DISSERTATION

PART A: A PROBABILISTIC FRAMEWORK FOR ASSESSING VULNERABILITY TO CLIMATE VARIABILITY AND CHANGE: THE CASE OF THE US WATER SUPPLY SYSTEM. PART B: DYNAMICS OF SELF-ORGANIZED VEGETATION PATTERNS

Submitted by

Romano Foti

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2011

Doctoral Committee:

Advisor: Jorge A. Ramirez

Mazdak Arabi
Thomas C. Brown
Stephanie Kampf

ABSTRACT

PART A: A PROBABILISTIC FRAMEWORK FOR ASSESSING VULNERABILITY TO CLIMATE VARIABILITY AND CHANGE: THE CASE OF THE US WATER SUPPLY SYSTEM. PART B: DYNAMICS OF SELF-ORGANIZED VEGETATION PATTERNS

Part A: Water use for human needs has increased dramatically all over the world, in response to tremendous population and economic growth. In the United States alone, water consumption increased over ten-fold during the twentieth century. While the increase in water use efficiency mitigated the pressure put on water resources, rising population and wealth together with expected climatic changes are aggravating hydrologic uncertainty. Taken together, these forces are making careful water management ever more important. Thus, a realistic broad-scale understanding of the vulnerability of water supply systems to shortage must be a component of any attempt to define the magnitude of the threat, and is essential in determining appropriate mitigation and adaptation measures.

In general, the environmental and socio-economic vulnerability of a system depends not only on its ability to withstand stresses and on the magnitude of those stresses but also on the inherent variability of their hydro-climatic and socio-economic drivers (*e.g.*, water supply capacity, water demand, precipitation, evapotranspiration, population, technology, water management infrastructure, etc.). Thus, in order to quantify current and future vulnerability one must do so probabilistically. Accordingly, we define vulnerability as the probability that, at a given place and time, the demand exceeds the supply. Water supply is evaluated as fresh water yield as altered by storage, trans-basin diversions, and other management impacts and is represented by a time dependent probability distribution function (PDF). Fresh water yield is estimated on an annual

basis as precipitation minus actual evapotranspiration by means of a statistical dynamical water balance model. The PDF of available water supply, when compared with the PDF of the desired water use, yields an estimate of the probability of shortage, and thus a measure of the vulnerability of the water supply system.

The vulnerability of the US water supply system is assessed for 98 basins covering the contiguous United States and for current and future conditions, in light of the socioeconomic and climatic changes projected by nine combinations of GCM models and IPCC-SRES scenarios for the 21st century.

Results are presented in general terms for the entire US while a deeper analysis is performed for a set of selected basins of the Colorado River Basin, California and the High Plains. The latter set of basins is ranked based on their response to current and future changes in the PDFs of supply and demand as projected by the set of GCM/scenario combination selected.

Our findings show that the Southwest and central and southern Great Plains are the more vulnerable areas to future climatic and socio-economic changes. In addition, future increases in the vulnerability of the US water supply will depend more on changes in water yield than on growth in water demand.

Among the selected basins, the Central California and the San Joaquin-Tulare are found to be the more sensitive to both current and future variability of demand and supply. Large sensitivity is also found for the entire area of the High Plains analyzed, while in the Colorado River Basin, the Lower Colorado is the sub-basin with the larger susceptibility to changes in future supply and demand. Future GCM/scenarios showed general disagreement both in terms of vulnerability and sensitivity of individual basins.

On the whole, the procedure outlined in this framework offers a versatile and consistent instrument to assess the vulnerability of physical systems to changes in inherently variable stressors and can be applied to any environmental and socio-economic vulnerability analysis. In

addition, it is the only methodology that both accounts the probabilistic character of the drivers and allows for explicit inclusion of thresholds.

Part B: Vegetation patterns are a common and well-defined characteristic of many arid and semi-arid landscapes. In this study we explore some of the physical mechanisms responsible for the establishment of self-organized, non-random vegetation patterns that arise at the hillslope scale in many areas of the world, especially in arid and semi-arid regions. In doing so we use a water and energy balance model and provide a fundamental mechanistic understanding of the dynamics of vegetation pattern formation and development. Within the modeling, reciprocal effects of vegetation on the hillslope energy balance, runoff production and run-on infiltration, root density, surface albedo and soil moisture content are analyzed.

In particular, we: 1) present a physically-based mechanistic description of the processes leading to vegetation pattern formation; 2) compare simulated vegetation coverage at the hillslope scale with observations; 3) quantify the relative impact of pattern-inducing dynamics on pattern formation; and 4) describe the relationships between vegetation patterns and the climatic, hydraulic and topographic characteristic of the system.

The model is validated by comparing hillslope-scale simulations with available observations for the areas of Niger near Niamey and Somalia near Garoowe, where respectively tiger bushes and banded vegetation patterns are present. The model validation includes comparison of simulated and observed vegetation coverage as well as simulated and measured water fluxes, showing both qualitative and quantitative agreement between simulations and observations.

The analysis of the system suggests that the main driver of pattern establishment is climate, in terms of average annual precipitation and incoming solar radiation. In particular, decreasing precipitation or, conversely, increasing incoming radiation are responsible for the system departure from a fully vegetated state with indistinguishable vegetation structures to a sparsely vegetated state with (self-organized) distinguishable patterns. In addition, within the range of climatic conditions that promote the formation of self-organized vegetation patterns, the

phenomenon is found as mainly driven by surface runoff production and run-on infiltration. On the other hand, the spatial interactions between adjacent vegetation groups and the effect of roots and surface reflectance on soil moisture redistribution have a determinant role on both the characteristics and stability of the pattern and on the total biomass that is established on the domain.

TABLE OF CONTENTS

Chapter 1 Introduction

1.1 General motivation

This dissertation consists of two parts: in Part A we propose a probabilistic approach to assess the sensitivity and the vulnerability of water supply systems to climate change, explicitly addressing the case of the US; in Part B we develop a mechanistic, physically based model to describe the dynamics of the self-organized vegetation patterns typically emerging at the hillslope scale in arid and semi-arid environments.

The document is organized as follows: in this chapter, the general motivation, objectives and the main conclusions of each part are presented. Part A of this dissertation is presented in detail in Chapter 2, while Chapter 3 deals with Part B. Detailed methodology and extended results are presented in the appendices.

1.2 Part A: General motivation

Water use for human needs has increased dramatically all over the word, in response to the tremendous population and economic growth. In the United States alone, water consumption has increased over ten-fold during the twentieth century (Brown, 2000). The increase in water use efficiency has only mitigated the enormous pressure put on water resources by the rising population and the rising incomes. Complicating the picture, climatic change is increasing hydrologic uncertainty. Taken together, these forces are making careful water management ever more important. Thus, a realistic broad-scale understanding of the vulnerability of the US water supply system to shortage must be a component of any attempt to define the magnitude of the threat, and is essential in determining appropriate mitigation and adaptation measures.

This large scale impact assessment is devoted to evaluating the potential impact of climatic and socio-economic changes on the water supply system of the United States. A probabilistic framework is here developed in order to estimate the probability that, at a given point in space and time, the water demand for human activities exceeds the water supply.

Water demands for the conterminous US are calculated at the Assessment Subregions (ASRs) level on an annual basis from 1985 to 2100 taking into consideration historical records of water withdrawals and future projections of climatic and socio economic conditions aided by the analysis of nine combinations of General Circulation Models (CGMs) and IPCC-SRES scenarios.

Water supply for the conterminous US is estimated as fresh water yield as modified by the effect of storage and surface water redistribution via natural streamflow and man-made water diversions. Water yield is calculated as precipitation minus actual evapotranspiration on a 5x5 km study grid covering the whole conterminous US by means of the Eagleson's model (Eagleson, 1978a-g), a statistical-dynamical representation of the annual water balance. Water yield is estimated from 1953 to 2100 making use of historical climatic records and model projections based on the above nine combinations of GCM and IPCC-SRES scenarios.

The US water supply network existing at the ASR spatial level is then identified and simulated under historical, current and projected future climatic and socio-economic scenarios in order to characterize the vulnerability of the water supply system.

The vulnerability assessments are made for each of the nine GCM/scenario combinations and for four target periods (2020, 2040, 2060 and 2080) each one based on 20-year simulations centered at the target year. In addition to its scientific implications, this large scale impact analysis can assist and benefit local communities, as well as stakeholders and decision-makers.

Results of the vulnerability analysis are provided in Chapter 2. The methodology is described in Appendix A through Appendix D.

1.1.1 Objectives

Within the overall development of the large-scale evaluation of vulnerability of the US water supply system to shortage, a number of specific objectives were pursued in this dissertation. They are discussed below.

1.1.1.1 Evaluation of long-term storm statistics at the 5x5 km scale grid for the conterminous US

Stochastic representation of precipitation regimes is often needed for the analysis of water resources systems and for watershed modeling. Manipulation of the probability density functions (PDFs) of precipitation events is often necessary in order to simulate rainfall-runoff behavior of catchments (Grayman and Eagleson 1971) or derivations of flood probability distributions (Chan and Bras 1979).

In this framework, long-term statistical properties of storm events are needed to apply Eagleson's model for the estimation of water yield for the conterminous US. In particular, Eagleson's model requires the following long-term storm statistics to be evaluated at the scale at which the model needs to be applied: mean storm duration, mean time between storms, mean interarrival time, mean storm intensity, mean storm depth, mean number of storms, mean annual precipitation, mean rainy season duration and parameter of the gamma distribution of storm depth.

Those parameters were evaluated using the NCDC hourly records for over 2000 rainfall stations across the country. Estimation of the storm statistic was performed both at the annual and at the monthly time scale. See Appendix A for details.

1.1.1.2 Application of a statistical-dynamical model for the estimation of current and future US water yield at the 5x5 km scale

Water supply in a given point in space is here quantified as water yield plus inflow from upstream subjected to the effect of management via reservoir storage and water diversions. Thus, estimation of water yield is the first step towards the characterization of water supply at any given location. Eagleson's one-dimensional, statistical-dynamical water balance model (1978; 1978; 1978; 1978; 1978; 1978; 1978) is here used to estimate yield as the difference between precipitation and actual evapotranspiration. On an annual basis, assuming there is no change in water stored in the surface or in the aquifers, total water yield also represents the sum of surface and groundwater runoff.

The water balance model is calibrated by minimizing the mean squared error between modeled water yield and historical streamflow records.

Three different historical streamflow datasets are used:

- 42-year series of annual streamflow records for 655 relatively undisturbed test basins across the conterminous U.S., provided by the U.S. Forest Service.
- Reconstructed virgin streamflow for a set of watersheds in the Colorado River Basin, provided by the Bureau of Reclamation.
- 30-year average reconstructed virgin streamflow estimated for the contiguous U.S. by the U.S. Geological Survey at the 8-digit basin level.

Calibration is performed at the basin level, either a test basin or 8-digit basin, or an 8-digit basin in combination with USBR streamflow records. In order to perform the calibration, all soil and climatic parameters needed were averaged at the basin level from the original datasets.

Once calibrated, the model is run in order to determine:

- annual water yield for the entire US for the historical period 1953-2005;
- annual projections of water yield for the entire US for the future years 2006-2100.

Detailed results of the calibration procedure and the historical yield estimation are provided in Appendix A.

1.1.1.3 Analysis of future climatic and socio-economic scenarios

The increasing globalization of the world economy and the possibility of substantial climatic change have created considerable uncertainty about future US water demand and supply, even for the near future. A way to capture this uncertainty is to examine a certain number of potential future storylines, characterized by different climatic and economic pathways. Climate and society, however, cannot be considered independent. It is clear that human activities have repercussions on the environment and climate and, on the other hand, climatic conditions have a strong impact on society. GCMs offer projections of future water and energy fluxes at large scale and can incorporate the effect of greenhouse gas emissions due to human activities. On the other hand, emission pathways are determined by the structure of the future society.

For this study, three scenarios (A1B, A2, and B2), each based on a different storyline, were chosen from the IPCC SRES set (Nakicenovic, Alcamo et al. 2000). Of the three scenarios used here, the A2 scenario is the most extreme and B2 the least extreme in terms of atmospheric CO_2 concentration, although the projections for CO_2 of these scenarios do not differ greatly until later in the 21st century.

Scenarios are here used in combination with the following GCMs: the Canadian Centre for Climate Modelling and Analysis Third Generation Coupled Global Climate Model Version 3.1 Medium Resolution (CGCM31MR) and the Australian Commonwealth Scientific and Industrial Research Organisation Mark 3.5 Climate System Model (CSIROMK35).

As mentioned above, each climatic projection must be used jointly with its corresponding socio-economic scenario in order to make a consistent assessment on future climate, water supply, water demand and, thus, future vulnerability.

Downscaling of the climatic GCM output was performed by others using the ANUSPLIN software (Price, McKenney et al. 2006) to a 10-km grid for the coterminous US while a simple inverse distance squared interpolation was executed as part of this study in order to match the spatial resolution of this study. Results are corrected for bias by matching averages of observations and projections for the period 2001-2008.

Analysis of future climatic projections and future water yield for the nine selected GCM/scenario combinations is supplied in Appendix C.

1.1.1.4 Characterization of the US water supply system

Water is a mobile source. It can be used *in loco* to meet the demands, moved to satisfy needs that are further away from the place where it is available or stored when it is in excess in order to be used in future circumstances. Ultimately, it simply flows from upstream to downstream. It follows that an estimation of the water available at a certain time and location cannot be properly made by simply calculating the difference between precipitation and actual evapotranspiration at that point. Hence, a thorough characterization of water supply for a large area necessitates an accurate understanding of the complex water network existing within the area itself. A substantial part of this study is devoted to the identification and the simulation of such network. The US water supply network is here characterized at the ASR spatial level and all simulations are carried out at the annual time step.

The water networks are identified by examining the water connections existing between ASRs via either natural routes or artificial diversions. Three main water networks plus a set of 15 isolated ASRs were identified. Each network is represented as a set of nodes interconnected by links. Nodes correspond to water inflows, water demands and storages, while links are used to represent natural or artificial water routes.

Once the physical structure of the US water supply network was determined, a set of operating rules was established in order to mimic water management actions and allow network simulation.

Results of the simulations under all of the nine GCM/scenario combinations enabled the probabilistic characterization of the vulnerability of the US water supply system to shortage. Detailed results are presented in Chapter 2 and Appendix B.

