET power law.



a) Deciduous broadleaf forest | b) Everngreen needle forest | c) Mixed cover

d) Grassland | e) Closed shrubland | f) Open shrubland

Figure 7.2. The clustering of various US LULC types along the ET mean~*CV* power law curve in Figure 7.1. a)-c) The forests are generally located at the right end of ET mean~*CV* power law curve with high ET and small ET variability; while d) the grassland, and e) and f) shrubland are located at the left end of ET mean~*CV* power law curve with low ET and large ET variability.

Vegetation physiology is one likely mechanism for the observed self-organization; a plant's isohydric or anisohydric nature [*Konings and Gentine*, 2016] (i.e., tendency to vary stomatal openings to conserve water or tendency to keep stomata open) determines how likely the plant is to outcompete others in a given environment and accordingly affects ET mean and variance. The emerging convergence of terrestrial ecosystems along the ET power law curve amid substantial climatic, landscape, vegetative, and anthropogenic diversity shows a close correspondence between the ET power law and the optimality of vegetation production [*Schymanski et al.*, 2009]. Given a climatic setting, an ecosystem may make trade-offs in water use pattern between high-consumption-and-low-variability and low-consumption-and-high variability as indicated by the ET power law curve. Thus the ET power law might be understood as an emerging property resulting from the interaction and co-evolution among plant water use strategy and biodiversity with climatic conditions, soil fertility, and a host of other abiotic processes. The emerging property provides a statistical preference to understand how terrestrial ecosystems respond to natural and artificial external drivers [*Huxman et al.*, 2004; *Franklin et al.*, 2014; *Manzoni et al.*, 2014] and organize themselves to achieve ecohydrologic optimality [*Eagleson*, 2002].

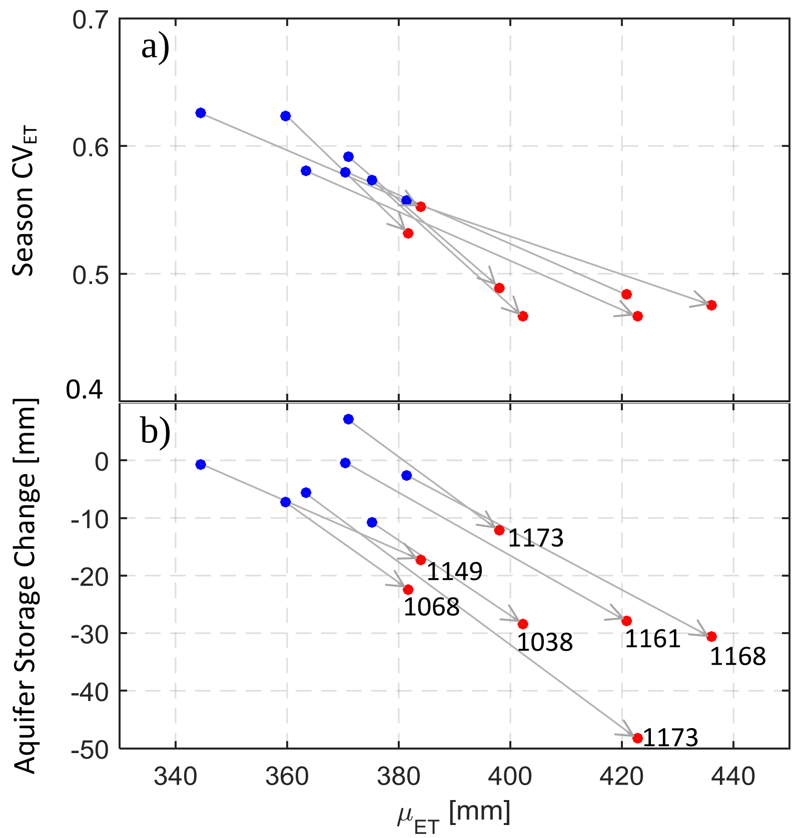


Figure 7.3. Switching from rain-fed crops to irrigated agriculture from groundwater pumping in the sub-basins (denoted by USGS HUC ID) within Republican River Basin results in a) higher and more stable ET at the expense of b) groundwater storage depletion.

In the context of ETVARD, vegetations' responses to climate (e.g., water uptake in the root zone and evaporative demand) is represented by the covariance terms (i.e., $cov_{P,\Delta S}$ and $cov_{PET,\Delta S}$). The power law relationship generates a hypothesis to test if a similar pattern can be found from existing processes-based eco-hydrologic models.

Finally, we demonstrate that the ET power law provides insight into the consequences of anthropogenic land cover and land use change, as displayed in major agriculture transitions in the American High Plains and Midwest. As farmers seek higher and more stable yields via irrigation, they increase the mean evapotranspirative water consumption of their land and decrease its variance in accordance with the ET power law, moving “down” the power law curve in Figures 7.3. Such transition occurred at a wide scale in the American High Plains in the 1970’s when a large fraction of agriculture land was switched from rain-fed to irrigated land. The shift of ET mean and variability following the power law presents an analogy with the power low relationship between the mean and variability of rain-fed and irrigated corn yield throughout the High Plains. However, the transformation has had dramatic ramifications on the water balance in the region. ET consumption is accompanied by groundwater or streamflow depletion as farmers must supplement water supply by withdrawing from aquifers and rivers in Figure 7.3. Moreover, farmers must continually withdraw water from storage in every crop year in order to maintain the new agricultural ecosystem out of the preceding ecosystem’s natural position along the ET power law curve, and do so for the entirety of the new ecosystem’s lifetime. As a result, water table levels in the High Plains have decreased significantly since the 1970 transition to irrigation [*McGuire*, 2014], after having been steady for previous decades. The ET power law thus provides a framework for understanding, to the first order, the water balance ramifications of anthropogenic induced landscape change.

Likewise, agriculture facilitated by drainage engineering in many Midwest watersheds also complies with the ET power law regarding natural processes and human interferences. Since the 1800s, vast amounts of land in the Midwest have been drained (via ditches and tile drainage systems) to convert prairies and wetlands to land more suitable for agriculture [*Blann et al.*, 2009]. Following drainage, observed local streamflow and flood frequency have increased while precipitation has remained relatively constant [*Raymond et al.*, 2008; *Gupta et al.*, 2015]. According to water balance, the increase in streamflow requires a commensurate decrease in ET, and the ET power law dictates that such a decrease in mean ET is accompanied by an increase in ET variability, which manifests itself in the seasonality of crops compared to the prior wetlands. Like the case of irrigation in the High Plains, human intervention transformed the land to a new ecosystem state in accordance with the ET power law, in this case moving “up” the power law curve (Figure 7.4). Contrary to the irrigation case of the High Plains, the drainage of the Midwest is achieved by a single initial land transformation rather than ongoing human intervention. However, both drainage and irrigation cases increase food production and meanwhile cause with environmental changes associated with the ET mean and variability relation. The agriculture supported by drainage in the Midwest has caused considerable consequences such as changes in flow regime and sediment and nutrient load, which directly contributes to the hypoxia in northern Mexico Gulf [*Rabalais et al.*, 2001].

Via the two cases of High Plains irrigation and Midwest drainage as described above (Figure 7.4), it is evident that the ET power law and clustering of ecosystems provide a framework for analyzing human interference in hydrologic systems. The ET power law provides insightful information regarding the constraints of human actions and impacts of human interferences on the hydrologic cycle and water balance. Future work to distinguish whether a desirable state, changed from the natural equilibrium state, would require continual human forcing as with the irrigation of the High Plains or could be achieved by a sufficient singular disturbance as with the drainage of the Midwest would usefully complement the ET power law. With such additional information, the ET power law could help predict the level of resource need (e.g., annual, continued storage depletion) and engineering need (e.g., irrigation or drainage infrastructure) to sustain an anthropogenic ecosystem transformation. Such information is of paramount importance for evaluating the sustainability of a wide range of coupled human-hydrologic systems, which are to be shifted to an alternative state with desired socioeconomic benefits. For example, recent studies attempt to predict the socioeconomic and environmental changes that will be associated with land adoption for cellulosic biofuel crops (e.g., switchgrass and Miscanthus). It is predicted that ET will significantly increase with Miscanthus in Midwest watersheds [*Le et al.*, 2011; *Housh et al.*, 2015], and therefore the ET variance will decrease as suggested by the ET power law. Compared to the current corn and soybean dominated landscape, on the lower law curve, the landscape will then move "down". The environmental and socioeconomic consequences resulting from the ET change and then associated streamflow change will be critical for the sustainability of the biofuel-economy in region.

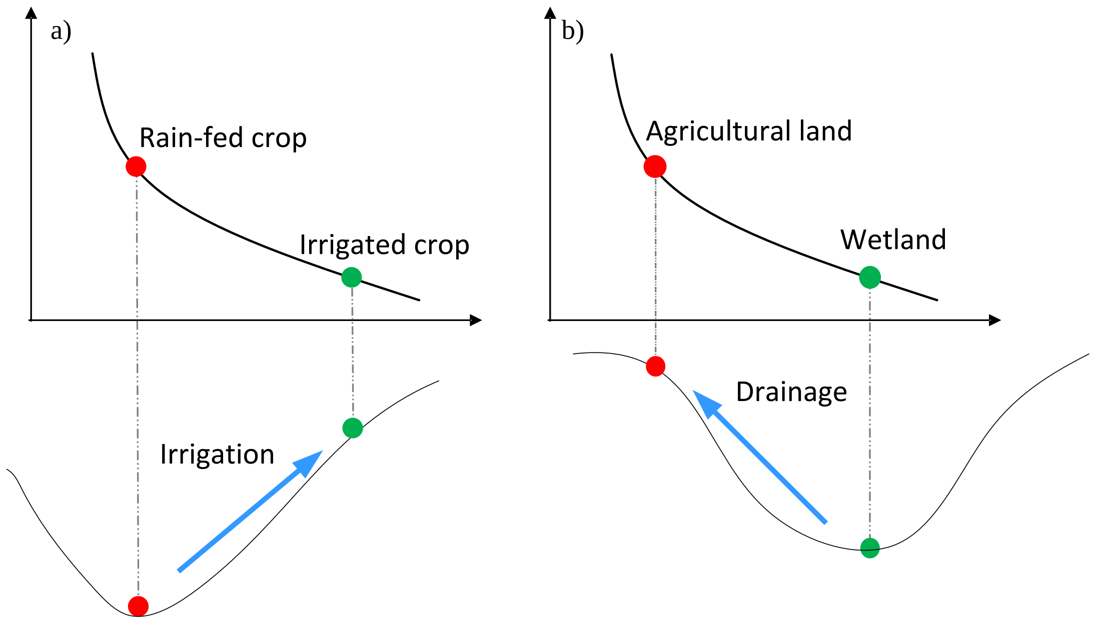


Figure 7.4. Anthropogenic interferences, including a) irrigation in the High Plains and b) drainage in the Midwest, modifying the system state in different directions along the ET power law curve to achieve suitable condition for crop production.

7.4. Conclusion

In summary, the ET power law, a specific case of Taylor's power law, identifies a preferential mean and variance relationship across ecosystem class and scale. The relationship describes the self-organization and co-evolution of ecosystem clusters in specific regions. Landscapes persistently follow the power law curve even upon human-induced transition from natural to managed (e.g., wetlands to agriculture land) landscape or from one managed state to an alternative state. The ET power law then provides valuable insight regarding the emerging behavior of complex, coupled human-hydrologic systems. More importantly, understanding which can be gleaned from the ET power law, and hopefully other complimentary descriptions of emerging behavior yet to be discovered, can be essential to responsible, sustainable management of our most valuable resources and systems. As the demands of society rise and anthropogenic influence on our environment rises, the stakes of management rise, the identified ET power law may prove to be a valuable guide for predicting the impacts and sustainability of anthropogenic landscape change.