1.1.1.5 Characterization of the future vulnerability of the US water supply system to shortage

Vulnerability of a system is a function of its ability to respond (i.e., cope with; adapt) to inherently variable stressors. In this study, given the uncertainty characterizing both the stressors and the capacity to withstand them, the necessity of quantifying vulnerability under a probabilistic framework is advocated. In particular, vulnerability is estimated as the probability that a critical system threshold, itself a function of both the capacity and the stressors of the system, will be crossed. In the context of the US water supply system, this definition translates into evaluating the probability of shortage.

We present a probabilistic, physically-based, spatial and temporal characterization of hydrologic fluxes (e.g., precipitation, evapotranspiration, surface and subsurface runoff), soil moisture storages, water demands, reservoir storage and water transfers.

By making use of the analyses described in the above sections, the vulnerability of water supply to shortage is estimated for each of the 98 ASRs of the conterminous US.

Vulnerability analysis is carried out for the entire 21st century in light of the socioeconomic and climatic changes projected by the nine combinations of GCM models and IPCC-SRES scenarios. Our probabilistic approach allows determination of the more vulnerable areas of the US water supply system, along with estimation of the sensitivity of each of them to changes in the system stressors. Results are presented both in terms of maps of current and future vulnerability under a given CGM/scenario combination and, more in detail, as sets of response surfaces for

individual basins. More detail on the procedure and on the general findings is provided in Chapter 2 and Appendix D.

1.1.2 Summary and conclusions

This study presents a general, consistent and systematic procedure to assess the impact of climatic and socio-economic changes on the US water supply system. We estimate the PDFs of annual water supply for each of the 98 ASRs making up the contiguous US by evaluating annual water yield as modified by the effect of storage and both natural and artificial water transfers. We then evaluate the probability of shortage by comparing the PDFs of water supply with PDFs of desired water uses. A subset of IPCC SRES climatic and socio-economic scenarios was used in combination with a set of GCMs to project the vulnerability estimates for the entire 21st century.

Results are presented in general terms for the entire US while a deeper analysis is performed for a set of selected basins of the Colorado River Basin, California and the High Plains.

Our approach allows us to examine the response of each given basin to potential changes in the future demand and supply, to identify the individual impact of changes in the moments of the PDFs of demand and supply and to rank basins according to their sensitivity to climatic and socio-economic changes.

We found that the Southwest and central and southern Great Plains are the more vulnerable areas to future climatic and socio-economic changes. In addition, we found that the future decrease in water yield is the main reason for the increasing probability of shortage in those areas. Although climate change is expected to increase water demand, in fact, future water use efficiency improvements will mitigate that impact so that overall increases in desired water use are expected to be modest in comparison with the effect of climatic changes on water yield and thus on water supply. The reductions in yield, on the other hand, are driven by temperature (and therefore potential and actual evapotranspiration) increases, especially where precipitation decreases or increases only slightly.

Looking at individual basins, the Central California and the San Joaquin-Tulare were found to be among the more sensitive to both current and future variability of demand and supply. Large sensitivity is also found for two selected basins of the High Plains, namely the Niobrara-Platte-Loup and Kansas, while the Lower Colorado was found to be the sub-basin of the Colorado River Basin with the larger susceptibility to changes in future supply and demand. However, general disagreement was found among future scenarios both in terms of vulnerability and sensitivity of individual basins.

Although in this study it is applied in the context of water resources, the procedure outlined in this framework is general and offers a versatile and consistent instrument to assess the vulnerability of any physical system to changes in inherently variable stressors.

1.2 Part B: General Motivation

Banded, spotted or labyrinthine vegetation patterns are a common and well-defined characteristic of many arid and semi-arid landscapes and develop at a wide range of spatial scales. Typical scales of individual vegetation pattern elements (i.e., thickness of a band or radius of a patch of vegetation) can span up to two orders of magnitude, ranging from 10^0 to almost 10^2 m. While most of the current literature on the topic approaches the problem of describing and explaining the dynamics of such patterns from a conceptual perspective by building purely mathematical frameworks reproducing those spatial mechanisms of facilitation and competition that are considered responsible for the emergence of self-organized structures, we here try to analyze the phenomenon from a mechanistic perspective. In doing so we explore (some of) the possible physical mechanisms responsible for the establishment of those self-organized, non-random vegetation patterns.

We use a water and energy balance model coupled with a spatial modeling of the mutual land-vegetation interactions in order to provide a fundamental mechanistic understanding of the dynamics of vegetation pattern formation and development. The analysis of the climate-soil-

vegetation system is carried out at the hillslope scale. Average annual precipitation, solar radiation, temperature, humidity and cloud coverage, along with the statistics of the rainy events (length, duration, and intensity of storms, length of the rainy season and number of storms per year) are the forcings of the system at the hillslope scale. The study domain is then subdivided into a study grid of NxN pixels. Reciprocal effects of vegetation on the hillslope thermodynamics, runoff production and run-on infiltration, root density, surface albedo and soil moisture content are modeled at the pixel level. Vegetation emergence and development is analyzed by tracking the vegetal fractional coverage, along with water and energy fluxes at each pixel of the study domain.

1.2.1 Hypotheses and objectives

The main assumption of this study is that vegetation patterns emerge as a result of reciprocal actions - physical, chemical and physiological, (i.e., feedback processes) - between vegetation, hydrologic and climatic processes, and soil properties. In addition, we assume that the spatial distribution of vegetation depends on the spatial distribution of water and energy and that, in turn, vegetation, water and energy are distributed on the hillslope in a way that allows optimal utilization of water and energy by the vegetation in the long term.

The overarching objective of this study is to explore and identify (some of) the physical mechanisms responsible for the establishment of non-random spatial vegetation patterns.

The specific objectives are to:

- develop a physically-based mechanistic understanding of the processes leading to vegetation pattern formation;
- implement such understanding in a mathematical model able to replicate the main physical characteristics of observed vegetation patterns (shapes, dimensions, etc.),
- individuate the relative impact of each process on pattern formation;

- capture the relationships between vegetation patterns and the climatic, hydraulic and topographic characteristics of the system.

1.2.2 Summary and conclusions

A set of climate-soil-vegetation systems was simulated at the hillslope scale, showing that our modeling of the system is able to reproduce the typologies of patterns commonly referred in the literature as bands, spots and labyrinths. The model was validated by comparing simulations with observed natural patterns in the areas of Niger near Niamey and Somalia near Garoowe. Our modeling of the system was able to capture the local dynamics inducing the formation of patterns and generated results that are qualitatively and quantitatively compatible with the observations and the literature.

The analyses of the processes involved in the formation of patterns suggest that the phenomenon is primarily driven by run-on infiltration and local mechanisms of facilitation/competition existing among adjacent vegetation groups. However, even in the presence of those mechanisms, patterns arise only when the climatic conditions, particularly annual precipitation and solar radiation, are favorable. In particular, we found that, with decreasing precipitation or, conversely, increasing incoming radiation, the system drifts from fully vegetated with undistinguishable vegetation structures to self-organized patterns.

In the range of climatic conditions that promote the formation of self-organized vegetation structures, the peculiar spatial features of patterns are determined by the peculiar characteristics of the spatial effects modeled at the pixel level. Our investigations suggest that all the dynamics considered (run-on infiltration, facilitation/inhibition dynamics between adjacent vegetation groups, effects of nutrient and litter transport and deposition, competition for soil moisture through roots and the effect of spatially inhomogeneous surface reflectance) have an impact on the spatial configuration of vegetation over the hillslope. Nonetheless, as anticipated before, our analyses indicate that the phenomenon is mainly driven by surface runoff production and run-on

infiltration. No self-organized structures, in fact, were observed in the absence of any surface runoff production and subsequent run-on infiltration. The effects of facilitation/competition existing between adjacent vegetation groups, roots and albedo have also a significant impact and affect pattern definition and shape, as well as the total biomass that establishes in the domain.

Chapter 2 Vulnerability of the US water supply to shortage

2.1 Abstract

The environmental and socio-economic vulnerability of a system depends not only on its ability to withstand stresses and on the magnitude of those stresses but, most important, on the inherent variability of their hydro-climatic and socio-economic drivers (*e.g.*, water supply capacity, water demand, precipitation, evapotranspiration, population, technology, water management infrastructure, etc.). Thus, in order to quantify current and future vulnerability one must do so probabilistically. Accordingly, we introduce a probabilistic framework for vulnerability analysis and use it to quantify current and future vulnerability of the highly interconnected water supply system of 98 basins covering the contiguous United States.

Water supply of each basin consists of fresh water yield as altered by storage, trans-basin diversions, and other management impacts and is represented by a time dependent probability distribution function (PDF). Fresh water yield is estimated as precipitation (P) minus evapotranspiration (E) for each point in a fine grid covering the study area, and then aggregated by river basin. The PDF of available water supply, when compared with the PDF of the desired water use, yields an estimate of the probability of shortage, and thus a measure of the vulnerability of the water supply system.

We determine the projected changes in vulnerability as well as the relative contributions to those changes from changes in climatic and socio-economic drivers. Vulnerability was found to vary in magnitude and spatial distribution depending on which IPCC emission scenario is chosen to represent future socio-economic conditions and on which global climate model is used. Nevertheless, the southwestern portion of the US including California in the west and the

southern Great Plains to the east was consistently found to face the greatest likelihood of future shortages—these areas being where socioeconomic and climatic changes converge most strongly to increase pressures on water resources. Although we focus here on the water supply system, the methodology is general. We contend that it should be applied to any environmental and socio-economic vulnerability analysis, as it is the only methodology that both accounts for the probabilistic character of the drivers and allows for explicit inclusion of thresholds.

2.2 Introduction

Offstream water use in the United States increased over ten-fold during the twentieth century in response to tremendous population and economic growth (Brown, 2000). Although water efficiency has improved in the last few decades as a result of technological advances and management, rising incomes and urbanization put pressure on water supply and call for measures to protect stream water quality and maintain habitat for endangered aquatic species (Gillilan and Brown, 1997). Complicating the picture, climatic change is increasing hydrologic uncertainty. Taken together, these forces are making careful water management ever more important. Thus, a realistic broad-scale understanding of the vulnerability of the US water supply system to shortage must be a component of any attempt to define the magnitude of the threat, and it is essential in determining appropriate mitigation and adaptation measures.

In general, vulnerability of a system is a function of the extent to which it can be damaged by the impact of an external hazard. The definition of vulnerability and the implication that it has on the approach used to estimate it has been the focus of several papers (Blaikie, Cannon et al. 1994),(Kelly and Adger 2000). In its forth assessment, the Intergovernmental Panel on Climate Change (IPCC) defined vulnerability as “the degree to which these systems are susceptible to, and unable to cope with, adverse impacts” (Schneider, Semenov et al. 2007). While this last definition seems to be widely accepted, it still leaves room for debate concerning the actual quantification of the “degree of susceptibility to adverse impacts”. In the context of water

resources, many studies have estimated the potential impact of future climate and socio-economic scenarios on water resources by estimating future projections of a set of water stresses indicators (Postel 2000),(Vörösmarty, Douglas et al. 2005), (Weib and Alcamo 2011). However, no generally accepted paradigm to quantify vulnerability has been established so far.

In agreement with the IPCC definition, we observe that, conceptually, the vulnerability of a system is a function of its ability to respond (*i.e.*, cope with; adapt) to inherently variable stressors. In this study, given the uncertainty characterizing both the stressors and the capacity to withstand them, we advocate the necessity of quantifying vulnerability under a probabilistic framework. In particular, we estimate vulnerability as the probability that a critical system threshold, itself a function of both the capacity and the stressors of the system, will be crossed. In the context of the US water supply system, this definition translates into evaluating the probability that, at given time and place, water demand exceeds water supply. In other words, we define vulnerability as the probability of shortage.

In order to estimate the vulnerability of the US water supply system to shortage, we developed a probabilistic, physically based, spatial and temporal characterization of hydrologic fluxes (e.g., precipitation, evapotranspiration, surface and subsurface runoff), soil moisture storages, water demands, reservoir storage and water transfers.

We quantify water supply at a certain location as water yield, *Y*, plus inflow from upstream subjected to the effect of management via reservoir storage and water diversions. We used Eagleson's (1978a-g) one-dimensional, statistical-dynamical water balance model to estimate water yield as the difference between precipitation, *P*, and actual evapotranspiration, *E*. We computed water demands as threshold amounts of desired water use based on historical records of water withdrawals and on projections of water use drivers and water use efficiency rates. Estimation of future demands include both deterministic and stochastic (climate depending) components. We finally simulate the US water supply system using a hydrologic network model in order to account for water routing and management via reservoir storage and water transfers.

We estimate the vulnerability of water supply to shortage for each of the 98 basins called Assessment Subregions (ASRs), which are subbasins of the 18 Water Resource Regions of the coterminous US (Figure 2.1). The vulnerability analysis was extended throughout the entire 21st century in light of projected socioeconomic and climatic changes.

2.3 Fresh water yield

Evaluating fresh water yield, *Y*, (that is, the sum of surface and sub-surface runoff), is the first step towards the estimation of the water supply. Water supply for each element of the US water network (that is, each one of the 98 ASRs), in fact, is determined as fresh water yield plus the effect of water redistribution (via natural flow and artificial diversions) and storage. In this framework we use Eagleson's (1978a-g) one-dimensional, statistical-dynamical physically-based annual water balance model to estimate water fluxes for each ASR and evaluate *Y* as *P-E*. The model uses probability distribution functions (PDFs) of precipitation, temperature and other hydrologic variables plus soil hydraulic parameters as input and produces, as output, the PDF of water fluxes (surface runoff, groundwater runoff, evapotranspiration), the average soil moisture content and the vegetation fractional coverage. The model was implemented on a 5x5 km grid for the US and used to determine water fluxes for historical (1953-2005) and projected future climate (2006-2100). Results were then aggregated at the basin level. See Appendix A for details.

2.3.1 Model description

Eagleson's model is a physically-based description of water fluxes across the soil-atmosphere interface—a one-dimensional representation of soil moisture dynamics as forced by a stochastic climate (Eagleson, 1978a-g). It describes the relationship between annual amounts of precipitation, runoff, infiltration and evapotranspiration as a function of volumetric soil moisture and soil and vegetation characteristics. The model is based on the steady-state solution of the water balance equation:

$$\int_0^t \left\{ i(t) - e_T(t) - \frac{\partial}{\partial t}\left[V_{ss}(t) - V_{rg}(t)\right] \right\} \cdot dt = \int_0^t \left[r_s(t) + r_g(t) \right] \cdot dt \qquad (2.1)$$

where $i(t)$, $e_T(t)$, $V_{ss}(t)$, $V_{rg}(t)$, $r_s(t)$ and $r_g(t)$ are respectively: the precipitation intensity, the evapotranspiration rate, the volume of water storage in the surface, the volume of water storage in the sub-surface, the surface runoff rate and the sub-surface runoff rate.

The integration of equation 1 for an arbitrary period of time is analytically intractable for three reasons: (1) the climatic forcing (i.e., precipitation and potential evapotranspiration) is stochastic; (2) all terms depend on soil moisture content, which is difficult to evaluate or measure; and (3) the integration requires the evaluation of carry-over storages.

However, under the assumption that the hydro-climatic system is in equilibrium in the long-term mean sense, it is possible to obtain an analytical solution of equation (2.1). This solution implies that the long-term average of the change in total soil moisture storage (above and below the surface) is zero and leads to the following final formulation of the long-term mean annual water balance:

$$E[P_A] - E[E_{T\,A}] = E[R_{S_A}] + E[R_{g_A}] \qquad (2.2)$$

where $E[P_A]$, $E[E_{TA}]$, $E[R_{SA}]$ and $E[R_{gA}]$ are respectively: the expected value of annual precipitation, the expected value of annual evapotranspiration, the expected value of annual surface runoff and the expected value of annual groundwater runoff.

We express all terms in equation (2.2) as analytical functions of soil moisture content, the characteristics of the stochastic precipitation input, the potential rate of evapotranspiration, the physical properties of the soil (e.g., porosity, intrinsic permeability, pore disconnectedness), and the properties of the vegetation (plant transpiration efficiency and fractional vegetation cover).

Because soil and climate parameters are fixed for the given control volume, the water balance equation is a function of two unknowns: the average soil moisture content, s_0, and the vegetation

fractional coverage, *M*. In order to obtain a unique solution, we further assume that, in the long term, vegetation operates under conditions of minimum stress (Eagleson 1978).

2.3.2 Data requirements

Eagleson's model requires a set of soil and vegetation parameters, as well as climatic input in the form of long term average statistics of the distribution of rainy events. Soil hydraulic parameters are: total porosity, saturated hydraulic conductivity, matric potential at effective saturation, pore size distribution index, pore disconnectedness index, and diffusivity index. Vegetation is characterized by the transpiration efficiency. Climate input is represented by annual values of precipitation and potential evapotranspiration and by the average long term statistics of storm events, namely: storm interarrival time, time between storms, number of storms per year, rainy season duration, storm duration, storm intensity, storm depth and the parameter of the gamma distribution of storm depth.

The VEMAP dataset (Kittel, Rosenbloom et al. 1995; Kittel, Rosenbloom et al. 1996) was used to estimate the soil hydraulic parameters. It provides bulk density, ρ, and soil texture (i.e., percentages of sand, silt and clay) at a 0.5º x 0.5º resolution for the whole US. A simple inverse distance squared method was used to convert the database from the original 0.5º spatial resolution to the 5x5 km grid. Estimates for the following soil hydraulic parameters were obtained from the VEMAP dataset: pore size distribution index, m, residual water content, θ_r, matric potential at effective saturation, $\Psi(1)$, saturated hydraulic conductivity, $K(1)$, pore disconnectedness index, c, and diffusivity index, d.

Precipitation data from 2088 hourly data gages, provided by the National Climatic Data Center (NCDC), were used to determine the storm statistics. Long term storm statistics were evaluated at each gauging station and then extended to the entire US at the 5x5 km spatial resolution.

Annual climatic data were retrieved from the PRISM database (Daly, Neilson et al. 1994) which provides annual values of precipitation, minimum and maximum temperature, and dew point temperature at the 5x5 km resolution. Annual potential evapotranspiration was calculated using a modification of Penman's equation by Linacre (1977).

Transpiration efficiency, k_v, is defined as the ratio of potential transpiration to potential evaporation from bare soil under the condition of unlimited water supply. Transpiration efficiency is specific for each given soil-climate-vegetation system and depends on both the vegetation species and the environmental conditions. Giving the lack of direct measurement or any reliable estimate available at the scale of our analysis, this parameter is estimated during model calibration rather than evaluated a priori.

2.3.3 Model calibration

Model was calibrated by matching the long term mean values of historical streamflow and modeled yields. The calibration procedure consisted in minimizing the mean squared error between the sequences of annual historical streamflow and modeled yields. The plant transpiration efficiency, k_v, was used as the primary calibration knob.

Model calibration was performed at the basin scale using the following historical datasets on streamflow:

- 42-year series (1953-1994) of annual streamflow records for 655 relatively unmolested test basins across the US, included in the Hydro-Climatic Data Network (HCDN) (Slack and Landwehr 1992; Hydrosphere Data Products 1996).
- 52-year series (1953-2004) of reconstructed natural flows for a set of watersheds in the Colorado River Basin (CRB), provided by the Bureau of Reclamation.
- 30-year average reconstructed natural streamflows estimated for the 2,148 cataloging units (hereafter referred to as 8-digit basins) of the US (Krug, Gebert et al. 1989).

A map showing the 655 test basins, the 8-digit basins and the set of watersheds of the CRB used for calibration is provided in Figure 2.2.

The 655 test basins were selected as the primary calibration dataset because of the length and the reliability of the records. In those areas of the US without test basins, with the exclusion of the Colorado River Basin, the model was calibrated over the 30-year average flows estimated at the 8-digit basins level. For the Colorado River Basin, the 8-digit basin information was used in combination with the streamflow records provided by the USBR. However, although the information provided by the USBR dataset was considered more reliable than the 8-digit basin averages provided by the USGS for the same area, the typical drainage area of the USBR dataset was considerably larger than the typical 8-digit basin. In order not to lose spatial detail, the calibration was performed at the 8-digit basin level, by scaling their average streamflow in such a way that the cumulative streamflow predicted by each set of 8-digit basins included in a USBR catchment would match the USBR average for that catchment. This procedure guarantees to match observations for those large catchments whose streamflow records are available from USBR and, simultaneously, to preserve the smaller scale variability at the 8-digit basin.

2.4 Water storage

Water storage capacity for each ASR was determined by aggregating storage capacities of natural and man-made impoundments. The June 2009 version of the National Inventory of Dams from the US Army Corps of Engineers was the primary source of information of reservoir storage. Not all of the reservoirs listed in the inventory were used in determining the aggregate storage capacity for each ASR. The final set of reservoirs selected includes only reservoirs with a normal surface area of at least 5 km^2 and excludes tailings ponds, cooling ponds, and reservoirs whose only purpose is flood control. The final list of reservoirs consists of 1196 reservoirs. The aggregate water storage capacities of the ASRs range from 0 to over 40 million acre-feet. Thirteen ASRs have at least 10 million acre-feet of storage.

Storage-surface area relationships were determined for each ASR, in order to compute evaporation from reservoirs. Except for ASRs 1404 and 1503, these equivalent volume-surface area relationships were obtained by regionalized regressions based on the individual relationships of the included reservoirs. For ASRs 1404 and 1503, because Lake Powell and Lake Mead constitute the only significant storages, the corresponding individual surface area-to-storage relations were directly used.

2.5 Water use classes

In this framework we individuated three classes of water use for each ASR: in-stream flow requirements, trans-ASR water diversions and consumptive uses.

2.5.1 In-stream flow requirements

In-stream flow requirements at any point in a basin refer to the magnitude and temporal distribution of flows required to ensure adequate supply for ecosystems maintenance. Careful determination of in-stream flow requirements involves a complicated mix of socio-economic, biological and environmental factors, which is not practical at the ASR scale. In this study we adopt the general guidelines delineated by Tennant (1976), and set the in-stream flow requirement of each ASR for both current and future conditions as 10% of the corresponding average historical streamflow for the period 1953-1985.

2.5.2 Water diversions

Trans-ASR diversions represent water diverted from one ASR to another as the result of legal agreements between the jurisdictions involved. Most of the information regarding inter-basin diversions was taken from two publications of the USGS regarding respectively the western (Petsch 1985) and the eastern (Mooty and Jeffcoat 1986) US; from these publications we used an average of the estimates for the three most recent years (1980-82). Those estimates were supplemented by more recent sources of information where available, specifically for the

Colorado River Basin and California (Litke and Appel 1989; California Department of Water Resources 1998; Colorado Water Conservation Board 1998; Colorado Water Conservation Board 2010). Data for inter-basin transfers were aggregated by ASR.

While each water agreement usually establishes a long-term average for the amount of water to be diverted from one basin to another, the actual amount of water transferred for each diversion typically fluctuates over time. Unfortunately, adequate documentation on the operating rules used to determine the actual amount in any one year was usually unavailable, and in any case tended to be too detailed to be of practical use in this large-scale study. For our purposes, therefore, the amount of water diverted each year was set equal to the historical average for each trans-ASR diversion. This amount was kept constant overtime both for current and future climatic and socio-economic scenarios.

2.5.3 Consumptive use

Estimates of water withdrawal across the US at a fairly fine scale are available at five-year intervals from the USGS for the period 1985-2005 (Solley, Merk et al. 1988; Solley, Pierce et al. 1993; Solley, Pierce et al. 1998; Hutson, Barber et al. 2004; Kenny, Barber et al. 2009). These data, along with data on water use drivers and water use efficiency rates, were used to simulate past and current conditions and as a basis for projecting future levels of desired water withdrawal (from surface and ground water combined) by ASR. Consumptive use proportions from the USGS for years 1985, 1990, and 1995 were then used in converting estimates of withdrawal to estimates of consumptive use (water depletion). The resulting projections of consumptive water use, called demand here, were produced for six categories of water use—domestic and public, industrial and commercial, thermoelectric, irrigation, livestock, and aquaculture—which were then aggregated to a single estimate of demand.

Demand (D) for a given water use category and future year was estimated as:

$$D = \alpha \cdot (U \cdot \Phi + \Delta W) \tag{2.3}$$

where U is number of demand units such as a person for domestic use or an irrigated acre for agricultural use, Φ is withdrawal per demand unit, also called a water use efficiency factor, ΔW is future withdrawal attributable to climatic or other changes that are largely unrelated to past levels of water use, and α is the portion of withdrawal that is consumptively used. Future levels of U and Φ were estimated by extending past trends that in all cases have been nonlinear. This approach assumes that water supply will be no more limiting to growth in demand than it has been in the recent past.

2.6 The US water supply system

Estimating annual water yield as the difference between precipitation and evapotranspiration does not provide, by itself, a valid estimation of the water effectively available at a certain point in space. A thorough characterization of water supply for any area, in fact, must account for the natural and man-made water networks that redistribute water on the surface. That makes the US water supply system a complex and highly interconnected structure that can be represented as a set of networks consisting of nodes and links.

We identify the water networks existing in the US at the ASR level by analyzing both natural and artificial flow routes. Two or more ASRs are considered part of the same network when they are connected by a sequence of water links, either natural (due to natural upstream to downstream flow) or artificial (*via* water diversions). The resulting water supply system consists of three multi-ASR networks and 15 single-ASR systems (Figure 2.3). The biggest of the three multi-ASR networks includes 69 ASRs in the central and western US. The other two networks include respectively 10 ASRs in the Northeast and four ASRs in the Southeast. Of the 15 single-ASR systems, eight drain to the ocean, five into Canada, and two are closed basins.

Simulations of the US water supply system at the ASR level were performed using MODSIM (Labadie, Pineda et al. 1984). The simulations provide annual values of water flows in any link,

storage levels in each ASR, water evaporated from storage, and water assigned to each demand, all of which depend both on climate and, critically, on the set of priorities for water allocation. Ideally, the priorities would represent the thousands of detailed agreements about water storage and allocation that exist across the country. Lacking information on many of these agreements, we implemented the following priorities: (1) in-stream flow requirements, (2) trans-ASR diversions, (3) consumptive water uses, and (4) reservoir storage. These priorities recognize the importance of guaranteeing a minimal amount of water for environmental and ecosystem needs before water is diverted for other uses, and allow trans-basin diversions to occur before within-basin diversions. For multi-ASR networks, water demands belonging to the same category were assigned the same priority regardless of their position in the network. Because storage was assigned the lowest priority level, water is stored in a given year only after all the demands reachable by that reservoir are satisfied. Water stored at the end of one year, minus the evaporation loss, is available for use the next year. Further details are provided in Appendix B.

2.7 Future projections

The analysis of the future US water system requires the estimation of future water demand and supply. This, in turn, calls for estimates of future water yield, storage and routes of water redistribution. Projections of future water yield were obtained by applying Eagleson's model using input climatology from a set of General Circulation Models (GCMs) used in combination with IPCC SRES socio-economic scenarios. See Appendix C for details.

2.7.1 Future GCM and scenarios

The increasing globalization of the world economy and the possibility of substantial climatic change have created considerable uncertainty about future US water supply and demand. One way to capture this uncertainty, adopted by the IPCC, is to examine various possible future scenarios. The emissions scenarios examined in the 4th IPCC Assessment differ in population

growth, economic growth, and energy use (among other things), and thus in greenhouse gas (GHG) emissions and CO_2 levels.

For this study, three scenarios (A1B, A2, and B2), each based on a different storyline, were chosen from the IPCC SRES set (Nakicenovic, Alcamo et al. 2000). Of the three scenarios used here, the A2 scenario is the most extreme and B2 the least extreme in terms of atmospheric CO_2 concentration, although the projections for CO_2 of these scenarios do not differ greatly until later in the 21st century.

Scenarios were used in combination with the following GCMs: the Canadian Centre for Climate Modelling and Analysis Third Generation Coupled Global Climate Model Version 3.1 Medium Resolution (CGCM31MR) and the Australian Commonwealth Scientific and Industrial Research Organisation Mark 3.5 Climate System Model (CSIROMK35).

For each one of the six GCM-scenario combinations, projections of monthly precipitation, minimum and maximum temperatures and potential evapotranspiration were obtained for years 2001 to 2100 (Joyce, Price et al. in preparation 1; Joyce, Price et al. in preparation 2). The GCM projections were downscaled to a uniform grid of 10-km cells over the conterminous US using the ANUSPLIN method (Price, McKenney et al. 2006).