CHAPTER 8
CONCLUSION

8.1 Conclusions

This dissertation provides new understanding of climatic and hydrologic controls on ET temporal variability, for both natural and managed watersheds. The six main chapters approach the research objectives from different but connected aspects by: 1) developing a theoretical ETVARD framework in Chapter 2; 2) quantifying the climatic and hydrologic controls on ET variance from real world case studies, where Chapter 3 focuses on the inter- and intra-annual scales in 32 global basins and Chapter 4 focuses on the seasonal scale in the High Plains; 3) examining the congruences among ETVARD (a theory), multi-source multi-variable hydroclimatic observations (Chapter 5) and multiple land surface models (Chapter 6.)

These chapters are inherently inter-connected and serve as cross-validation for each other. In Chapter 2, we hypothesize that dominant controls from storage components on ET variance should be limited to basins with relatively small precipitation or potential evaporation flux. Chapter 3 confirms this hypothesis is valid at the intra-annual scale, while not obvious at the inter-annual scale. In addition to the climate and terrestrial water storage, Chapter 3 refines the knowledge from theoretical development in Chapter 2 and adds the temporal scale into the analysis of ET variance. This helps us identify the proper time scale in the irrigation impact study in Chapter 4. Terrestrial water storage changes are caused by many processes, such as groundwater recharge/discharge, soil moisture change, snow thawing/melting, which are associated with different time scales. With soil moisture dominant at the monthly scale and climatic fluctuation dominant at the annual scale, the impact of groundwater-based irrigation is most significant at seasonal scale. The seasonal scale also has an advantage for the analysis by eliminating the impacts of snow processes. Although there are many studies on impact of irrigation on land surface processes, Chapter 4 is unique in providing a comprehensive picture of ET variance before and after extensive irrigation. The groundwater monitoring wells capture the aquifer depletion signals, and ETVARD further carries the human interference signals embedded in groundwater to changes of ET temporal and spatial patterns.

In the context of watersheds as coupled nature-human systems (CNHS), Chapter 4 starts from extending hydrologic knowledge (i.e, ETVARD) to managed systems with an assumption that human interferences are well captured by observations (groundwater table in this case). Though conducting controlled experiments is difficulty in hydrology at a long-time scale and a wide spatial scale, the High

Plain case study provides a unique case. Compared to natural systems where many ecologic services balancing each other, managed agricultural systems have relatively simple functioning or services, that is, providing better economic benefit at most cases. Under the dominance of human activities, famers' goal (conceptualized as tradeoff between profit maximization and risk aversion) propagates to the hydrologic signal, as shown by the correspondence between ET and crop mean-variance. This inspires the author another way to understand CNHS: can we start the hypothesis from a managed system with a known signature (such as optimization or trade-off among several objectives) and find the analogy in natural systems? Chapter 7 shows our effort in this approach. Chapter 7 extends the hypothesis of tradeoff between mean and variability from a managed system to general ecosystems. The single hypothesis is still applicable to different ecosystems. Chapter 7 further illustrates the differences as indicated by the clusters in the different zones of the power law relationship. Human deals with the tradeoffs between profit and risk, so do natural systems. Chapter 7 provides a possible bridge to link the optimization in water resources management to the idea of co-evolution in ecohydrology. Surely, the power law relationship is obtained at the system level, and further studies should explore how these relationships emerge from the process level.

Another perspective of the thesis is about contribution to the methodology, mainly in Chapter 5 and Chapter 6. Unlike streamflow, a unique feature of ET is that we do not know its accurate value at watershed level. Whether the model, observation or the theory is true is challenging for validation. Then, what is the criterion of hydrologic knowledge confirmation? Hydrology is a science that should have its own basic laws, theories and hypotheses; on the other hand, hydrology is an earth science where logics cannot fully synthesize all the observations and phenomenon. Furthermore, a large portion of hydrologic experiments are conducted by numerical models. With that being said, we bring up the theory-observation-simulation triplets as a diagnosis framework for hydrologic knowledge discovery. The role of ETVARD derived in Chapter 2 connects between observations (Chapter 5) and simulations (Chapter 6). We believe that the congruence in the theory-observation-simulation triplets is more likely to provide the true knowledge than the traditional approach that is based on the agreement between observations and simulations.

8.2 Future work

The six main chapters provide a consistent and comprehensive framework in addressing the research objectives to understand climatic and terrestrial water storage controls on ET variability at different scales for both nature and managed systems. In terms of future work, three perspectives can be extended from the work of this thesis.

8.2.1 ETVARD as a constraining mechanism for hydroclimatic data assimilation

In Chapter 5, ETVARD serves as a theoretic constraint to check the consistency of multi-variable hydroclimatic observations. As a theatrical constraint, ETVARD can also be extended for other approaches on hydrologic data processing, for example, data assimilation. The data used in Chapter 3 for 32 global watersheds is processed through a constrained Kalman Filter with water balance as a constraint [*Pan et al.*, 2011]. In the constrained Kalman Filter, water balance is treated as a hard constraint in addition to minimizing the data uncertainty in the conventional Kalman Filter. In adopting water balance as a hard constraint, water balance as a fundamental law must be satisfied at each time step. As a confirmed hydrologic theory, ETVARD captures the physical dynamics of multiple hydrologic variables, and it can be adopted as a soft constraint in Kalman Filter. The soft constraint means that some degree of violating ETVARD is allowable when some assumptions with ETVARD are not satisfied (e.g., terrestrial water storage change is caused by groundwater lateral flow and trans-basin water delivery).

8.2.2 Incorporating ET variance into the Budyko curve to understand the long-term ET average

The Budyko curve is plotted as a single line that relates evaporative ratio (i.e., *ET/P*) to the aridity index (i.e., *PET/P*). Although it has been shown that basins around the world generally follow the Budyko curve, the various deviations from the Budyko curve have been assessed and error [*Yang et al.*, 2014a] are identified and attributed to other factors such as seasonality [*Ning et al.*, 2017]. Chapter 7 have identified a statistical power law relationship between the monthly ET variance and the long-term average. This provides an approach to incorporate ET variability to explain the deviations from the Budyko curve. Under the same aridity index, the hydroclimatic variability at a smaller temporal scale (e.g., seasonal, monthly or daily) will adjust ET variance through the variance/covariance terms by enhancing/buffering the water consumption. As shown in the power law curve in Chapter 7, the increases in the coefficient of variance is associated with the decreases in ET long-term average. Therefore, given the same aridity index, the power law relationship between ET variance and average can be used to examine e deviations in Budyko curve. For example, in Chapter 4, although basins in the High Plain have experienced negligible changes in climate condition (i.e., the aridity index), their evaporative ratios (i.e., *ET/P*) increase due to irrigation associated with a decrease in ET coefficient of variance. Therefore, the differences in ET variance (either by natural or anthropogenic factors) can be reflected in the long-term average through the power law relationship. Higher or lower long-term average than the Budyko curve will be explained by differences in ET variance, rather than as errors in previous study. For example, the out-of-phase *P* and *PET* pattern in Mediterranean climate damps ET variance and also decreases ET mean value. If two regions have the same aridity index, the evaporative ratio in the basin with *P* and *PET* out-of-phase will be lower than that with *P* and *PET* in-phase.

8.2.3 Assessing climatic and terrestrial water storage controls on runoff variability at different temporal scales

The framework by *Koster and Suarez* [1999] is followed by researchers to study the runoff variability [*Sankarasubramanian et al.*, 2001; *Sankarasubramanian and Vogel*, 2003]. The study in this thesis incorporates terrestrial water storage into the ET variance assessment framework, and it can be extended to runoff assessment. The author has obtained the formulation of the runoff variability decomposition, which is similar to the ET variability decomposition with the same six variance/covariance terms. The differences lay in the analytical expressions of the weighting factors quantifying the contribution of each variance/covariance term to runoff variance. With the expression for runoff variance decomposition, the basin studies in Chapter 2 and 3 can be conducted in the same manner as ET variance assessment in this dissertation. The cross-validation between observations and model simulations can also be conducted for runoff variability. The runoff measurement is relatively more accurate than ET observations, while current land surface models still face challenges in stream flow prediction. The theory-observation-simulation congruence will provide useful information to diagnose the runoff simulations in land surface models. Since surface runoff is measured at the watershed outlet, the assessment should be better conducted according to the watershed boundary rather than the grids used for ET assessment.

In addition, the lateral flow of groundwater may also contribute to the terrestrial storage change and thus runoff variability. In addition, *McMahon et al.* [2007] reported a relationship between runoff mean value and inter-annul standard deviation, which is similar to the ET power law relationship discovered in Chapter 7 (although the ET power law is found at month scale). The expression for runoff variance can be further used to explore the emergence to the relationship between runoff mean value and inter-annul standard deviation. Some preliminary results on inter- and intra-annual runoff variance is shown in Appendix B with the same data set in Chapter 3.

REFERENCES

Aeschbach-Hertig, W., and T. Gleeson (2012), Regional strategies for the accelerating global problem of groundwater depletion, *Nature Geosci*, *5*(12), 853-861.

Allen, R., M. Tasumi, A. Morse, and R. Trezza (2005), A Landsat-based energy balance and evapotranspiration model in Western US water rights regulation and planning, *Irrig Drainage Syst*, *19*(3-4), 251-268.

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, *FAO, Rome*, *300*(9), D05109.

Anderson, M. C., C. Hain, J. Otkin, X. Zhan, K. Mo, M. Svoboda, B. Wardlow, and A. Pimstein (2013), An Intercomparison of Drought Indicators Based on Thermal Remote Sensing and NLDAS-2 Simulations with U.S. Drought Monitor Classifications, *Journal of Hydrometeorology*, *14*(4), 1035-1056.

Arora, V. K. (2002), The use of the aridity index to assess climate change effect on annual runoff, *Journal of Hydrology*, *265*(1–4), 164-177.

Badgley, G., J. B. Fisher, C. Jiménez, K. P. Tu, and R. Vinukollu (2015), On uncertainty in global terrestrial evapotranspiration estimates from choice of input forcing datasets, *Journal of Hydrometeorology*.

Best, M. J., M. Pryor, D. B. Clark, G. G. Rooney, R. L. H. Essery, C. B. Ménard, J. M. Edwards, M. A. Hendry, A. Porson, N. Gedney, L. M. Mercado, S. Sitch, E. Blyth, O. Boucher, P. M. Cox, C. S. B. Grimmond, and R. J. Harding (2011), The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, *Geosci. Model Dev.*, *4*(3), 677-699.

Best, M. J., G. Abramowitz, H. R. Johnson, A. J. Pitman, G. Balsamo, A. Boone, M. Cuntz, B. Decharme, P. A. Dirmeyer, J. Dong, M. Ek, Z. Guo, V. Haverd, B. J. J. van den Hurk, G. S. Nearing, B. Pak, C. Peters-Lidard, J. A. Santanello, L. Stevens, and N. Vuichard (2015), The Plumbing of Land Surface Models: Benchmarking Model Performance, *Journal of Hydrometeorology*, *16*(3), 1425-1442.

Beven, K. (2012), Causal models as multiple working hypotheses about environmental processes, *Comptes Rendus Geoscience*, *344*(2), 77-88.

Beven, K., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *Journal of Hydrology*, *249*(1–4), 11-29.

Blann, K. L., J. L. Anderson, G. R. Sands, and B. Vondracek (2009), Effects of Agricultural Drainage on Aquatic Ecosystems: A Review, *Critical Reviews in Environmental Science and Technology*, *39*(11), 909-1001.

Bonan, G. B., P. J. Lawrence, K. W. Oleson, S. Levis, M. Jung, M. Reichstein, D. M. Lawrence, and S. C. Swenson (2011), Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data, *Journal of Geophysical Research: Biogeosciences*, *116*(G2), n/a-n/a.

Brooks, P. D., P. A. Troch, M. Durcik, E. Gallo, and M. Schlegel (2011), Quantifying regional scale ecosystem response to changes in precipitation: Not all rain is created equal, *Water Resources Research*, *47*(10), W00J08.