The downscaled GCM projections were corrected for biases before they were used to estimate future demands and future water yields. Biases in precipitation, temperature and evapotranspiration were determined as the difference between downscaled mean projections for the period 2001-2008 and the corresponding observations retrieved from the PRISM datasets.

Predicted changes in annual precipitation are quite erratic both in magnitude and direction and little agreement is found between different GCM-scenario combinations. In general, while local changes can be large, aggregated US precipitation is not expected to change dramatically over the next decades. The US projections by the CGCM31, for example, predict an increase in mean precipitation of only 2.3% and 1.8% for scenarios A1B and A2, and a 3.6% decrease for the B2 scenario by 2060, while predicting decreases of 2.1%, 10.8% and 14.3% for the respective

scenarios when only the Colorado River Basin is considered. Figure 2.4 shows that the mean precipitation projected by the CGCM/A1B scenario for the 21st century is expected to consistently increase in most of the Northeast and in Texas and decrease in the West. Besides this large scale behavior, however, coherent patterns of changes in precipitation are not easily identifiable and little coherence if found in the spatial distributions of precipitation projected by the various GCM/scenarios.

When looking at potential evapotranspiration rates, all GCM-scenarios agree in projecting consistent increases all the way to the end of the century. The CSIROMK35 model is the most extreme and produces increases in the average potential evapotranspiration for the US that exceed 30% by 2060. Changes predicted by CGCM31 model are smaller and suggest increases of 15.8%, 25.9% and 14.1% respectively for the A1B, A2 and B2 scenarios by 2060, where for the CRB alone the same model predicts increases of 21.1%, 25.1% and 18.3%.

2.7.2 Future water yield

Projections for precipitation, temperature and potential evapotranspiration given by the six GCM-scenario combinations were used in the water balance model to estimate future traces of water yield for the period 2006-2100. Because the projected changes in water yield are the direct result of the projected changes in precipitation and potential evapotranspiration rates, some of the trends of future yield can easily be anticipated. The overall picture, in fact, shows that water yield decreases throughout the 21st century, with an average decrease in yield of 15.8%, 21.7% and 17.2% predicted by 2060 by the CGCM31/A1B, A2 and B2 scenarios. Figure 2.5 shows that the mean water yield projected for the CGCM/A1B scenario is expected to consistently decrease in most of the US with the exception of the Southwest. This is apparently in contradiction with the fact that precipitation is expected to decrease and temperatures are projected to increase in the Southwest. However, it can be explained by observing that knowing only the changes in the mean values of precipitation and evapotranspiration may not, by itself, suffice to anticipate the change

in yield. In some circumstances, in fact, the higher moments of their distribution functions affect strongly the distribution function of yield, leading to some apparently counterintuitive results. This is especially the case of the lower Colorado River Basin, where the CGCM/A1B predicts an average increase in yield larger than 20% by 2060, despite of a decreasing precipitation and an increasing potential evapotranspiration. This large percentage increase (although in its absolute value corresponds to only a fraction of a cm), is caused by the increases in the variance of both predicted precipitation and potential evapotranspiration. As a result of a larger variance of the climatic forcing, in fact, the distribution of extreme events can be affected in a measure that leads to an increase rather than a decrease in the average water yield. This is especially true in arid climates, as a consequence of the highly skewed distributions of precipitation and water yield.

2.7.3 Future demands

The population and income projections of the three IPCC scenarios (A1B, A2 and B2) were updated for the US and disaggregated to the county level. The updates utilize the US Census Bureau's (U.S. Bureau of Census 2004) national moderate growth population projection as an update of the original A1B scenario estimate for US population - A1B being the scenario that most clearly represents a continuation of business as usual in US population growth. The recent Census Bureau population projection incorporates the 2000 Census, which the original A1B scenario did not. The IPCC projections for scenario A2 were then updated in relation to the revised A1B projection by maintaining the proportional differences among the projection paths for the US of the original IPCC scenarios. Estimates were disaggregated to the county level utilizing Census data and a socio-economic projection model (Zarnoch, Cordell et al. 2008). To allocate county estimates to ASRs, year 2000 census tract data were used to determine the proportion of a county's population occurring in each ASR. Details on projections of other drivers are available from the authors.

In the absence of climate change, water withdrawal efficiency - domestic and public withdrawal per capita, industrial and commercial withdrawal per dollar of income, thermoelectric withdrawal per kWh of electricity produced, agricultural withdrawal per acre irrigated, and livestock and aquaculture withdrawal per capita - was projected to improve in all sectors but aquaculture. Changes in most drivers of water use - especially population and per-capita income - are expected to increase pressure on water supplies, but projected decreases in irrigated agriculture in the West will help alleviate pressures. Combining these factors, in the absence of climate change but assuming the A1B scenario increases in population and income, aggregate US demand is projected to increase by only 5% from 2005 to 2060 despite a 51% increase in population.

Climate change is projected to increase demand substantially. For example, with the A1B scenario, and averaging results from the three global climate models, aggregate US demand is projected to increase from 2005 to 2060 by 27% as compared to only 5% without climate change. Of the 22% difference, 76% is due to increases in agricultural irrigation, 10% to increases in landscape irrigation, and 14% to increases in withdrawals at thermoelectric plants to handle the increase in space cooling demand.

The irrigation changes are computed as a function of changes in precipitation and potential evapotranspiration assuming that irrigation fully meets crop water demand, and thermoelectric changes are computed as function of changes in temperature based on the work of Sailor (Sailor 2001; Sailor and Pavlova 2003). Because of the many climatic and socioeconomic differences across the US, ASRs differ considerably in projected water demand. Based on a multi-model average, withdrawals from the current period to the 2060 period with the A1B scenario are projected to drop in 11 ASRs, increase by less than 25% in 34 ASRs, increase by from 25% to 50% in 35 ASRs, and increase by more than 50% in the remaining 18 ASRs. All of the ASRs where withdrawals are projected to drop are in the East, but ASRs where withdrawals are projected to increase by more than 50% are scattered through the US (see Figure 2.6).

2.8 Vulnerability assessment

In general, vulnerability of a system is a function of the system's resilience and robustness with respect to the inherent variability of the main design variables. Thus, in order to quantify vulnerability one must do so probabilistically. We define vulnerability in the manner of Korchendorfer and Ramirez (1996) and quantify the vulnerability of the US water supply to shortage by determining the probability that water demand will exceed water supply. This is done by using the time-dependent probability distribution functions of supply *(S)* and demand *(D)* and estimate vulnerability (V) for each ASR as:

$$V(t) = \Pr[S(t) < D(t)] \tag{2.4}$$

Probability distributions for D(t) and S(t) are estimated for each ASR for current conditions and for future projections by simulating the US water supply system for each GCM-scenario combination. As a consequence of such definition, vulnerability of water supply to shortage is a function of not only the mean water supply and demand at a given location, but also of their respective variance and covariance and, in general, of all the moments of their distributions. In a context of hydro-climatic and socio-economic variability, then, it is not sufficient to quantify the effects of changes in the mean values of hydro-climatic and socio-economic variables of interest, but most important, it is necessary to quantify the effects of changes in their inherent variability.

PDFs of supply and demand were obtained by simulating the US water supply system from 1953 to 2090. All simulations used reservoirs half full as initial condition on water storage and considered the period 1953-1985 as transient. Current vulnerability was evaluated for each ASR over the 20-year period of 1986-2005. Future vulnerability was estimated for four 20-year periods centered at 2020, 2040, 2060 and 2080 assuming no changes in storage, in installed trans-ASR diversion capacity, in-stream flow requirements and in the physical structure of the US water supply network. More details are provided in Appendix D.

2.8.1 Current conditions

The climate of the period 1986-2005 was taken as the *current* climate. Annual records of precipitation, temperature and potential evapotranspiration for the period 1986-2005 were used as input for the water balance model in order to estimate the annual water yield, while withdrawal records were used as explained earlier to determine the corresponding water demands.

The analysis of the period 1986-2005 is carried out not only to provide an estimate of the current probability of shortage, but mostly to set the benchmark to which future projections can be compared.

Defining water surplus, Z, as the difference between water supply and water demand, vulnerability is the probability that the water surplus is zero or negative. By simply looking at the first moments of the water surplus PDF, one can notice that vulnerability increases as the mean of the surplus μ_z decreases and as its variance σ^2_z increases (provided that $\mu_z>0$). Taking into account both effects simultaneously, one may also quantify vulnerability as a function of the ratio of the mean surplus to the corresponding standard deviation, $\beta = \mu_z/\sigma_z$, referred to hereafter as the reliability index. The reliability index quantifies in units of the standard deviation how far from shortage a given location is.

Maps of current vulnerability and reliability ratios are presented in Figure 2.7 and show that the water supply system for much of the US west of the Mississippi river is vulnerable under current hydro-climatic and socio-economic conditions. However, only a few areas show vulnerability values exceeding 0.05 at the ASR scale, and they tend to be those that rely heavily on mining of groundwater.

2.8.2 Future projections: maps and response surfaces

Future vulnerability is evaluated for each of the six GCM-scenario combinations for the target 20-year periods centered around 2020, 2040, 2060 and 2080.

In addition to future vulnerability maps of the entire US, we present a deeper analysis of ten selected basins by using climate impact response surfaces. These response surfaces are a useful tool to represent the first order response of a given system to changes in two specific triggers (in our case supply and demand) and allow a widespread analysis of the impact that future variations in those triggers have on the system itself (Weib and Alcamo 2011). In this paper, we produce response surfaces to changes in future average and standard deviation of water supply and demand for selected basins of the High Plains, Colorado River Basin and California.

2.8.3 Sensitivity of vulnerability to changes in the drivers

Changes in future vulnerability of water supply to shortage are not only a function of the actual changes in future climate and in future demands, but also depend on the sensitivity of vulnerability to changes in demand and climate. Understanding how a given location responds to potential changes in climatic and socio-economic conditions is essential for future water management planning and for the individuation of effective measures of adaptation.

We express the differential vulnerability as follows:

$$dV = \frac{\partial V}{\partial \mu_S} d\mu_S + \frac{\partial V}{\partial \mu_D} d\mu_D + \frac{\partial V}{\partial \sigma_S} d\sigma_S + \frac{\partial V}{\partial \sigma_D} d\sigma_D + \frac{\partial V}{\partial \operatorname{cov}(S,D)} d\operatorname{cov}(S,D) \qquad (2.5)$$

Expression 5 captures the individual contributions of the drivers of vulnerability to its total change and shows that the total change in vulnerability depends not only on the actual changes in demands and supply, but also on the sensitivity of vulnerability to unit changes in demand and supply. In turn, those sensitivities are functions of the mean, variance and covariance of P, E, and D.

A map of the sensitivity of vulnerability to changes in mean supply, μ_S, that is, the change in vulnerability per unit change in mean supply, μ_S, for the CGCM/A1B scenario is shown in Figure 2.8. The map shows that western US and, in general, the areas exhibiting the lower reliability index are the most sensitive. A similar behavior, also shown in Figure 2.8, occurs with respect to

changes in mean demand, μ_D, but with opposite sign. Therefore, these areas, in addition to being quite vulnerable under the current conditions, are more prone to large increases in vulnerability for the same change in S, and D.

Response surfaces showing the sensitivity of vulnerability to changes in mean water supply and mean water demand (given the current standard deviations of supply and demand) for CGCM/A1B scenario are shown in Figure 2.9. Obviously, sensitivity to changes in water supply is always negative, meaning that at the increase of mean water supply the probability of shortage decreases. However, the magnitude of the sensitivity is found to be dependent on both space (confirming spatial variability observed in Figure 2.8) and time, since projected changes in the triggers (supply and demand) will affect the future sensitivity of each system. Notably, all selected basins, with the exception of the Little Colorado River Basin, the Gila River Basin and Central California are expected to become more sensitive to changes in the average water supply overtime as consequence of the projected demand and supply. The same happens for the sensitivity to changes in water demand indicated in the response surfaces in Figure 2.10. It is interesting to notice that, for each basin, the magnitude of the sensitivity to mean demand and supply (for a given set of standard deviation of demand and supply) has a maximum when the average demand equals the average supply, as analytically shown in the appendix. This condition is crossed or reached by 2080 by several of the selected basins, namely the Kansas River Basin, the Lower Colorado River Basin, the San Joaquin-Tulare River Basin and the Central California. Once the mean demand exceeds the mean supply, the actual vulnerability to shortage becomes larger than 0.5, although the vulnerability itself becomes less sensitive to further changes in mean demand and supply.

Changes in the mean of the drivers, however, are not the only source of changes in vulnerability; the probability of shortage of each system is also impacted by changes in the variance and co-variance of supply and demand. Given that the current average supply exceeds the average demand for all selected basins (and, in general, for the entire US with the exception

of the Sevier Lake Basin), as the variance of either supply or demand increases so does vulnerability, as shown in Figure 2.8.

Response surfaces of sensitivity of vulnerability to changes in standard deviation of supply and demand (given the current standard deviations of supply and demand) for CGCM/A1B scenario are shown in Figure 2.11 and in Figure 2.12. Because in all the cases considered the mean supply is larger than the mean demand, the sensitivity of vulnerability to changes in standard deviations of those two triggers is always positive, meaning that vulnerability will increase as a result of a more variable supply or demand. Unlike the case of sensitivity to mean supply and demand, the absolute maximum of the response functions to changes in standard deviations does not have an intuitive interpretation. However, the sensitivity of vulnerability to changes in standard deviation of water supply (/demand) is zero when the standard deviation of water supply (/demand) is itself zero. Figure 2.11 shows that the sensitivity of vulnerability to the standard deviation of water supply is projected to decrease throughout the 21st century for all the selected basins, with the exception of the San Joaquin-Tulare and Central California. Sensitivity to standard deviation of water demand, on the other hand, is expected to decrease in the Colorado-San Juan, in the Little Colorado, Gila, and in the Sacramento-Lahontan, Figure 2.12.

2.8.4 Vulnerability assessment: the CGCM/A1B scenario

Changes in future vulnerability reflect changes in the probability distribution functions of S and D. The US maps shown in Figure 2.17 show the vulnerability predicted for the entire US by CGCM31/A1B scenario for the periods 2020, 2040, 2060, and 2080. It is noticeable how the southwestern US and the Great Plains areas are projected to face the greatest increases in vulnerability, in addition to being already the areas where shortages are more likely. Large increases in vulnerability are expected throughout the entire 21st century for the lower CRB, the central Great Plains, and the central California. Large increases are also expected in the Rio Grande basin, Texas and Utah. Noticeably, all those areas affected by the larger increase in

vulnerability are characterized by having a current reliability index less than two. Interestingly, vulnerability increases overtime in most of the country, with the highest increases affecting those areas where the current likelihood of shortage is larger. Decreases in vulnerability, on the other hand, are quite small and are only expected in few areas of the eastern US, Midwest and northern US.