Brutsaert, W. (2005), *Hydrology: an introduction*, 605 pp., Cambridge University Press, Cambridge, U.K.

Brutsaert, W., and H. Stricker (1979), An advection-aridity approach to estimate actual regional evapotranspiration, *Water Resources Research*, *15*(2), 443-450.

Budyko, M. I. (1974), *Climate and life*, Academic press, San Diego, CA.

Burt, O. R., M. Baker, and G. A. Helmers (2002), Statistical estimation of streamflow depletion from irrigation wells, *Water Resources Research*, *38*(12), 1296.

Cai, X., D. C. McKinney, and M. W. Rosegrant (2003), Sustainability analysis for irrigation water management in the Aral Sea region, *Agricultural Systems*, *76*(3), 1043-1066.

Cai, X., X. Zhang, P. H. Noël, and M. Shafiee-Jood (2015), Impacts of climate change on agricultural water management: a review, *Wiley Interdisciplinary Reviews: Water*, n/a-n/a.

Cai, X., Z.-L. Yang, Y. Xia, M. Huang, H. Wei, L. R. Leung, and M. B. Ek (2014), Assessment of simulated water balance from Noah, Noah-MP, CLM, and VIC over CONUS using the NLDAS test bed, *Journal of Geophysical Research: Atmospheres*, *119*(24), 2014JD022113.

Cheng, L., Z. Xu, D. Wang, and X. Cai (2011), Assessing interannual variability of evapotranspiration at the catchment scale using satellite-based evapotranspiration data sets, *Water Resources Research*, *47*(9), W09509.

Choudhury, B. (1999), Evaluation of an empirical equation for annual evaporation using field observations and results from a biophysical model, *Journal of Hydrology*, *216*(1–2), 99-110.

Clark, M. P., D. Kavetski, and F. Fenicia (2011), Pursuing the method of multiple working hypotheses for hydrological modeling, *Water Resources Research*, *47*(9), n/a-n/a.

Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resources Research*, *44*(12).

Clark, M. P., B. Schaefli, S. J. Schymanski, L. Samaniego, C. H. Luce, B. M. Jackson, J. E. Freer, J. R. Arnold, R. Dan Moore, E. Istanbulluoglu, and S. Ceola (2016), Improving the theoretical underpinnings of process-based hydrologic models, *Water Resources Research*, n/a-n/a.

Clark, M. P., B. Nijssen, J. D. Lundquist, D. Kavetski, D. E. Rupp, R. A. Woods, J. E. Freer, E. D. Gutmann, A. W. Wood, L. D. Brekke, J. R. Arnold, D. J. Gochis, and R. M. Rasmussen (2015a), A unified approach for process-based hydrologic modeling: 1. Modeling concept, *Water Resources Research*, *51*(4), 2498-2514.

Clark, M. P., Y. Fan, D. M. Lawrence, J. C. Adam, D. Bolster, D. J. Gochis, R. P. Hooper, M. Kumar, L. R. Leung, D. S. Mackay, R. M. Maxwell, C. Shen, S. C. Swenson, and X. Zeng (2015b), Improving the representation of hydrologic processes in Earth System Models, *Water Resources Research*, *51*(8), 5929-5956.

Clark, M. P., B. Nijssen, J. D. Lundquist, D. Kavetski, D. E. Rupp, R. A. Woods, J. E. Freer, E. D. Gutmann, A. W. Wood, D. J. Gochis, R. M. Rasmussen, D. G. Tarboton, V. Mahat, G. N. Flerchinger, and D. G. Marks (2015c), A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies, *Water Resources Research*, *51*(4), 2515-2542.

Condon, L. E., and R. M. Maxwell (2014), Feedbacks between managed irrigation and water availability: Diagnosing temporal and spatial patterns using an integrated hydrologic model, *Water Resources Research*, *50*(3), 2600-2616.

Cosgrove, B. A., D. Lohmann, K. E. Mitchell, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, C. Marshall, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, J. D. Tarpley, and J. Meng (2003), Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, *Journal of Geophysical Research: Atmospheres*, *108*(D22), 8842.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris (2008), Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, *International Journal of Climatology*, *28*(15), 2031-2064.

DeAngelis, A., F. Dominguez, Y. Fan, A. Robock, M. D. Kustu, and D. Robinson (2010), Evidence of enhanced precipitation due to irrigation over the Great Plains of the United States, *Journal of Geophysical Research: Atmospheres*, *115*(D15), D15115.

Demissie, Y., A. Valocchi, X. Cai, N. Brozovic, G. Senay, and M. Gebremichael (2015), Parameter Estimation for Groundwater Models under Uncertain Irrigation Data, *Groundwater*, *53*(4), 614-625.

Döll, P., H. Hoffmann-Dobrev, F. T. Portmann, S. Siebert, A. Eicker, M. Rodell, G. Strassberg, and B. R. Scanlon (2012), Impact of water withdrawals from groundwater and surface water on continental water storage variations, *Journal of Geodynamics*, *59–60*(0), 143-156.

Donohue, R. J., M. L. Roderick, and T. R. McVicar (2010), Can dynamic vegetation information improve the accuracy of Budyko's hydrological model?, *Journal of Hydrology*, *390*(1–2), 23-34.

Döring, T. F., S. Knapp, and J. E. Cohen (2015), Taylor's power law and the stability of crop yields, *Field Crops Research*, *183*, 294-302.

Draper, S. E. (2007), Administration and Institutional Provisions of Water Sharing Agreements, *Journal of Water Resources Planning and Management*, *133*(5).

Dunn, A. L., C. C. Barford, S. C. Wofsy, M. L. Goulden, and B. C. Daube (2007), A long-term record of carbon exchange in a boreal black spruce forest: means, responses to interannual variability, and decadal trends, *Global Change Biology*, *13*(3), 577-590.

Eagleson, P. S. (2002), *Ecohydrology: Darwinian expression of vegetation form and function*, 484 pp., Cambridge University Press, Cambridge, U. K.

Famiglietti, J., M. Lo, S. Ho, J. Bethune, K. Anderson, T. Syed, S. Swenson, C. de Linage, and M. Rodell (2011), Satellites measure recent rates of groundwater depletion in California's Central Valley, *Geophysical research letters*, *38*(3), L03403.

Famiglietti, J. S. (2014), The global groundwater crisis, *Nature Clim. Change*, *4*(11), 945-948.

FAO (2016), AQUASTAT Main Database, Food and Agriculture Organization of the United Nations (FAO), edited.

Fasullo, J., and P. J. Webster (2003), A Hydrological Definition of Indian Monsoon Onset and Withdrawal, *Journal of Climate*, *16*(19), 3200-3211.

Fatichi, S., and V. Y. Ivanov (2014), Interannual variability of evapotranspiration and vegetation productivity, *Water Resources Research*, *50*(4), 3275-3294.

Feng, X., G. Vico, and A. Porporato (2012), On the effects of seasonality on soil water balance and plant growth, *Water Resources Research*, *48*(5), W05543.

Ferguson, I. M., and R. M. Maxwell (2010), Role of groundwater in watershed response and land surface feedbacks under climate change, *Water Resources Research*, *46*(10), W00F02.

Foster, S., and H. Garduño (2004), Towards sustainable groundwater resource use for irrigated agriculture on the North China Plain, *World Bank, Washington, DC*.

Foster, T., N. Brozović, and A. P. Butler (2014), Modeling irrigation behavior in groundwater systems, *Water Resources Research*, *50*(8), 6370-6389.

Franklin, O., S. Palmroth, and T. Näsholm (2014), How eco-evolutionary principles can guide tree breeding and tree biotechnology for enhanced productivity, *Tree Physiology*, *34*(11), 1149-1166.

Fu, B. (1981), On the calculation of the evaporation from land surface (in Chinese), *Sci. Atmos. Sin.*, *5*(1), 23–31.

Gao, H., Q. Tang, C. R. Ferguson, E. F. Wood, and D. P. Lettenmaier (2010a), Estimating the water budget of major US river basins via remote sensing, *International Journal of Remote Sensing*, *31*(14), 3955-3978.

Gao, H., Q. Tang, X. Shi, C. Zhu, T. Bohn, F. Su, J. Sheffield, M. Pan, D. Lettenmaier, and E. F. Wood (2010b), Water budget record from Variable Infiltration Capacity (VIC) model, *Algorithm Theoretical Basis Document for Terrestrial Water Cycle Data Records*.

Glauber, J. W. (2004), Crop Insurance Reconsidered, *American Journal of Agricultural Economics*, *86*(5), 1179-1195.

Gordon, L. J., W. Steffen, B. F. Jönsson, C. Folke, M. Falkenmark, and Å. Johannessen (2005), Human modification of global water vapor flows from the land surface, *Proceedings of the National Academy of Sciences of the United States of America*, *102*(21), 7612-7617.

Gulden, L. E., E. Rosero, Z.-L. Yang, M. Rodell, C. S. Jackson, G.-Y. Niu, P. J. F. Yeh, and J. Famiglietti (2007), Improving land-surface model hydrology: Is an explicit aquifer model better than a deeper soil profile?, *Geophysical Research Letters*, *34*(9), n/a-n/a.

Gupta, H. V., T. Wagener, and Y. Liu (2008), Reconciling theory with observations: elements of a diagnostic approach to model evaluation, *Hydrological Processes*, *22*(18), 3802-3813.

Gupta, S. C., A. C. Kessler, M. K. Brown, and F. Zvomuya (2015), Climate and agricultural land use change impacts on streamflow in the upper midwestern United States, *Water Resources Research*, *51*(7), 5301-5317.

Gutentag, E. D. (1984), Geohydrology of the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming, *Geological Survey professional paper, 1400-B*.

Haacker, E. M. K., A. D. Kendall, and D. W. Hyndman (2015), Water Level Declines in the High Plains Aquifer: Predevelopment to Resource Senescence, *Groundwater*, n/a-n/a.

Han, S., H. Hu, D. Yang, and Q. Liu (2011), Irrigation impact on annual water balance of the oases in Tarim Basin, Northwest China, *Hydrological Processes*, *25*(2), 167-174.

Hargreaves, G. H., and Z. A. Samani (1982), Estimating potential evapotranspiration, *Journal of the Irrigation and Drainage Division*, *108*(3), 225-230.

Hejazi, M., X. Cai, X. Yuan, X. Liang, and P. Kumar (2014), Incorporating Reanalysis-Based Short-Term Forecasts from a Regional Climate Model in an Irrigation Scheduling Optimization Problem, *Journal of Water Resources Planning and Management*, *140*(5), 699-713.

Hejazi, M. I., and X. Cai (2009), Input variable selection for water resources systems using a modified minimum redundancy maximum relevance (mMRMR) algorithm, *Advances in Water Resources*, *32*(4), 582-593.

Herold, N., L. V. Alexander, M. G. Donat, S. Contractor, and A. Becker (2016), How much does it rain over land?, *Geophysical Research Letters*, n/a-n/a.

Housh, M., M. A. Yaeger, X. Cai, G. F. McIsaac, M. Khanna, M. Sivapalan, Y. Ouyang, I. Al-Qadi, and A. K. Jain (2015), Managing Multiple Mandates: A System of Systems Model to Analyze Strategies for Producing Cellulosic Ethanol and Reducing Riverine Nitrate Loads in the Upper Mississippi River Basin, *Environmental Science & Technology*, *49*(19), 11932-11940.

Housh, M., X. Cai, T. Ng, G. McIsaac, Y. Ouyang, M. Khanna, M. Sivapalan, A. Jain, S. Eckhoff, S. Gasteyer, I. Al-Qadi, Y. Bai, M. Yaeger, S. Ma, and Y. Song (2014), System of Systems Model for Analysis of Biofuel Development, *Journal of Infrastructure Systems*, *21*(3), 04014050.

Hoyos, C. D., and P. J. Webster (2007), The Role of Intraseasonal Variability in the Nature of Asian Monsoon Precipitation, *Journal of Climate*, *20*(17), 4402-4424.

Hsu, J. S., J. Powell, and P. B. Adler (2012), Sensitivity of mean annual primary production to precipitation, *Global Change Biology*, *18*(7), 2246-2255.

Huxman, T. E., M. D. Smith, P. A. Fay, A. K. Knapp, M. R. Shaw, M. E. Loik, S. D. Smith, D. T. Tissue, J. C. Zak, J. F. Weltzin, W. T. Pockman, O. E. Sala, B. M. Haddad, J. Harte, G. W. Koch, S. Schwinning, E. E. Small, and D. G. Williams (2004), Convergence across biomes to a common rain-use efficiency, *Nature*, *429*(6992), 651-654.

Istanbulluoglu, E., T. Wang, O. M. Wright, and J. D. Lenters (2012), Interpretation of hydrologic trends from a water balance perspective: The role of groundwater storage in the Budyko hypothesis, *Water Resources Research*, *48*(3), W00H16.

Joodaki, G., J. Wahr, and S. Swenson (2014), Estimating the human contribution to groundwater depletion in the Middle East, from GRACE data, land surface models, and well observations, *Water Resources Research*, n/a-n/a.

Jung, M., M. Reichstein, and A. Bondeau (2009), Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, *Biogeosciences*, *6*(10), 2001-2013.

Jung, M., M. Reichstein, P. Ciais, S. I. Seneviratne, J. Sheffield, M. L. Goulden, G. Bonan, A. Cescatti, J. Chen, R. de Jeu, A. J. Dolman, W. Eugster, D. Gerten, D. Gianelle, N. Gobron, J. Heinke, J. Kimball, B. E. Law, L. Montagnani, Q. Mu, B. Mueller, K. Oleson, D. Papale, A. D. Richardson, O. Roupsard, S. Running, E. Tomelleri, N. Viovy, U. Weber, C. Williams, E. Wood, S. Zaehle, and K. Zhang (2010), Recent decline in the global land evapotranspiration trend due to limited moisture supply, *Nature*, *467*(7318), 951-954.

Karam, H. N., and R. L. Bras (2008), Climatological Basin-Scale Amazonian Evapotranspiration Estimated through a Water Budget Analysis, *Journal of Hydrometeorology*, *9*(5), 1048-1060.

Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resources Research*, *42*(3), n/a-n/a.

Konikow, L., and E. Kendy (2005), Groundwater depletion: A global problem, *Hydrogeology Journal*, *13*(1), 317-320.

Konikow, L. F. (2011), Contribution of global groundwater depletion since 1900 to sea-level rise, *Geophysical Research Letters*, *38*(17), L17401.

Konikow, L. F. (2015), Long-Term Groundwater Depletion in the United States, *Groundwater*, *53*(1), 2-9.

Konings, A. G., and P. Gentine (2016), Global Variations in Ecosystem-Scale Isohydricity, *Global Change Biology*, n/a-n/a.

Koster, R. D., and M. J. Suarez (1999), A simple framework for examining the interannual variability of land surface moisture fluxes, *Journal of Climate*, *12*(7), 1911-1917.

Koster, R. D., B. M. Fekete, G. J. Huffman, and P. W. Stackhouse (2006), Revisiting a hydrological analysis framework with International Satellite Land Surface Climatology Project Initiative 2 rainfall, net radiation, and runoff fields, *Journal of Geophysical Research: Atmospheres*, *111*(D22), D22S05.

Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, C.-H. Lu, S. Malyshev, B. McAvaney, K. Mitchell, D. Mocko, T. Oki, K. Oleson, A. Pitman, Y. C. Sud, C. M. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada (2004), Regions of Strong Coupling Between Soil Moisture and Precipitation, *Science*, *305*(5687), 1138-1140.

Kumar, P. (2011), Typology of hydrologic predictability, *Water Resources Research*, *47*(3).

Kumar, P. (2015), Hydrocomplexity: Addressing water security and emergent environmental risks, *Water Resources Research*, *51*(7), 5827-5838.

Kuwayama, Y., and N. Brozović (2013), The regulation of a spatially heterogeneous externality: Tradable groundwater permits to protect streams, *Journal of Environmental Economics and Management*, *66*(2), 364-382.

Landerer, F. W., and S. C. Swenson (2012), Accuracy of scaled GRACE terrestrial water storage estimates, *Water Resources Research*, *48*(4), W04531.

Lawston, P. M., J. A. S. Jr., B. F. Zaitchik, and M. Rodell (2015), Impact of Irrigation Methods on Land Surface Model Spinup and Initialization of WRF Forecasts, *Journal of Hydrometeorology*, *16*(3), 1135-1154.

Le, P. V. V., P. Kumar, and D. T. Drewry (2011), Implications for the hydrologic cycle under climate change due to the expansion of bioenergy crops in the Midwestern United States, *Proceedings of the National Academy of Sciences*, *108*(37), 15085-15090.

Lettenmaier, D. P., and J. S. Famiglietti (2006), Hydrology: Water from on high, *Nature*, *444*(7119), 562-563.

Lettenmaier, D. P., and P. C. D. Milly (2009), Land waters and sea level, *Nature Geosci*, *2*(7), 452-454.

Li, D. (2014), Assessing the impact of interannual variability of precipitation and potential evaporation on evapotranspiration, *Advances in Water Resources*, *70*(0), 1-11.

Li, D., M. Pan, Z. Cong, L. Zhang, and E. Wood (2013), Vegetation control on water and energy balance within the Budyko framework, *Water Resources Research*, *49*(2), 969-976.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *Journal of Geophysical Research: Atmospheres*, *99*(D7), 14415-14428.

Liu, J., C. Zheng, L. Zheng, and Y. Lei (2008), Ground Water Sustainability: Methodology and Application to the North China Plain, *Ground Water*, *46*(6), 897-909.

Liu, J., T. Dietz, S. R. Carpenter, M. Alberti, C. Folke, E. Moran, A. N. Pell, P. Deadman, T. Kratz, and J. Lubchenco (2007), Complexity of coupled human and natural systems, *Science*, *317*(5844), 1513-1516.

Long, D., L. Longuevergne, and B. R. Scanlon (2015), Global analysis of approaches for deriving total water storage changes from GRACE satellites, *Water Resources Research*, *51*(4), 2574-2594.

Lute, A. C., and J. T. Abatzoglou (2014), Role of extreme snowfall events in interannual variability of snowfall accumulation in the western United States, *Water Resources Research*, *50*(4), 2874-2888.

Mahrt, L., and M. Ek (1984), The influence of atmospheric stability on potential evaporation.

Mahrt, L., and M. Ek (1984), The Influence of Atmospheric Stability on Potential Evaporation, *Journal of Climate and Applied Meteorology*, *23*(2), 222-234.

Manzoni, S., G. Vico, G. Katul, S. Palmroth, and A. Porporato (2014), Optimal plant water-use strategies under stochastic rainfall, *Water Resources Research*, *50*(7), 5379-5394.

McGuire, V. L. (2009), Changes in water levels and storage in the High Plains aquifer, predevelopment to 2007.

McGuire, V. L. (2012), Water-Level and Storage Changes in the High Plains Aquifer, Predevelopment to 2011 and 2009–11, *U.S. Geological Survey Scientific Investigations Report*, 2012-5291, 2015 p. (Also available at http://pubs.usgs.gov/sir/2012/5291/.).

McGuire, V. L. (2014), Water-level changes and change in water in storage in the High Plains aquifer, predevelopment to 2013 and 2011-13, *U.S. Geological Survey Scientific Investigations Report 2014 - 5218, 14p.*

McGuire, V. L., K. D. Lund, and B. K. Densmore (2012), Saturated thickness and water in storage in the High Plains aquifer, 2009, and water-level changes and changes in water in storage in the High Plains aquifer, 1980 to 1995, 1995 to 2000, 2000 to 2005, and 2005 to 2009, *U.S. Geological Survey Scientific Investigations Report*, 2012-5177, 2028 p. (Also available at http://pubs.usgs.gov/sir/2012/5177/.).

McMahon, T. A., R. M. Vogel, M. C. Peel, and G. G. S. Pegram (2007), Global streamflows – Part 1: Characteristics of annual streamflows, *Journal of Hydrology*, *347*(3–4), 243-259.

McVicar, T. R., M. L. Roderick, R. J. Donohue, L. T. Li, T. G. Van Niel, A. Thomas, J. Grieser, D. Jhajharia, Y. Himri, N. M. Mahowald, A. V. Mescherskaya, A. C. Kruger, S. Rehman, and Y. Dinpashoh (2012), Global review and synthesis of trends in observed terrestrial near-surface wind speeds: Implications for evaporation, *Journal of Hydrology*, *416–417*(0), 182-205.

Milly, P. C. D., and K. A. Dunne (2001), Trends in evaporation and surface cooling in the Mississippi River Basin, *Geophysical Research Letters*, *28*(7), 1219-1222.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008), Stationarity Is Dead: Whither Water Management?, *Science*, *319*(5863), 573-574.

Mitchell, K. E., D. Lohmann, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, B. A. Cosgrove, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, J. D. Tarpley, D. P. Lettenmaier, C. H. Marshall, J. K. Entin, M. Pan, W. Shi, V. Koren, J. Meng, B. H. Ramsay, and A. A. Bailey (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *Journal of Geophysical Research: Atmospheres*, *109*(D7), D07S90.

Montanari, A., and G. Di Baldassarre (2013), Data errors and hydrological modelling: The role of model structure to propagate observation uncertainty, *Advances in Water Resources*, *51*, 498-504.

Montanari, A., G. Young, H. H. G. Savenije, D. Hughes, T. Wagener, L. L. Ren, D. Koutsoyiannis, C. Cudennec, E. Toth, S. Grimaldi, G. Blöschl, M. Sivapalan, K. Beven, H. Gupta, M. Hipsey, B. Schaefli, B. Arheimer, E. Boegh, S. J. Schymanski, G. Di Baldassarre, B. Yu, P. Hubert, Y. Huang, A. Schumann, D. A. Post, V. Srinivasan, C. Harman, S. Thompson, M. Rogger, A. Viglione, H. McMillan, G. Characklis, Z. Pang, and V. Belyaev (2013), "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013–2022, *Hydrological Sciences Journal*, *58*(6), 1256-1275.

Mu, Q., M. Zhao, and S. W. Running (2011), Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sensing of Environment*, *115*(8), 1781-1800.

Mu, Q., M. Zhao, J. S. Kimball, N. G. McDowell, and S. W. Running (2013), A Remotely Sensed Global Terrestrial Drought Severity Index, *Bulletin of the American Meteorological Society*, *94*(1), 83-98.

Mu, Q., M. Zhao, F. A. Heinsch, M. Liu, H. Tian, and S. W. Running (2007), Evaluating water stress controls on primary production in biogeochemical and remote sensing based models, *Journal of Geophysical Research: Biogeosciences*, *112*(G1), n/a-n/a.

Mueller, B., S. I. Seneviratne, C. Jimenez, T. Corti, M. Hirschi, G. Balsamo, P. Ciais, P. Dirmeyer, J. B. Fisher, Z. Guo, M. Jung, F. Maignan, M. F. McCabe, R. Reichle, M. Reichstein, M. Rodell, J. Sheffield, A. J. Teuling, K. Wang, E. F. Wood, and Y. Zhang (2011), Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations, *Geophysical Research Letters*, *38*(6), L06402.

Munier, S., A. Polebistki, C. Brown, G. Belaud, and D. P. Lettenmaier (2015), SWOT data assimilation for operational reservoir management on the upper Niger River Basin, *Water Resources Research*, *51*(1), 554-575.