In an effort to isolate individual contributions of changes in the distributions of supply and demand to total vulnerability, we produced maps showing the effect on vulnerability due to changes in the mean and in the standard deviation of *S* and *D* for the CGCM/A1B in Figure 2.18 and Figure 2.19 for the target years 2020, 2040, 2060 and 2080.

The cumulative effect of changes in mean S and in the standard deviation of S is expected to increase vulnerability for the central Great Plains throughout the entire 21st century. The Rio Grande Basin and the Colorado River Basin, on the other hand, are expected to undergo alternate periods of increasing and decreasing vulnerability due to changes the PDF of S. A similar behavior is detected in central California, where changes in the PDF of S are expected to determine an increase in vulnerability for the periods 2020 and 2060 and a decrease in 2040 and 2080. As for the rest of the country, the contribution of changes in the moments of S is projected to have negligible effect or to lead to decreases in vulnerability. Although not shown, the latter situation is primarily due to smaller variance in water supply projected through the course of the 21st century for the northwest, the northern Great Plains, the northern California and Texas. Storage capacity is not projected to be limiting as vulnerability increases. In fact, simulations predict a steady decline in the water levels of the main reservoirs of all the areas where vulnerability is expected to increase overtime, indicating that water scarcity is primarily due to demand-supply imbalance rather than to insufficient storage capacity.

Unlike the case of the water supply, the effect of changes in both mean and standard deviation of water demand is projected to always increase future vulnerability, with the noticeable exception of the Sevier Lake and the Rio Grande basin in 2020. The effect of water demand

change is larger in the Southwest, in central California and in the southern Great Plains, while being essentially negligible in the rest of the country. Changes in vulnerability resulting from projected changes in the *S* and *D* are of the same magnitude or larger than those from changes in the corresponding means over most of the US, except in central coastal California, in the San Joaquin river basin and in southern Florida.

A deeper analysis of the individual impact of changes in the mean supply and demand for the selected basins of the Colorado River Basin, California and the High Plains is shown through the response surfaces of Figure 2.13. All selected basins, with the exception of the Little Colorado and Gila are expected to become more vulnerable overtime as the result of projected changes in mean supply and demand. The highest increases are expected in the Lower Colorado, Kansas and Central California, where the average demand is projected to exceed the supply respectively by 2060, 2040 and 2020, leading to vulnerability higher than 0.5. When the effect of standard deviation of supply and demand are considered, on the other hand, we notice that the only two basins where future projections result in progressive vulnerability increases are the Little Colorado and Kansas. However, while in the case of Kansas that is due to the simultaneous increase in variability of both supply and demand, the case of the Little Colorado is driven by the increase in standard deviation of water supply, which offsets the projected decrease in variability of demand.

We use the volume below the response surfaces, calculated as shown in Figure 2.15, as indicator of the susceptibility of the selected basins to changes in PDFs of water supply and demand. We assume that the greater the volume is, the more sensitive the given basin is to future changes is (Weib and Alcamo 2011). In order to use this indicator to directly compare the response of the selected basins to projected changes in supply and demand, the surfaces were created using ranges of $[0.5\cdot\mu_S, 1.5\cdot\mu_S]$, $[0.5\cdot\mu_D, 1.5\cdot\mu_D]$, $[0.5\cdot\sigma_S, 1.5\cdot\sigma_S]$ and $[0.5\cdot\sigma_D, 1.5\cdot\sigma_D]$ for each basin. Volumes below the response surfaces were then normalized by their respective average. As shown in Figure 2.16, the San Joaquin-Tulare basin and the Central California are the

most sensitive basins to changes in current supply and demand, while the Colorado-Gunnison and the Colorado-San Juan are the least sensitive. Normalized sensitivities to changes in mean water supply and demand are very close to the normalized sensitivities to change in the standard deviation of supply and demand for most of the basins. Gila is the basin where the difference between the normalized current sensitivity to changes in mean supply and the normalized sensitivity to changes in standard deviations of demand and supply is larger; in particular, Gila presents a sensitivity to changes in the mean of demand and supply considerably larger than the average of the selected basins (the normalized current sensitivity is about 1.5), while it is as sensitive as the average of the selected basins to changes in standard deviations of demand and supply.

2.8.5 Vulnerability assessment: GCM and scenario dependence

The analysis presented so far is based on the CGCM/A1B hydro-climatic and socio-economic projections. Obviously, different pictures of the future arise when other combinations of GCM and scenarios are used. Analyzing alternative storylines (that is, alternative GCM/scenarios projections) is a way to measure the level of uncertainty that characterizes those portraits of the future.

Composite maps of the maximum and minimum values of vulnerability from among the projections by the CSIRO, and CGCM models under scenarios A1B, A2 and B2 show that although there is general agreement that the water supply system of the southwestern US is the most vulnerable to hydro-climatic variability and socio-economic changes, there is also a great deal of disagreement on the magnitude of that vulnerability, as observed in Figure 2.20 for 2060. The disagreement is greatest in the central and southern Great Plains, the Rio Grande basin, the lower Colorado River basin, the San Joaquin river basin in California, and southern Idaho.

Normalized volumes below the response surfaces for future target year and six GCM-SRES scenario combinations are presented in Figure 2.21 for each of the selected basins. Each bar in the

figure represents the sum of the volumes below the response curves of vulnerability with respect to changes mean and standard deviation of water supply and demand normalized by the average of the volumes of the ten basins for current conditions. The analysis of Figure 2.21 permits to appreciate that the future response of individual basins varies significantly among future scenarios. Central California and Gila (with the exception of the CGCM/B2 scenario projection for 2020) seem to be the only two cases where the various scenarios are in general agreement, while little general inference can be made for the other basins. In an attempt to capture the future trends of the response of each basin, we calculated the averages across the six GCM-SRES scenarios of the normalized volumes below the response curves, as shown in Figure 2.22a. The figure shows that the San Joaquin-Tulare and the Central California are expected to be the more sensitive basins also in the future. Large increases in sensitivity are also expected for the Niobrara-Platte-Loup and Kansas, while the sensitivity of the Sacramento-Lahontan and the other basins of the Colorado River Basin is not projected to change significantly in the future. The sum of the volumes below response functions across basins for each scenario and each target period is shown in Figure 2.22b. The scenarios that predict the overall larger responses are the CSIRO/B2 and the CGCM/A2, especially for the periods 2040, 2060 and 2080. The lower responses, on the other hand are predicted by the CSIRO/A2 and CGCM/A1B, although the latter is projected to consistently increase of throughout the 21st century.

2.9 Summary and conclusions

In this study we developed a procedure to estimate current and future vulnerability of US water supply to shortage. Vulnerability was evaluated on an annual basis as the probability that the water supply is insufficient to meet the demand by simulating the US water supply system at the spatial level of the 98 ASR that make up the contiguous 48 United States. Current and future water supplies were estimated by evaluating local fresh water yield plus the contribution of water transfers and storage, while demands were estimated based on available data of current

consumptive use and projected climatic and socio-economic scenarios. A sub-set of IPCC SRES climatic and socio-economic scenarios was used to determine the vulnerability of the water supply for the conterminous US.

We determined the current and future vulnerability of the US water supply to shortage as well as the sensitivity of vulnerability to changes in current and future water supply and demand. Results are presented in general terms for the entire US while a deeper analysis is performed for a set of selected basins of the Colorado River Basin, California and the High Plains. The latter set of basins is ranked based on their response to current and future changes in PDF of supply and demand as projected by the set of GCM/scenario combination selected.

In agreement with other large-scale assessments (Hurd, Leary et al. 1999), our findings show that the Southwest and central and southern Great Plains are the more vulnerable areas to future climatic and socio-economic changes. In addition, this analysis adds to that prior work in several ways including an accounting for reservoir storage, trans-basin diversions and routing of water among basins, a more comprehensive effort to project future desired water use, and a probabilistic approach to vulnerability.

Contrary to a prior global scale conclusion (Vörösmarty, Green et al. 2000) and in concert with a recent US study (Roy, L. et al. 2010), we find that future increases in the vulnerability of the US water supply will depend more on changes in water yield than on growth in water demand. This is supported by the fact that water use has leveled off in recent years (Kenny, Barber et al. 2009) as irrigated area in the West has diminished and the efficiency of water withdrawals in nearly all sectors has improved (Brown 2000). Moreover, although climate change are expected to increase water demand, future water use efficiency improvements will mitigate that impact so that overall increases in desired water use are expected to be modest in comparison with the effect of climatic changes on water yield and thus on water supply. The reductions in yield, on the other hand, are driven by temperature (and therefore potential and actual evapotranspiration) increases, especially where precipitation decreases or increases only slightly.

Among the selected basins, the Central California and the San Joaquin-Tulare were found to be the more sensitive to both current and future variability of demand and supply. Large sensitivity is also found for the two selected basins of the High Plains, namely the Niobrara-Platte-Loup and Kansas, while the Lower Colorado was found the sub-basin of the Colorado River Basin with the larger susceptibility to changes in future supply and demand. A general disagreement was found among future scenarios both in terms of vulnerability and response of individual basins.

On the whole, the procedure outlined in this framework offers a versatile and consistent instrument to assess the vulnerability of physical system to changes in inherently variable stressors and can be applied to any environmental and socio-economic vulnerability analysis. In addition, it is the only methodology that accounts for both the probabilistic character of the drivers and allows for explicit inclusion of thresholds.

The findings of this analysis assume no major modifications to the physical structure of US water networks. In addition, in-stream flow requirements and trans-ASR diversions were set constant, thereby ignoring possible future changes in surface water redistribution. Indeed, it is the purpose of this assessment to point to those locations where adaptation (i.e. enlarged trans-basin diversion capacity or, more likely, within-basin water transfers and enhanced water conservation efforts) will be most needed. Because all simulations project a steady decline in the water levels of the main Southwestern reservoirs, indicating that water scarcity is primarily due to demand-supply imbalance rather than to insufficient storage capacity, increasing storage capacity within existing ASR networks does not appear to be a successful adaptation strategy.

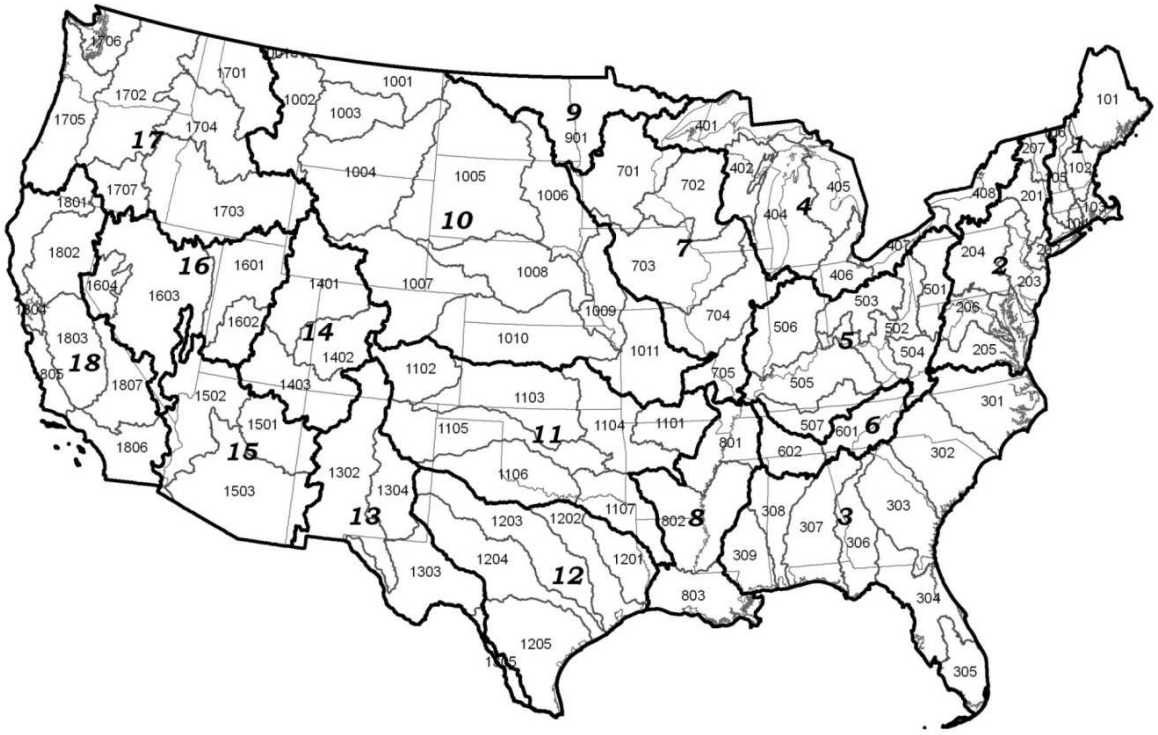


Figure 2.1 Water Resource Regions and Assessment Subregions.

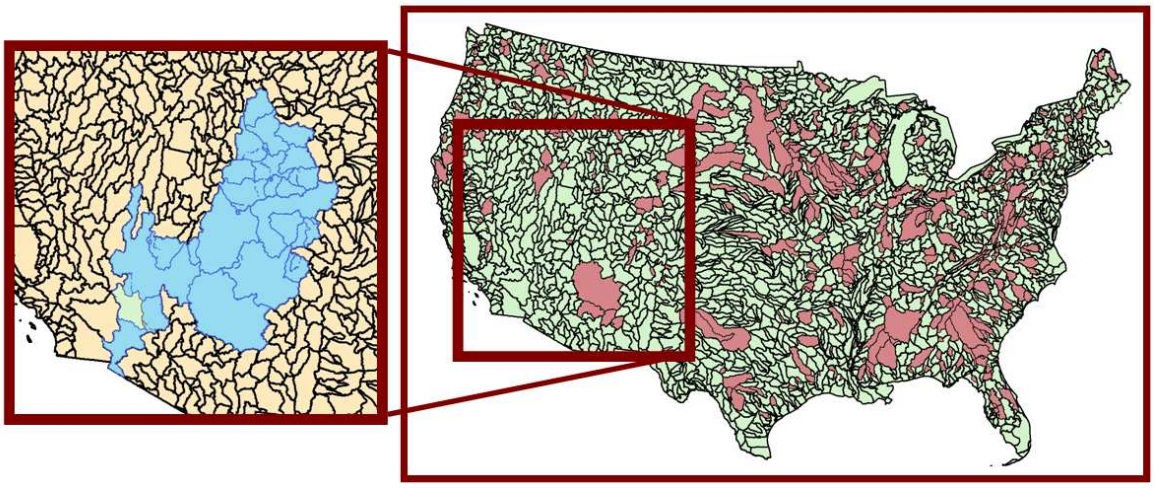


Figure 2.2 Map of the 655 test basins (red), 8-digit basins (green) with zoom of the Colorado River Basin with the watersheds available in USBR records (light blue).