Mutiibwa, D., and S. Irmak (2013), AVHRR-NDVI-based crop coefficients for analyzing long-term trends in evapotranspiration in relation to changing climate in the U.S. High Plains, *Water Resources Research*, *49*(1), 231-244.

Nepstad, D. C., C. R. de Carvalho, E. A. Davidson, P. H. Jipp, P. A. Lefebvre, G. H. Negreiros, E. D. da Silva, T. A. Stone, S. E. Trumbore, and S. Vieira (1994), The role of deep roots in the hydrological and carbon cycles of Amazonian forests and pastures, *Nature*, *372*(6507), 666-669.

Ng, T. L., J. W. Eheart, X. Cai, and J. B. Braden (2011), An agent-based model of farmer decision-making and water quality impacts at the watershed scale under markets for carbon allowances and a second-generation biofuel crop, *Water Resources Research*, *47*(9), n/a-n/a.

Niemann, J. D., and E. A. B. Eltahir (2005), Sensitivity of regional hydrology to climate changes, with application to the Illinois River basin, *Water Resources Research*, *41*(7), W07014.

Ning, T., Z. Li, and W. Liu (2017), Vegetation dynamics and climate seasonality jointly control the interannual catchment water balance in the Loess Plateau under the Budyko framework, *Hydrol. Earth Syst. Sci.*, *21*(3), 1515-1526.

Niu, G.-Y., Z.-L. Yang, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, A. Kumar, K. Manning, D. Niyogi, E. Rosero, M. Tewari, and Y. Xia (2011), The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, *Journal of Geophysical Research: Atmospheres*, *116*(D12), n/a-n/a.

Noël, P. H., and X. Cai (2017), On the role of individuals in models of coupled human and natural systems: Lessons from a case study in the Republican River Basin, *Environmental Modelling & Software*, *92*, 1-16.

Orth, R., E. Dutra, and F. Pappenberger (2016), Improving weather predictability by including land-surface model parameter uncertainty, *Monthly Weather Review*.

Ozdogan, M., M. Rodell, H. K. Beaudoing, and D. L. Toll (2010), Simulating the Effects of Irrigation over the United States in a Land Surface Model Based on Satellite-Derived Agricultural Data, *Journal of Hydrometeorology*, *11*(1), 171-184.

Pan, M., and E. F. Wood (2006), Data Assimilation for Estimating the Terrestrial Water Budget Using a Constrained Ensemble Kalman Filter, *Journal of Hydrometeorology*, *7*(3), 534-547.

Pan, M., A. K. Sahoo, T. J. Troy, R. K. Vinukollu, J. Sheffield, and E. F. Wood (2011), Multisource Estimation of Long-Term Terrestrial Water Budget for Major Global River Basins, *Journal of Climate*, *25*(9), 3191-3206.

Pan, M., A. K. Sahoo, T. J. Troy, R. K. Vinukollu, J. Sheffield, and E. F. Wood (2012), Multisource Estimation of Long-Term Terrestrial Water Budget for Major Global River Basins, *Journal of Climate, 25*(9), 3191-3206.

Pan, M., C. K. Fisher, N. W. Chaney, W. Zhan, W. T. Crow, F. Aires, D. Entekhabi, and E. F. Wood (2015), Triple collocation: Beyond three estimates and separation of structural/non-structural errors, *Remote Sensing of Environment*, *171*, 299-310.

Popper, K. (2005), *The logic of scientific discovery*, Routledge.

Porporato, A., X. Feng, S. Manzoni, Y. Mau, A. J. Parolari, and G. Vico (2015), Ecohydrological modeling in agroecosystems: Examples and challenges, *Water Resources Research*, *51*(7), 5081-5099.

Potter, N. J., and L. Zhang (2007), Water balance variability at the interstorm timescale, *Water Resources Research*, *43*(5), W05405.

Potter, N. J., and L. Zhang (2009), Interannual variability of catchment water balance in Australia, *Journal of Hydrology*, *369*(1–2), 120-129.

Potter, N. J., and L. Zhang (2009), Interannual variability of catchment water balance in Australia, *Journal of Hydrology*, *369*(1-2), 120-129.

Potter, N. J., L. Zhang, P. C. D. Milly, T. A. McMahon, and A. J. Jakeman (2005), Effects of rainfall seasonality and soil moisture capacity on mean annual water balance for Australian catchments, *Water Resources Research*, *41*(6), W06007.

Rabalais, N. N., R. E. Turner, and W. J. Wiseman (2001), Hypoxia in the Gulf of Mexico, *Journal of Environmental Quality*, *30*(2), 320-329.

Rana, G., and N. Katerji (2000), Measurement and estimation of actual evapotranspiration in the field under Mediterranean climate: a review, *European Journal of Agronomy*, *13*(2–3), 125-153.

Raymond, P. A., N.-H. Oh, R. E. Turner, and W. Broussard (2008), Anthropogenically enhanced fluxes of water and carbon from the Mississippi River, *Nature*, *451*(7177), 449-452.

Renner, M., R. Seppelt, and C. Bernhofer (2012), Evaluation of water-energy balance frameworks to predict the sensitivity of streamflow to climate change, *Hydrol. Earth Syst. Sci.*, *16*(5), 1419-1433.

Riegger, J., and M. J. Tourian (2014), Characterization of runoff-storage relationships by satellite gravimetry and remote sensing, *Water Resources Research*, *50*(4), 3444-3466.

Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, *460*(7258), 999-1002.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C. J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin*, J. P. Walker, D. Lohmann, and D. Toll (2004), The Global Land Data Assimilation System, *Bulletin of the American Meteorological Society*, *85*(3), 381-394.

Rodell, M., H. K. Beaudoing, T. S. L'Ecuyer, W. S. Olson, J. S. Famiglietti, P. R. Houser, R. Adler, M. G. Bosilovich, C. A. Clayson, D. Chambers, E. Clark, E. J. Fetzer, X. Gao, G. Gu, K. Hilburn, G. J. Huffman, D. P. Lettenmaier, W. T. Liu, F. R. Robertson, C. A. Schlosser, J. Sheffield, and E. F. Wood (2015), The Observed State of the Water Cycle in the Early Twenty-First Century, *Journal of Climate*, *28*(21), 8289-8318.

Roderick, M. L., and G. D. Farquhar (2011), A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties, *Water Resources Research*, *47*(12), W00G07.

Rossman, N. R., and V. A. Zlotnik (2013), Review: Regional groundwater flow modeling in heavily irrigated basins of selected states in the western United States, *Hydrogeology Journal*, *21*(6), 1173-1192.

Ryu, Y., D. D. Baldocchi, S. Ma, and T. Hehn (2008), Interannual variability of evapotranspiration and energy exchange over an annual grassland in California, *Journal of Geophysical Research: Atmospheres*, *113*(D9), D09104.

Sankarasubramanian, A., and R. M. Vogel (2002), Annual hydroclimatology of the United States, *Water Resources Research*, *38*(6), 19-11-19-12.

Sankarasubramanian, A., and R. M. Vogel (2003), Hydroclimatology of the continental United States, *Geophysical Research Letters*, *30*(7), 1363.

Sankarasubramanian, A., R. M. Vogel, and J. F. Limbrunner (2001), Climate elasticity of streamflow in the United States, *Water Resources Research*, *37*(6), 1771-1781.

Scanlon, B. R., L. Longuevergne, and D. Long (2012), Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA, *Water Resources Research*, *48*(4), W04520.

Scanlon, B. R., C. C. Faunt, L. Longuevergne, R. C. Reedy, W. M. Alley, V. L. McGuire, and P. B. McMahon (2012), Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley, *Proceedings of the National Academy of Sciences*, *109*(24), 9320-9325.

Schnier, S., and X. Cai (2014), Prediction of regional streamflow frequency using model tree ensembles, *Journal of Hydrology*, *517*, 298-309.

Schurle, B. (1996), The Impact of Size on Yield Variability and Crop Insurance Premiums, *Review of Agricultural Economics*, *18*(3), 415-422.

Schymanski, S. J., M. Sivapalan, M. L. Roderick, L. B. Hutley, and J. Beringer (2009), An optimality-based model of the dynamic feedbacks between natural vegetation and the water balance, *Water Resources Research*, *45*(1).

Seaber, P. R., F. P. Kapinos, and G. L. Knapp (1987), Hydrologic Unit Maps, *U.S. Geological Survey, Water Supply Paper 2294*.

Sheffield, J., and E. F. Wood (2007), Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle, *Journal of Geophysical Research: Atmospheres*, *112*(D17), D17115.

Sheffield, J., G. Goteti, and E. F. Wood (2006), Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling, *Journal of Climate*, *19*(13), 3088-3111.

Sheffield, J., C. R. Ferguson, T. J. Troy, E. F. Wood, and M. F. McCabe (2009), Closing the terrestrial water budget from satellite remote sensing, *Geophysical Research Letters*, *36*(7), L07403.

Shuttleworth, W., R. Gurney, A. Hsu, and J. Ormsby (1989), FIFE: the variation in energy partition at surface flux sites, *IAHS Publ*, *186*, 67-74.

Shuttleworth, W. J. (2007), Putting the "vap" into evaporation, *Hydrol. Earth Syst. Sci.*, *11*(1), 210-244.

Siebert, S., J. Burke, J. M. Faures, K. Frenken, J. Hoogeveen, P. Döll, and F. T. Portmann (2010), Groundwater use for irrigation – a global inventory, *Hydrol. Earth Syst. Sci. Discuss.*, *7*(3), 3977-4021.

Sivapalan, M., H. H. G. Savenije, and G. Blöschl (2012), Socio-hydrology: A new science of people and water, *Hydrological Processes*, *26*(8), 1270-1276.

Sivapalan, M., M. A. Yaeger, C. J. Harman, X. Xu, and P. A. Troch (2011), Functional model of water balance variability at the catchment scale: 1. Evidence of hydrologic similarity and space-time symmetry, *Water Resources Research*, *47*(2), W02522.

Smith, L. M., D. A. Haukos, S. T. McMurry, T. LaGrange, and D. Willis (2011), Ecosystem services provided by playas in the High Plains: potential influences of USDA conservation programs, *Ecological Applications*, *21*(sp1), S82-S92.

Stoffelen, A. (1998), Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, *Journal of Geophysical Research: Oceans*, *103*(C4), 7755-7766.

Strassberg, G., B. R. Scanlon, and D. Chambers (2009), Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United States, *Water Resources Research*, *45*(5), n/a-n/a.

Swenson, S. C., and D. M. Lawrence (2015), A GRACE-based assessment of interannual groundwater dynamics in the Community Land Model, *Water Resources Research*, n/a-n/a.

Syed, T. H., V. Lakshmi, E. Paleologos, D. Lohmann, K. Mitchell, and J. S. Famiglietti (2004), Analysis of process controls in land surface hydrological cycle over the continental United States, *Journal of Geophysical Research: Atmospheres*, *109*(D22), D22105.

Szilagyi, J. (2001), Identifying Cause of Declining Flows in the Republican River, *Journal of Water Resources Planning and Management*, *127*(4), 244-253.

Tang, Q., E. Rosenberg, and D. Lettenmaier (2009), Use of satellite data to assess the impacts of irrigation withdrawals on Upper Klamath Lake, Oregon, *Hydrology and Earth System Sciences*, *13*(5), 617-627.

Tang, Q., S. Peterson, R. H. Cuenca, Y. Hagimoto, and D. P. Lettenmaier (2009), Satellite-based near-real-time estimation of irrigated crop water consumption, *Journal of Geophysical Research: Atmospheres*, *114*(D5), D05114.

Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins (2004), GRACE Measurements of Mass Variability in the Earth System, *Science*, *305*(5683), 503-505.

Taylor, L. R. (1961), Aggregation, Variance and the Mean, *Nature*, *189*(4766), 732-735.

Taylor, S. L., M. E. Payton, and W. R. Raun (1999), Relationship between mean yield, coefficient of variation, mean square error, and plot size in wheat field experiments, *Communications in Soil Science and Plant Analysis*, *30*(9-10), 1439-1447.

Tiedeman, C. R., and C. T. Green (2013), Effect of correlated observation error on parameters, predictions, and uncertainty, *Water Resources Research*, *49*(10), 6339-6355.

Ukkola, A. M., and I. C. Prentice (2013), A worldwide analysis of trends in water-balance evapotranspiration, *Hydrol. Earth Syst. Sci.*, *17*(10), 4177-4187.

Vico, G., and A. Porporato (2013), Probabilistic description of crop development and irrigation water requirements with stochastic rainfall, *Water Resources Research*, *49*(3), 1466-1482.

Vogel, R. (2011), Hydromorphology, *Journal of Water Resources Planning and Management*, *137*(2), 147-149.

Vogel, R. M., A. Rosner, and P. H. Kirshen (2013), Brief Communication: Likelihood of societal preparedness for global change: trend detection, *Nat. Hazards Earth Syst. Sci.*, *13*(7), 1773-1778.

Vogel, R. M., U. Lall, X. Cai, B. Rajagopalan, P. K. Weiskel, R. P. Hooper, and N. C. Matalas (2015), Hydrology: The interdisciplinary science of water, *Water Resources Research*, n/a-n/a.

Wada, Y., and M. F. P. Bierkens (2014), Sustainability of global water use: past reconstruction and future projections, *Environmental Research Letters*, *9*(10), 104003.

Wada, Y., L. P. H. van Beek, C. M. van Kempen, J. W. T. M. Reckman, S. Vasak, and M. F. P. Bierkens (2010), Global depletion of groundwater resources, *Geophysical Research Letters*, *37*(20), L20402.

Wang, D. (2012), Evaluating interannual water storage changes at watersheds in Illinois based on long-term soil moisture and groundwater level data, *Water Resources Research*, *48*(3), W03502.

Wang, D., and X. Cai (2009), Detecting human interferences to low flows through base flow recession analysis, *Water resources research*, *45*(7), W07426.

Wang, T., E. Istanbulluoglu, J. Lenters, and D. Scott (2009), On the role of groundwater and soil texture in the regional water balance: An investigation of the Nebraska Sand Hills, USA, *Water Resources Research*, *45*(10), W10413.

Weiskel, P. K., R. M. Vogel, P. A. Steeves, P. J. Zarriello, L. A. DeSimone, and K. G. Ries (2007), Water use regimes: Characterizing direct human interaction with hydrologic systems, *Water Resources Research*, *43*(4), W04402.

Weiskel, P. K., D. M. Wolock, P. J. Zarriello, R. M. Vogel, S. B. Levin, and R. M. Lent (2014), Hydroclimatic regimes: a distributed water-balance framework for hydrologic assessment and classification, *Hydrol. Earth Syst. Sci. Discuss.*, *11*(3), 2933-2965.

Williams, C. A., M. Reichstein, N. Buchmann, D. Baldocchi, C. Beer, C. Schwalm, G. Wohlfahrt, N. Hasler, C. Bernhofer, T. Foken, D. Papale, S. Schymanski, and K. Schaefer (2012), Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers, *Water Resources Research*, *48*(6), W06523.

Xia, Y., M. T. Hobbins, Q. Mu, and M. B. Ek (2015a), Evaluation of NLDAS-2 evapotranspiration against tower flux site observations, *Hydrological Processes*, *29*(7), 1757-1771.

Xia, Y., C. D. Peter-Lidard, M. Huang, H. Wei, and M. Ek (2015b), Improved NLDAS-2 Noah-simulated hydrometeorological products with an interim run, *Hydrological Processes*, *29*(5), 780-792.

Xia, Y., D. Mocko, M. Huang, B. Li, M. Rodell, K. E. Mitchell, X. Cai, and M. B. Ek (2017), Comparison and Assessment of Three Advanced Land Surface Models in Simulating Terrestrial Water Storage Components over the United States, *Journal of Hydrometeorology*, *18*(3), 625-649.

Xia, Y., B. A. Cosgrove, K. E. Mitchell, C. D. Peters-Lidard, M. B. Ek, M. Brewer, D. Mocko, S. V. Kumar, H. Wei, J. Meng, and L. Luo (2016), Basin-scale assessment of the land surface water budget in the National Centers for Environmental Prediction operational and research NLDAS-2 systems, *Journal of Geophysical Research: Atmospheres*, n/a-n/a.

Xia, Y., K. Mitchell, M. Ek, B. Cosgrove, J. Sheffield, L. Luo, C. Alonge, H. Wei, J. Meng, B. Livneh, Q. Duan, and D. Lohmann (2012a), Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow, *Journal of Geophysical Research: Atmospheres*, *117*(D3), D03110.

Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. Wood, L. Luo, C. Alonge, H. Wei, J. Meng, B. Livneh, D. Lettenmaier, V. Koren, Q. Duan, K. Mo, Y. Fan, and D. Mocko (2012b), Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *Journal of Geophysical Research: Atmospheres*, *117*(D3), D03109.

Xu, T., and A. J. Valocchi (2015), A Bayesian approach to improved calibration and prediction of groundwater models with structural error, *Water Resources Research*, *51*(11), 9290-9311.

Xu, T., A. J. Valocchi, J. Choi, and E. Amir (2014), Use of Machine Learning Methods to Reduce Predictive Error of Groundwater Models, *Groundwater*, *52*(3), 448-460.

Xu, T., A. J. Valocchi, M. Ye, and F. Liang (2017a), Quantifying model structural error: Efficient Bayesian calibration of a regional groundwater flow model using surrogates and a data-driven error model, *Water Resources Research*, n/a-n/a.

Xu, T., A. J. Valocchi, M. Ye, F. Liang, and Y.-F. Lin (2017b), Bayesian calibration of groundwater models with input data uncertainty, *Water Resources Research*, *53*(4), 3224-3245.

Xu, X., W. Liu, B. R. Scanlon, L. Zhang, and M. Pan (2013), Local and global factors controlling water-energy balances within the Budyko framework, *Geophysical Research Letters*, *40*(23), 2013GL058324.

Yang, D., F. Sun, Z. Liu, Z. Cong, G. Ni, and Z. Lei (2007), Analyzing spatial and temporal variability of annual water-energy balance in nonhumid regions of China using the Budyko hypothesis, *Water Resour. Res.*, *43*(4), W04426.

Yang, H., D. Yang, and Q. Hu (2014a), An error analysis of the Budyko hypothesis for assessing the contribution of climate change to runoff, *Water Resources Research*, *50*(12), 9620-9629.

Yang, H., D. Yang, Z. Lei, and F. Sun (2008), New analytical derivation of the mean annual water-energy balance equation, *Water Resources Research*, *44*(3), W03410.

Yang, H., J. Qi, X. Xu, D. Yang, and H. Lv (2014b), The regional variation in climate elasticity and climate contribution to runoff across China, *Journal of Hydrology*, *517*(0), 607-616.

Yang, Y.-C. E., X. Cai, and D. M. Stipanović (2009), A decentralized optimization algorithm for multiagent system–based watershed management, *Water Resources Research*, *45*(8), n/a-n/a.

Yang, Y., J. Fang, W. Ma, and W. Wang (2008), Relationship between variability in aboveground net primary production and precipitation in global grasslands, *Geophysical Research Letters*, *35*(23), L23710.

Yohe, G., N. Andronova, and M. Schlesinger (2004), To Hedge or Not Against an Uncertain Climate Future?, *Science*, *306*(5695), 416-417.

Zanardo, S., C. J. Harman, P. A. Troch, P. S. C. Rao, and M. Sivapalan (2012), Intra-annual rainfall variability control on interannual variability of catchment water balance: A stochastic analysis, *Water Resources Research*, *48*(6), W00J16.

Zeng, R., and X. Cai (2014), Analyzing streamflow changes: irrigation-enhanced interaction between aquifer and streamflow in the Republican River basin, *Hydrol. Earth Syst. Sci.*, *18*(2), 493-502.

Zeng, R., and X. Cai (2015), Assessing the temporal variance of evapotranspiration considering climate and catchment storage factors, *Advances in Water Resources*, *79*(0), 51-60.

Zeng, R., and X. Cai (2016), Climatic and terrestrial storage control on evapotranspiration temporal variability: Analysis of river basins around the world, *Geophysical Research Letters*, *43*(1), 185-195.
Zhang, K., J. S. Kimball, R. R. Nemani, and S. W. Running (2010), A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006, *Water Resources Research*, *46*(9), W09522.
Zhang, L., W. R. Dawes, and G. R. Walker (2001), Response of mean annual evapotranspiration to vegetation changes at catchment scale, *Water Resources Research*, *37*(3), 701-708.
Zhang, L., K. Hickel, W. R. Dawes, F. H. S. Chiew, A. W. Western, and P. R. Briggs (2004), A rational function approach for estimating mean annual evapotranspiration, *Water Resources Research*, *40*(2), W02502.

APPENDIX A

SENSITIVITY ANALYSIS WITH VARIOUS PET CALCULATION METHODS

We calculate *PET* in additional three methods besides the method used in the text in order to test the robustness of the ET variance result subject to *PET* uncertainties. The four *PET* calculation methods are:

1) The *PET* calculated from Penman equation denoted as "Penman" is calculated as [*Brutsaert*, 2005]:

$$PET = \frac{\Delta}{\Delta+\gamma}(R_n - G) + \frac{\gamma}{\Delta+\gamma}0.26(1 + 0.54\bar{u}_2)(e_s^* - e_a) \qquad \text{Eqn.(A1)}$$

where Δ is the slope of the saturation water vapor pressure curve at air temperature (Pa K^{-1}); γ is the psychrometric constant (taken as 0.67 hPa K^{-1}); R_n is the net radiation and G is ground heat flux (mm d^{-1}); e_s^* is the saturated vapor pressure (hPa); e_a is the actural water vapor pressure (hPa); $\bar{u}_2$ is the mean wind speed at 2m about the ground (m s^{-1}).

2) The Penman methods without ground energy flux, denoted as "Penman no ground flux", is calculated by setting G=0 in Eqn S1 to avoid the uncertainty caused by the VIC model.

3) *PET* calculated from Priesley-Taylor method, denoted as "Priesley-Taylor", follows:

$$PET_{PT} = \alpha\frac{\Delta}{\Delta+\gamma}(R_n - G) \qquad \text{Eqn.(A2)}$$

where $\alpha = 1.3$.

4) *PET* calculated from FAO Penman-Monteith method [*Allen et al.*, 1998], denoted as "FAO", is calculated as:

$$PET_{FAO} = \frac{0.408\Delta(R_n-G)+\gamma\frac{9000}{T+273}\bar{u}_2(e_s^*-e_a)}{\Delta+\gamma(1+0.34\bar{u}_2)} \qquad \text{Eqn.(A3)}$$

where T is the daily mean temperature (K).

In terms of mean annual *PET* as shown in Figure A1.1, the FAO method results in very close *PET* as the Penman method in most of the 32 basins. Priesley-Taylor method gives relatively low *PET* among the four methods, especially in arid basins such as Niger, Nile and Senegal.