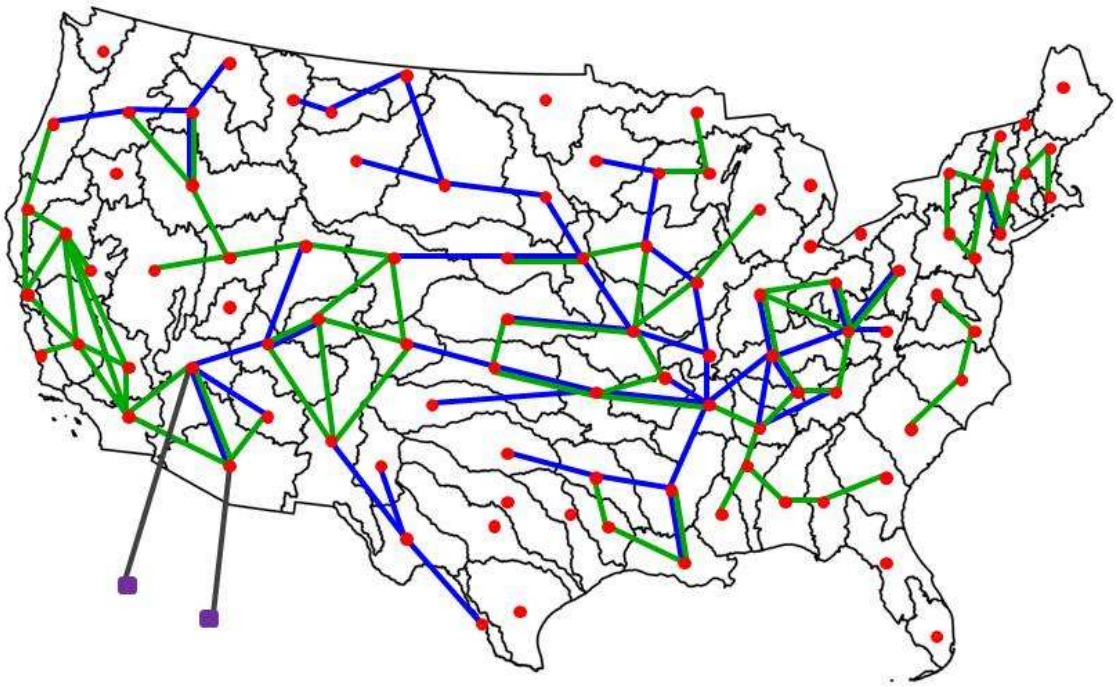


Figure 2.3 Water networks across the conterninous US at the ASR level. Natural links are indicated with blue lines, artificial links (water diversions) with green lines. Gray lines indicate diversions to Mexico.

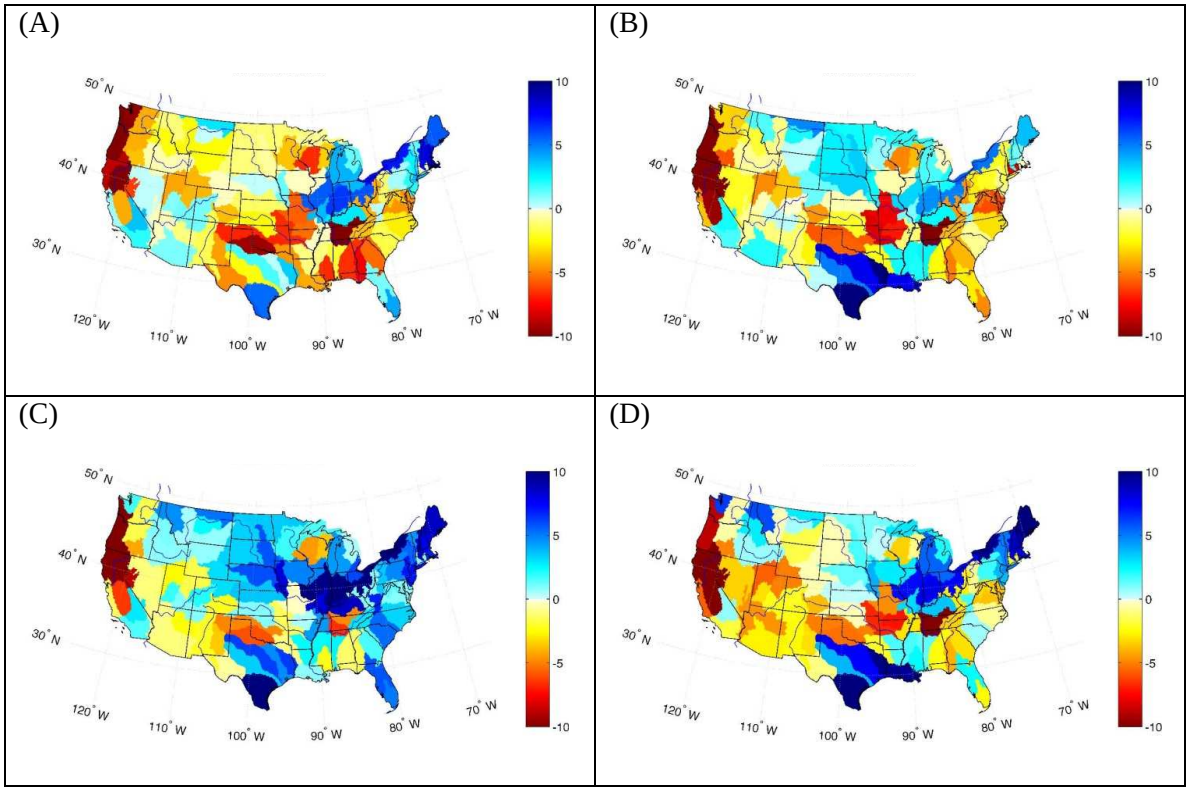


Figure 2.4 Changes in mean precipitation (in cm) projected by the CGCM/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

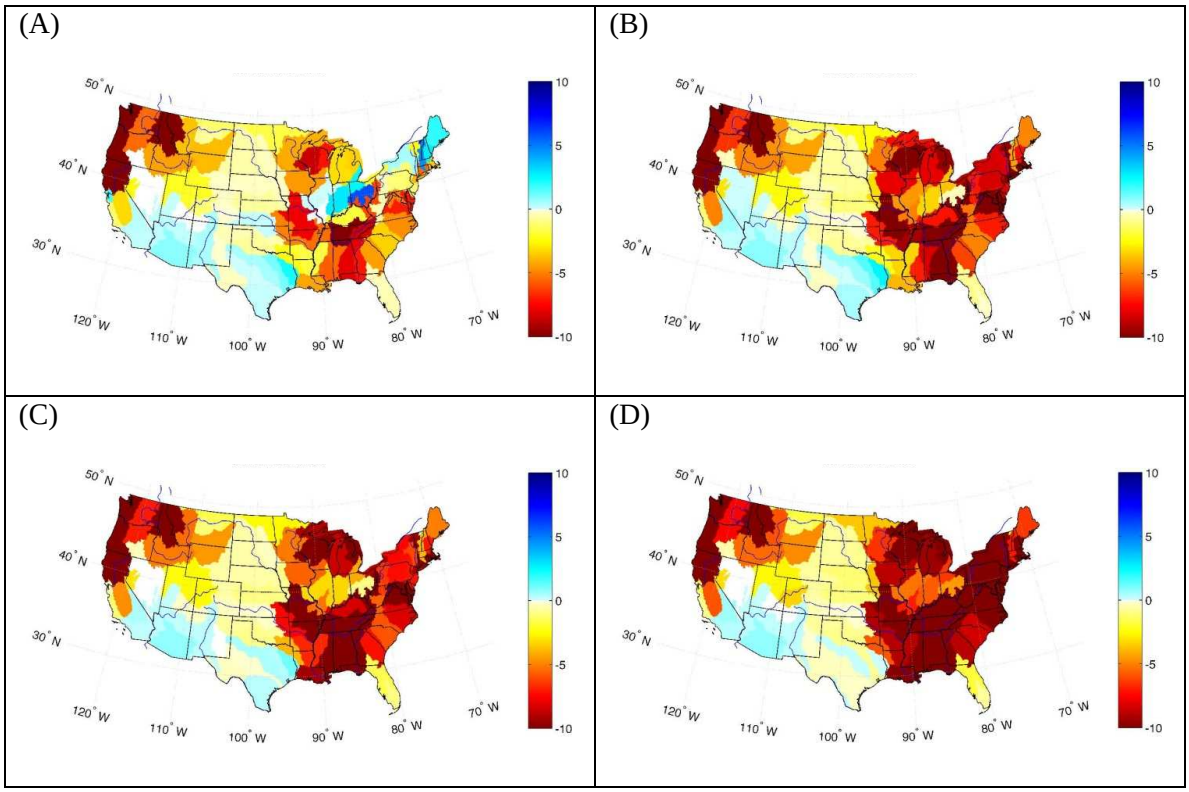


Figure 2.5 Changes in mean water yield (in cm) for the CGCM/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

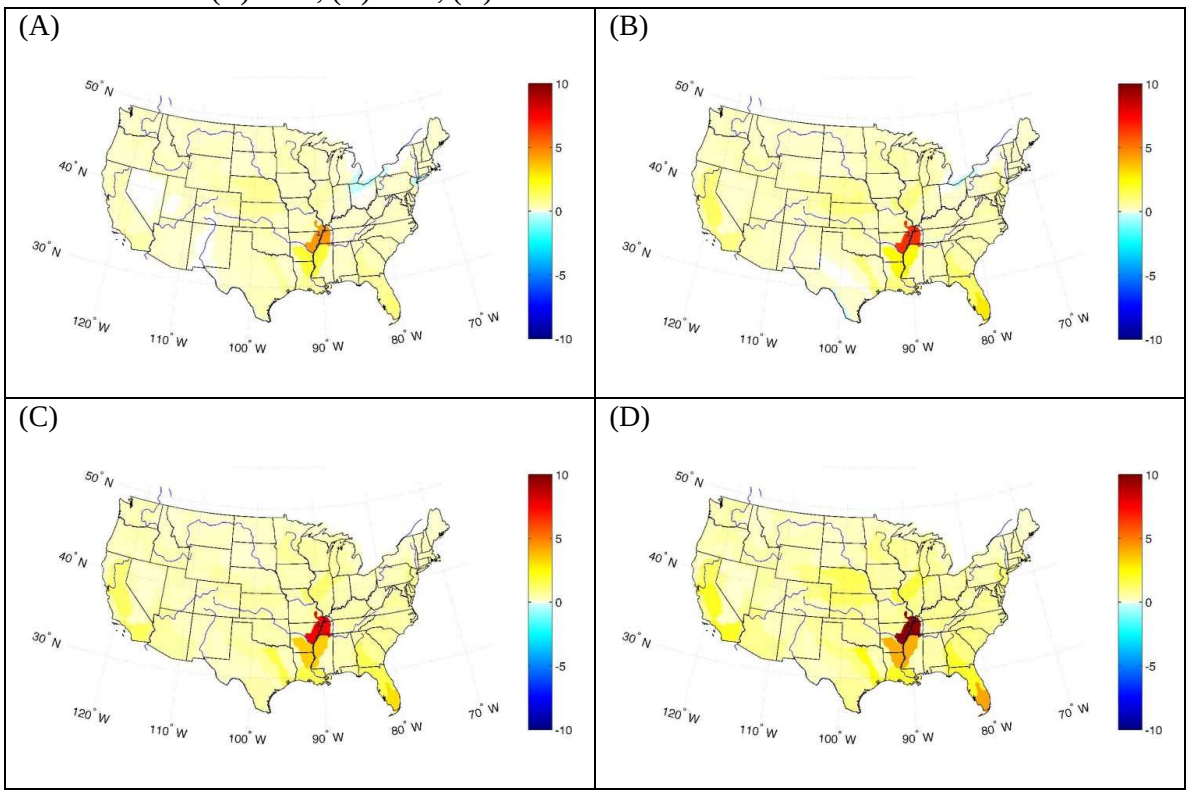


Figure 2.6 Changes in mean water demand (in cm) for the CGCM/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

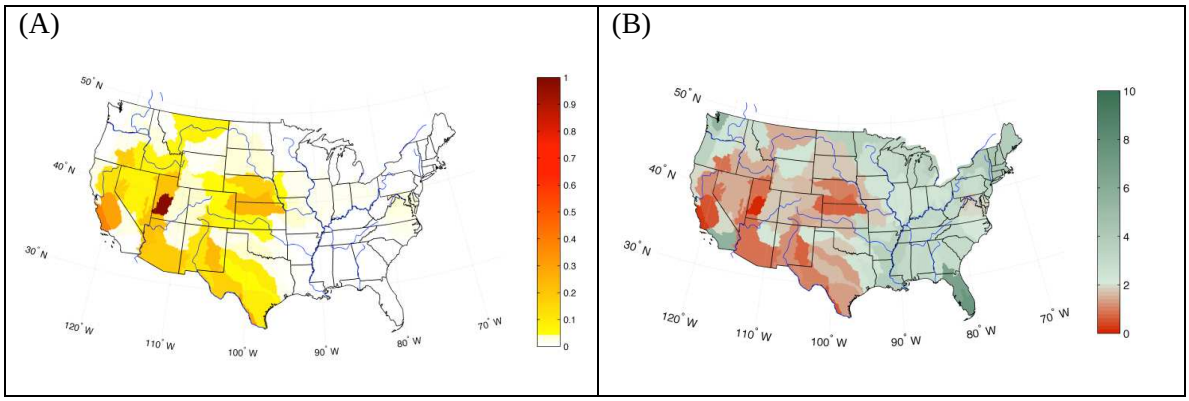


Figure 2.7 Current vulnerability (A) and reliability index (B).

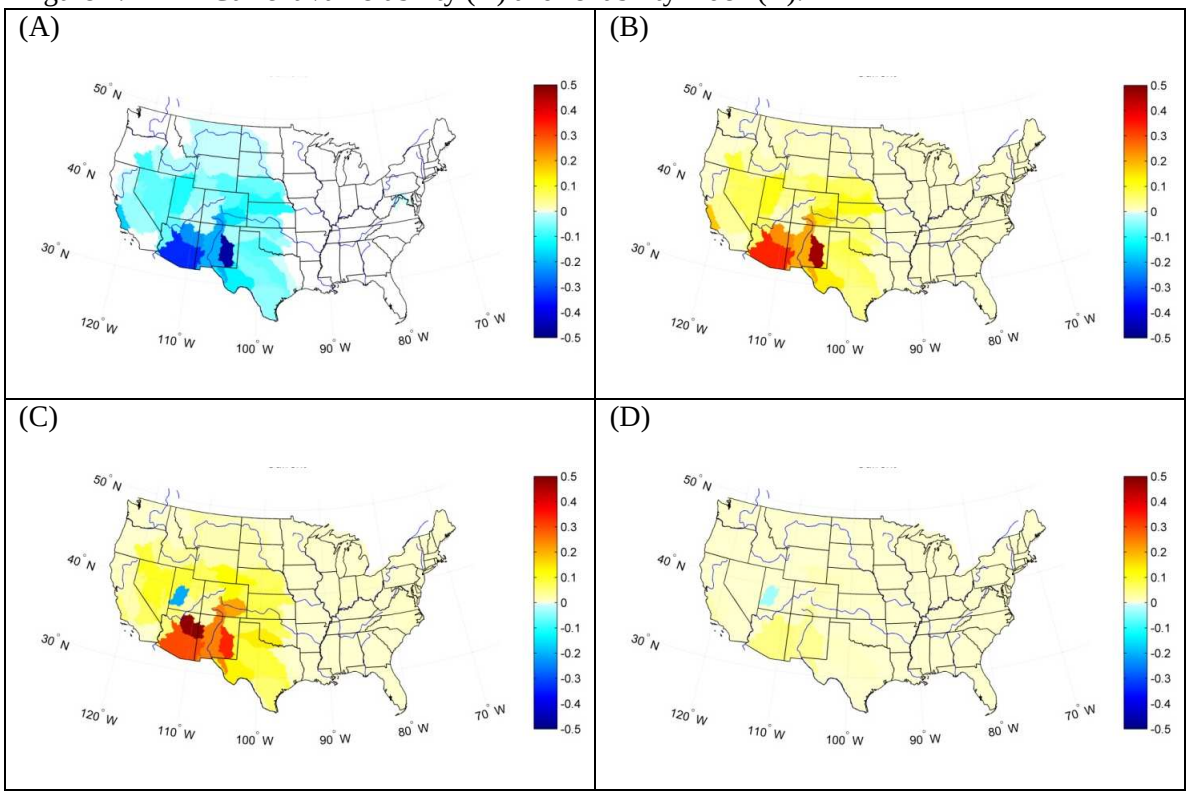


Figure 2.8 Current sensitivity (in cm^{-1}) of vulnerability to unit changes in: (A) mean water supply; (B) mean water demand; (C) standard deviation of water supply; (D) standard deviation of water demand; for the CGCM/A1B scenario.

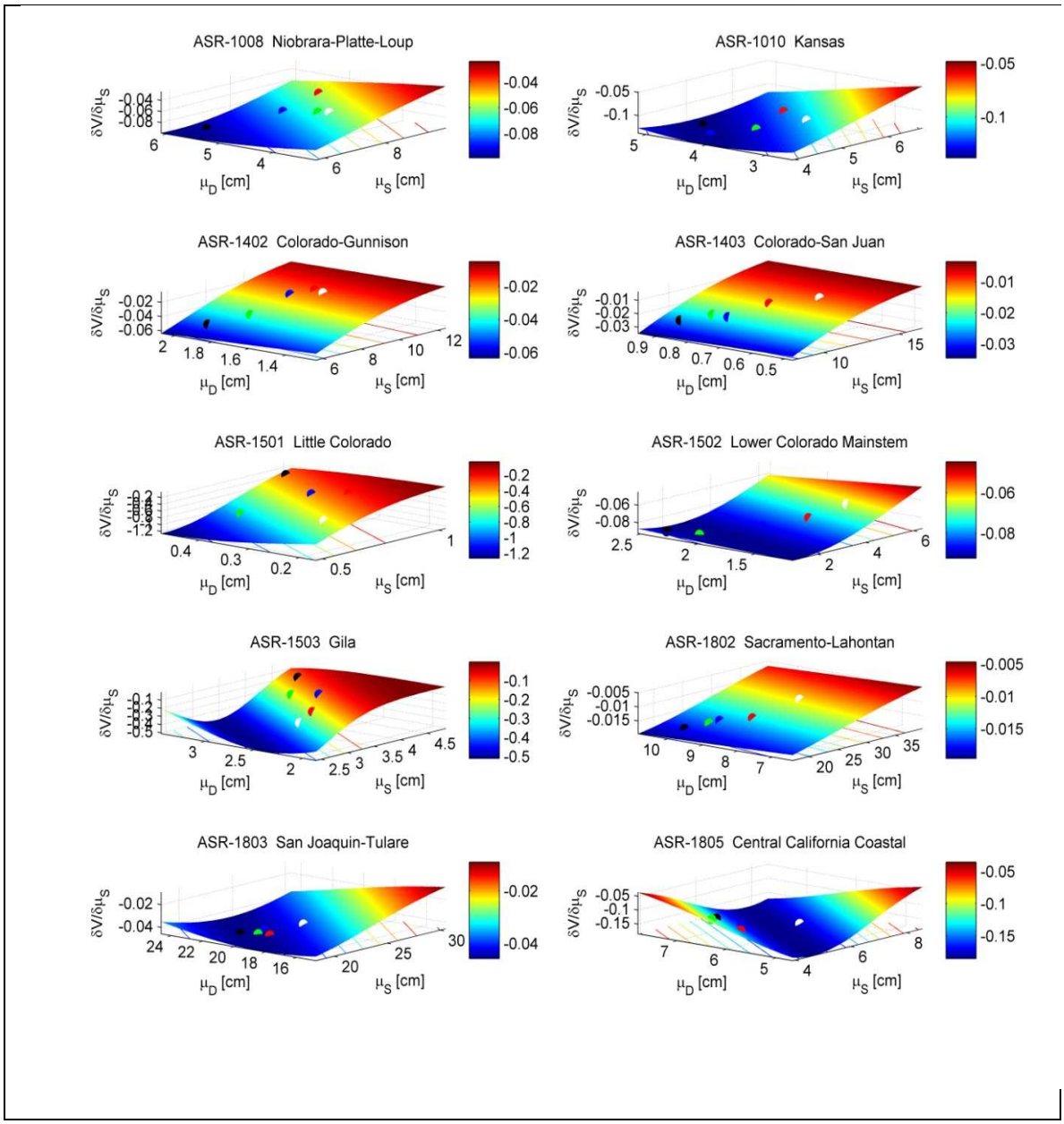


Figure 2.9 Response surfaces of sensitivity of vulnerability to changes in mean water supply for selected basins of the High Plains, the Colorado River Basins and California. Current status is represented for each surface by the white marker. Red, blue, green and black markers indicate respectively the 2020, 2040, 2060 and 2080 periods. Maps are relative to the CGCM/A1B scenario.

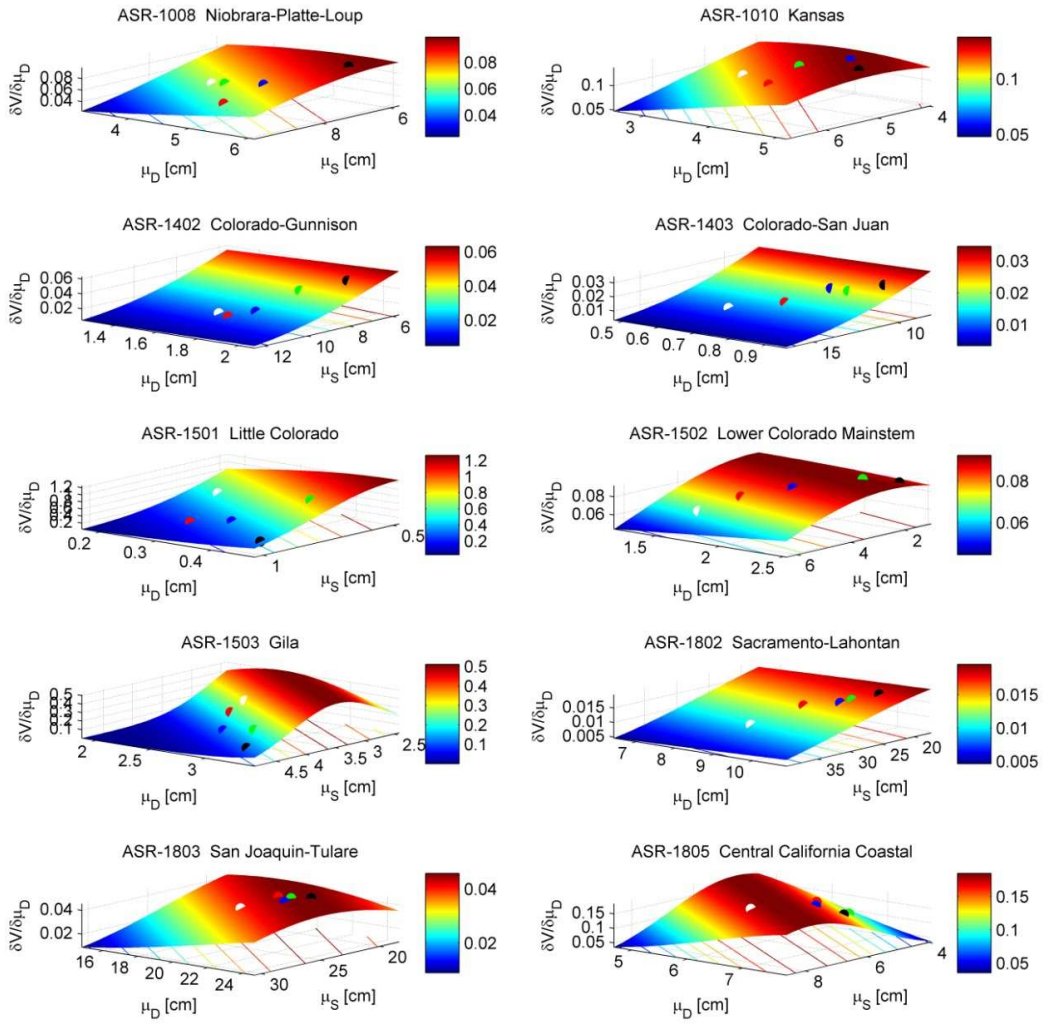


Figure 2.10 Response surfaces of sensitivity of vulnerability to changes in mean water demand for selected basins of the High Plains, the Colorado River Basins and California. Current status is represented for each surface by the white marker. Red, blue, green and black markers indicate respectively the 2020, 2040, 2060 and 2080 periods. Maps are relative to the CGCM/A1B scenario.

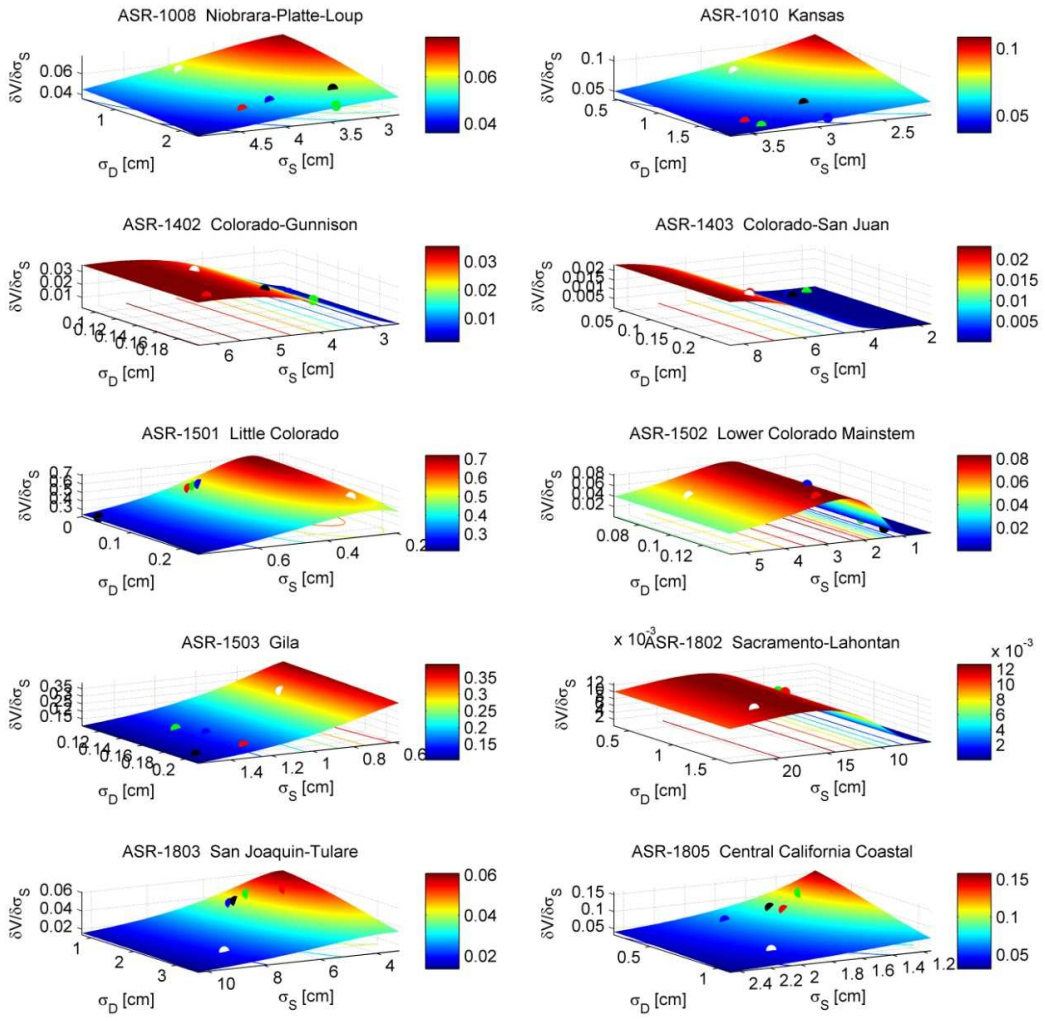


Figure 2.11 Response surfaces of sensitivity of vulnerability to changes in standard deviation of water supply for selected basins of the High Plains, the Colorado River Basins and California. Current status is represented for each surface by the white marker. Red, blue, green and black markers indicate respectively the 2020, 2040, 2060 and 2080 periods. Maps are relative to the CGCM/A1B scenario.