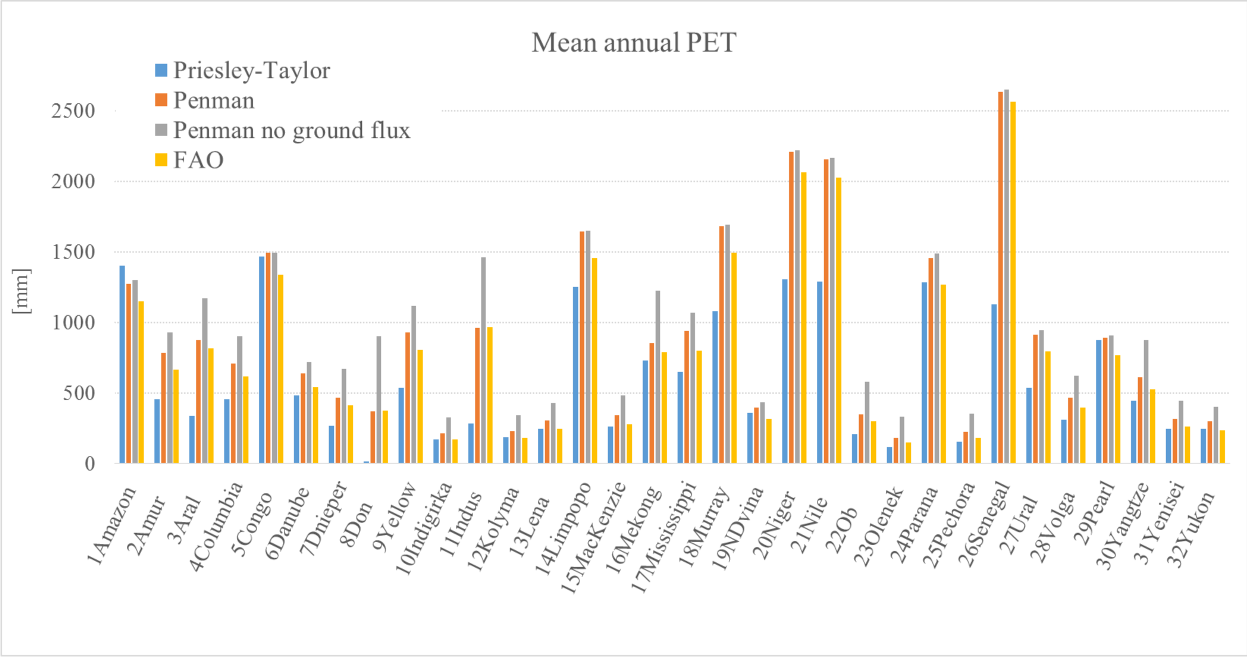


Figure A1.1. Mean annual *PET* calculated from the four methods

The inter-annual *PET* variances by these methods are quite consistent as shown in Figure A1.2, except that the Priesley-Taylor method yields smaller *PET* variance in some basins such as Dnieper, Don, Mississippi and Ural.

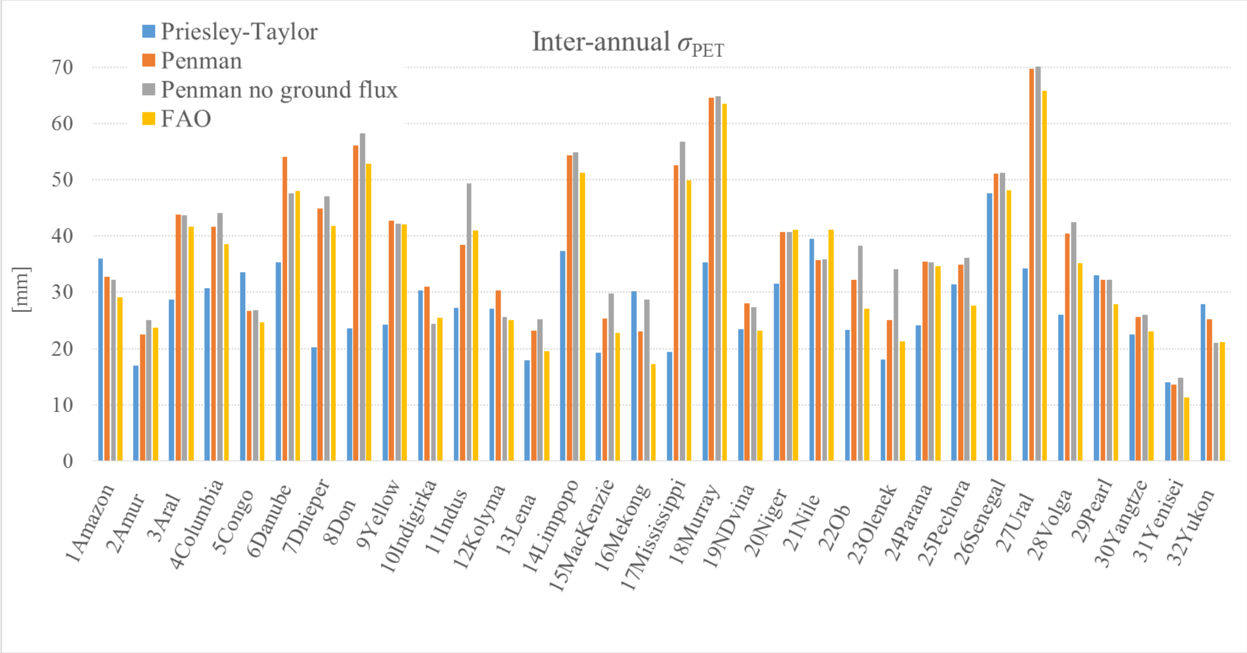


Figure A1.2. Inter-annual σ_{PET} calculated from the four methods

The intra-annual *PET* variance is shown in Figure A1.3. The Penman method without ground flux yields larger intra-annual *PET* variance in most of the basins since the buffer effect of ground energy is neglected with the calculation. The other three methods yield similar results.

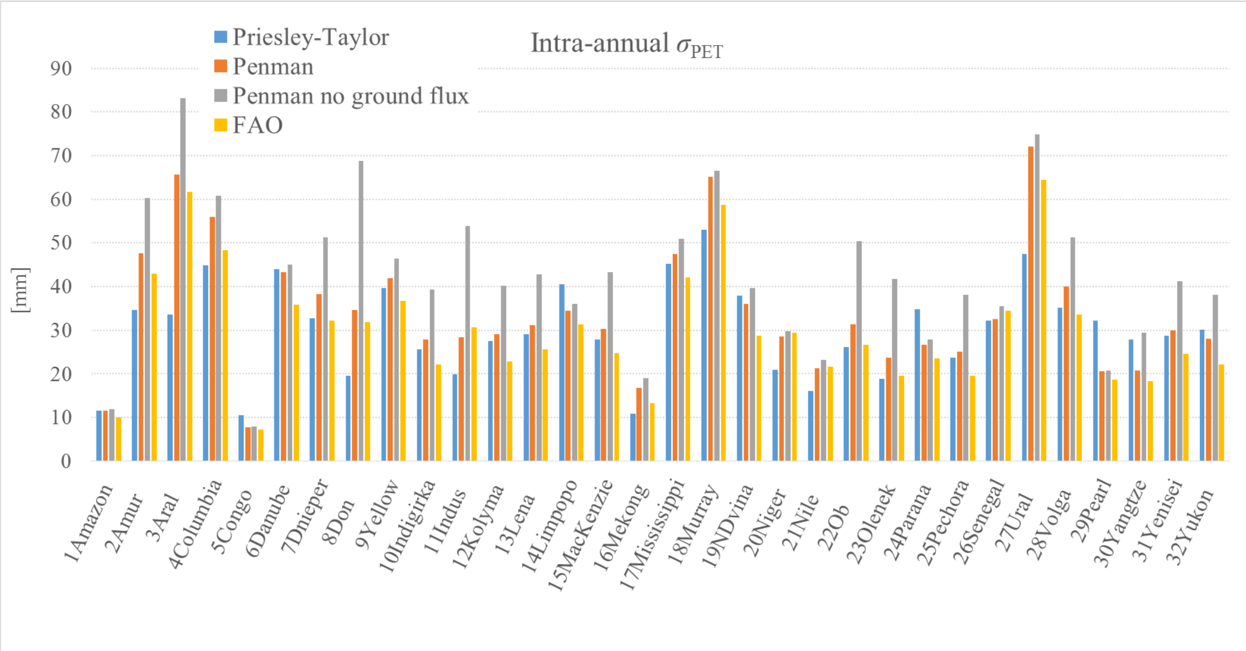


Figure A1.3. Intra-annual σ_{PET} calculated from the four methods

Both the mean value and variability of *PET* affect ET variance, and the inter-annual σ_{ET} with the four *PET* calculation methods are shown in Fiugre A1.4. For inter-annual σ_{ET}, the various *PET* calculation methods yield quite similar results. For example, the inter-annual σ_{ET} from the Penman methods is within 10% of the ensemble mean for all the basins except Dnieper and Don.

Specifically, for the three basins with *PET*-dominated inter-annual σ_{ET} (e.g., Amazon, North Divan and Pechora), the four *PET* calculation methods give very similar results (within 5% deviation among the four methods).

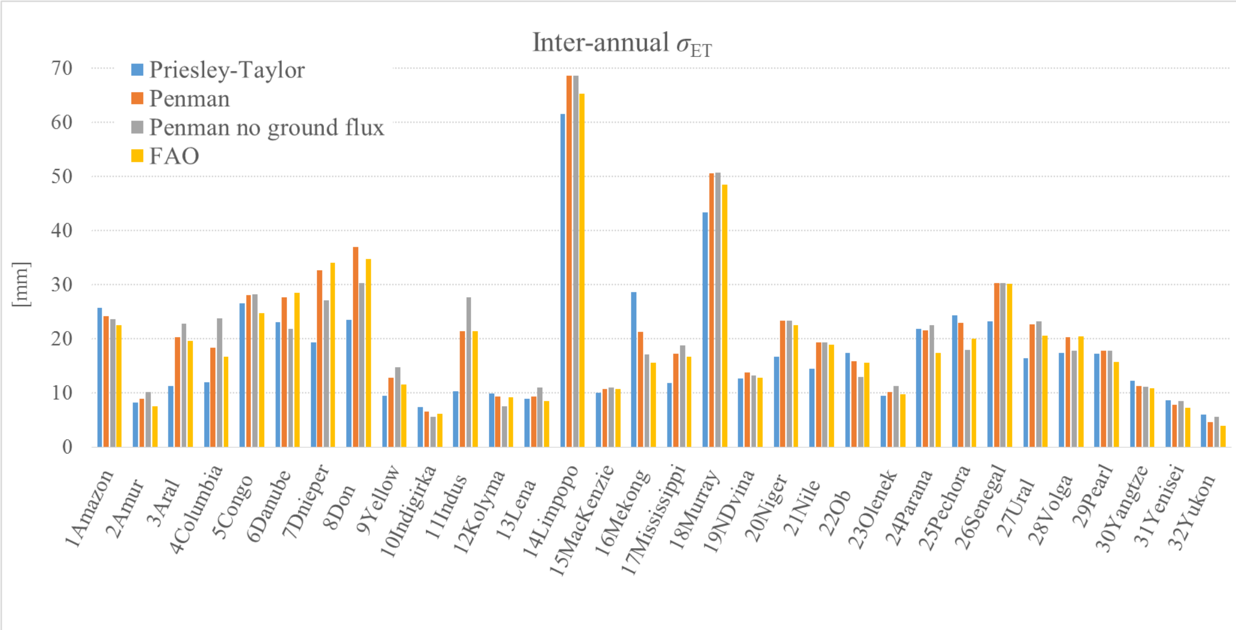


Figure A1.4. Inter-annual σ_{ET} calculated from the four methods for *PET*

The intra-annual σ_{ET} with the four *PET* calculation methods are shown in Figure A1.5. For intra-annual σ_{ET}, the Penman method without ground energy flux yield larger intra-annual σ_{ET} than other three methods in basins such as Ob, Mekong, MacKenzie and Yenisei. As discussed before, the over-estimation is due to the over-estimation of *PET* fluctuation by neglecting the buffer effect of ground energy flux. The over-estimation is more obvious at monthly scale than that at annual scale.

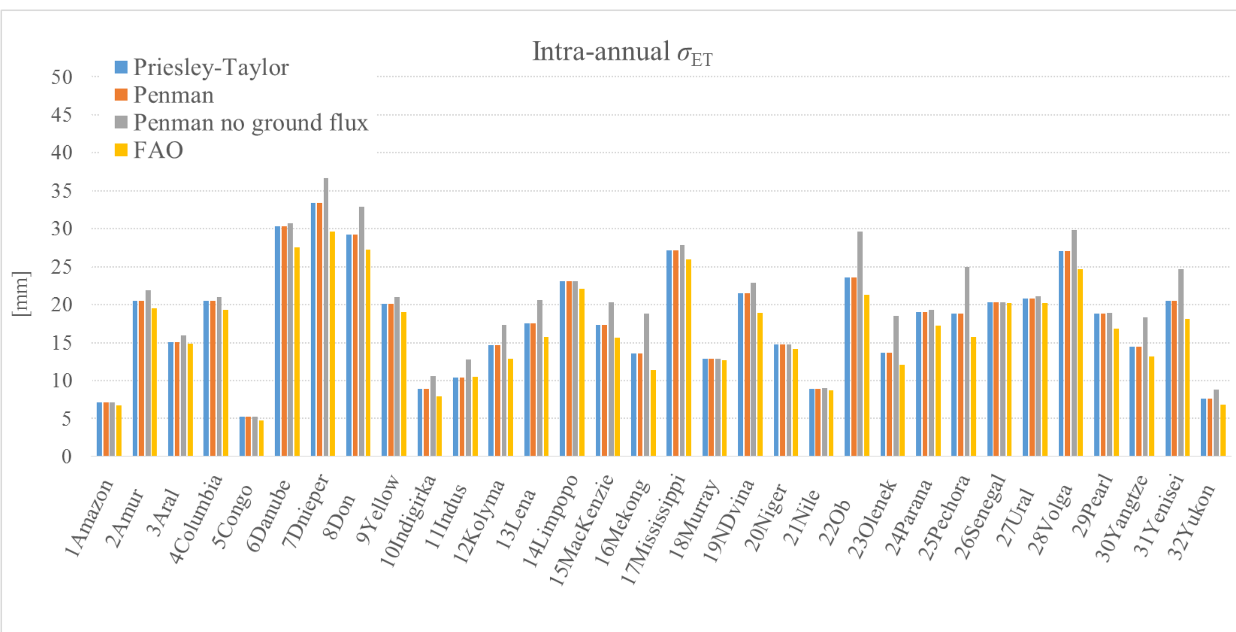


Figure A1.5. Inter-annual σ_{ET} calculated from the four methods for *PET*

In summary, the ET variance decomposition framework is robust to the four *PET* calculation methods discussed above. Although the long-term average *PET* varies by the calculation methods, the variance is consistence among these methods.

APPENDIX B

INTER- AND INTRA- ANNUAL RUNOFF VARIANCE

1. Inter-annual runoff variance

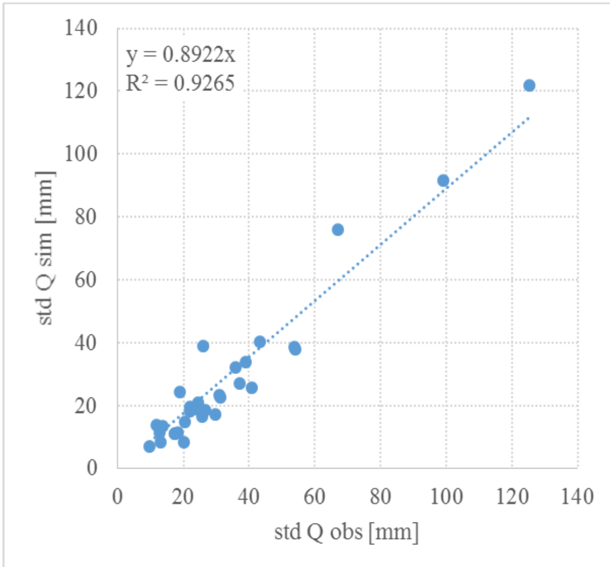


Figure A2.1. Observed vs. simulated standard deviation of annual runoff calculated in 32 basins.

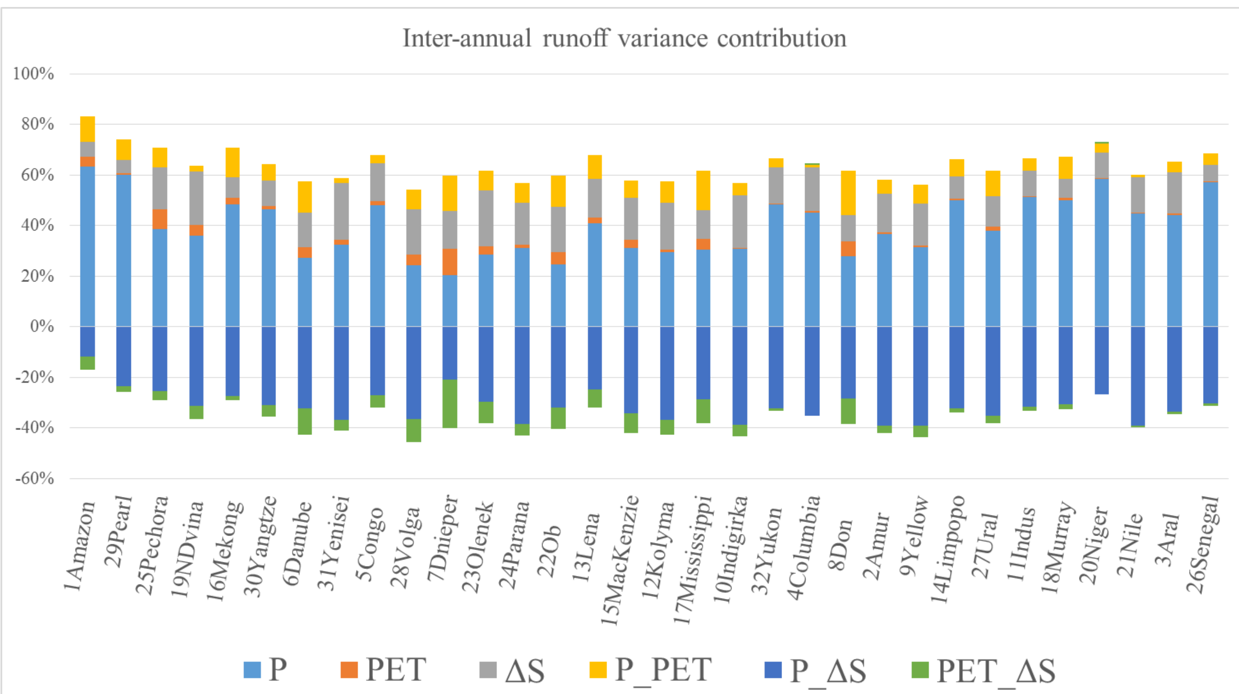


Figure A2.2. The proportional contributions from *P, PET* and *ΔS* variance and covariance to runoff variance. At annual scale, *P* variance and *P_ΔS* coupling control runoff variance.

2. Intra-annual runoff variance

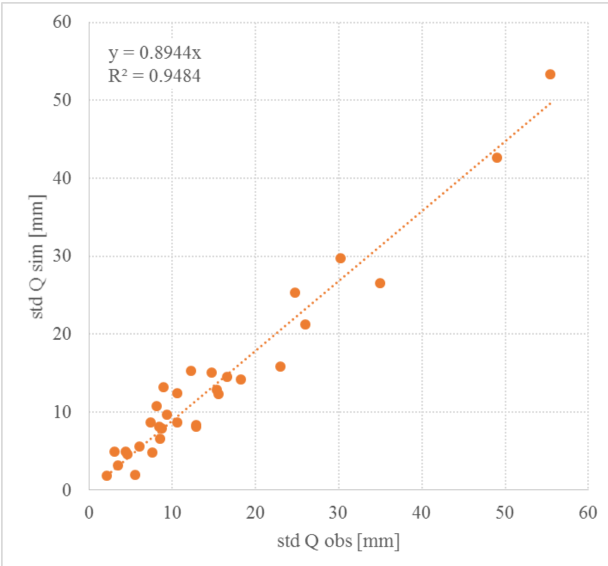


Figure A2.3. The observed vs. simulated standard deviation of monthly runoff calculated in 32 basins.

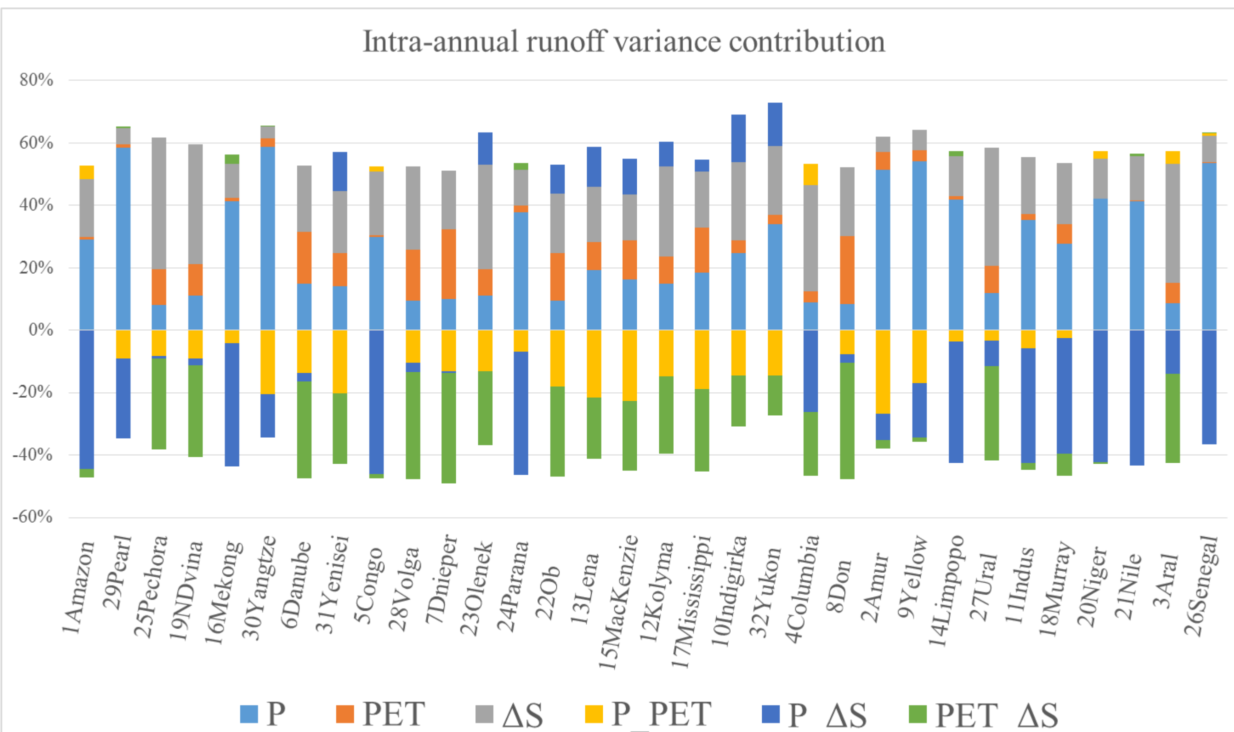


Figure A2.4. The proportional contributions from *P, PET* and *ΔS* variance and covariance to runoff variance.

At intra-annual scale, there is no uniform pattern of runoff variance like that at inter-annual scale. But there are some patterns we can visually get. The most important findings is the impact of *P&ΔS* and *PET&ΔS* coupling on runoff variance. In those basins with significant *PET&ΔS* signal (e.g., Pechora, North Dvina) the *P&ΔS* signal is negligible; in those basins with significant *P&ΔS* signal (e.g., Amazon, Congo and Senegal) the *PET&ΔS* signal is negligible. This shows how storage behaves differently under different climate. That is, *ΔS* is driven by energy in humid/cold basins (most of them are in Russia, with significant snow processes) and by atmospheric water supply in arid basins.