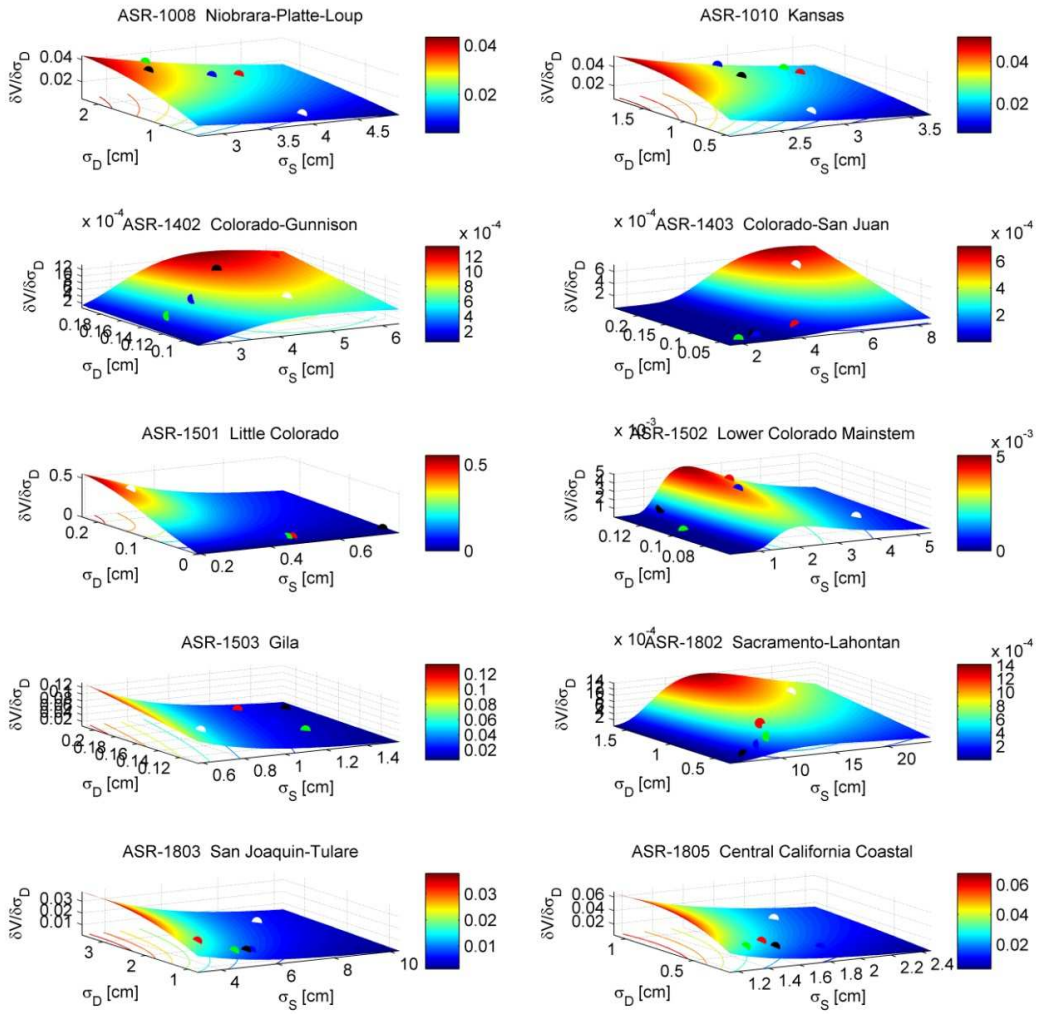


Figure 2.12 Response surfaces of sensitivity of vulnerability to changes in standard deviation of water demand for selected basins of the High Plains, the Colorado River Basins and California. Current status is represented for each surface by the white marker. Red, blue, green and black markers indicate respectively the 2020, 2040, 2060 and 2080 periods. Maps are relative to the CGCM/A1B scenario.

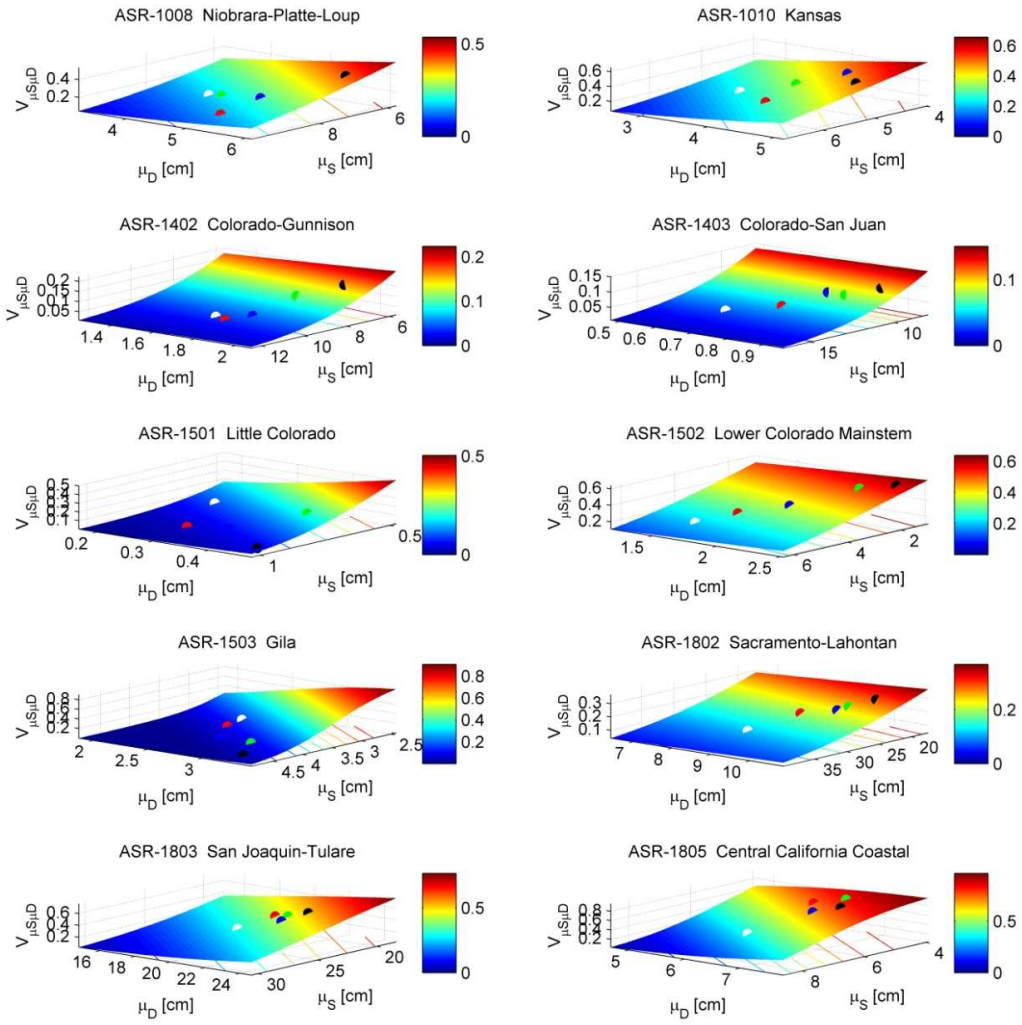


Figure 2.13 Response surfaces of vulnerability as function of mean water supply and water demand for selected basins of the High Plains, the Colorado River Basins and California. Current status is represented for each surface by the white marker. Red, blue, green and black markers indicate respectively the 2020, 2040, 2060 and 2080 periods. Maps are relative to the CGCM/A1B scenario.

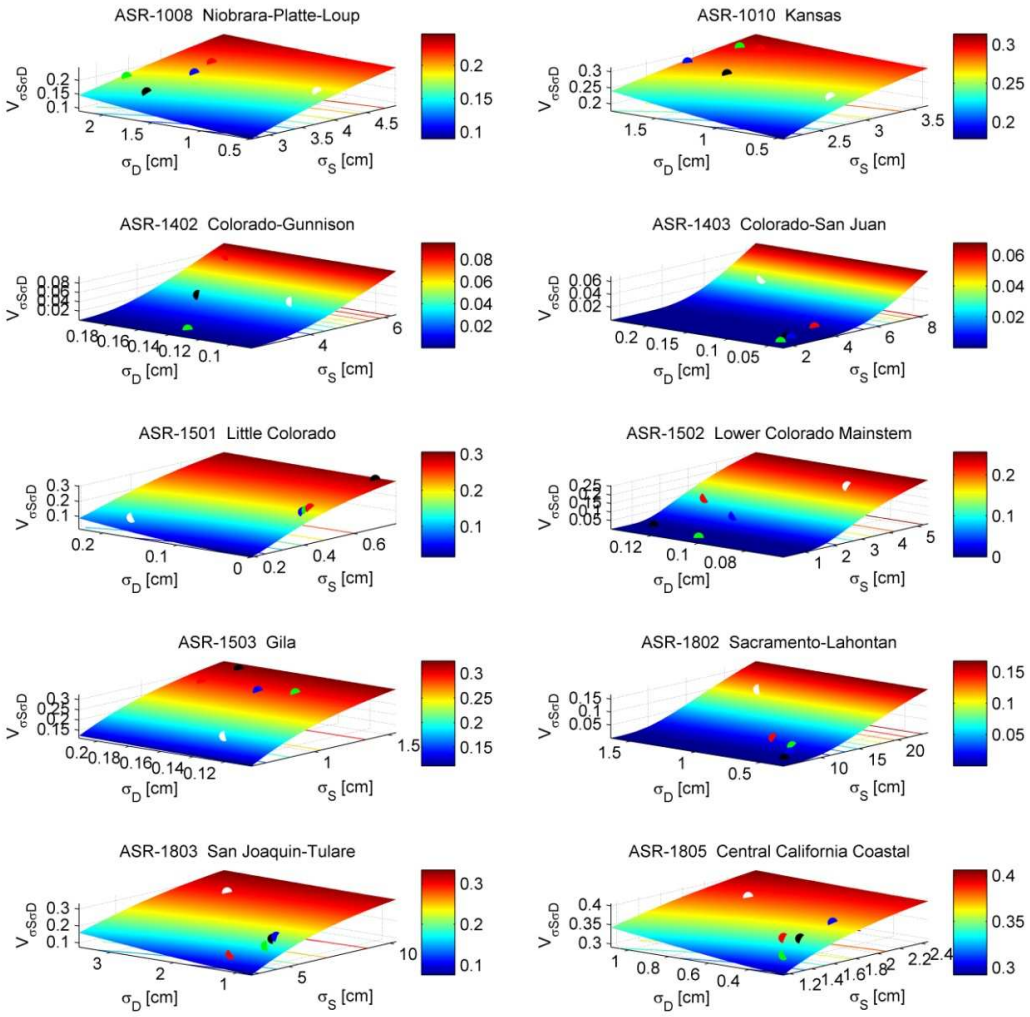


Figure 2.14 Response surfaces of vulnerability as function of standard deviation of water supply and water demand for selected basins of the High Plains, the Colorado River Basins and California. Current status is represented for each surface by the white marker. Red, blue, green and black markers indicate respectively the 2020, 2040, 2060 and 2080 periods. Maps are relative to the CGCM/A1B scenario.

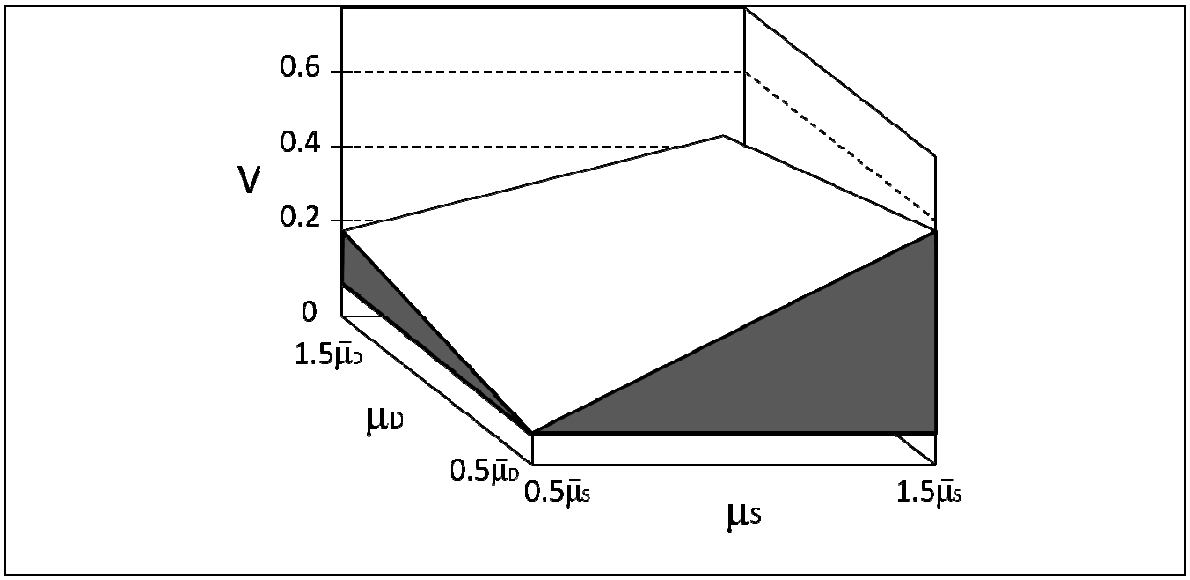


Figure 2.15 Volume below the response surface (in grey) as indicator of the basin response to changes in supply and demand.

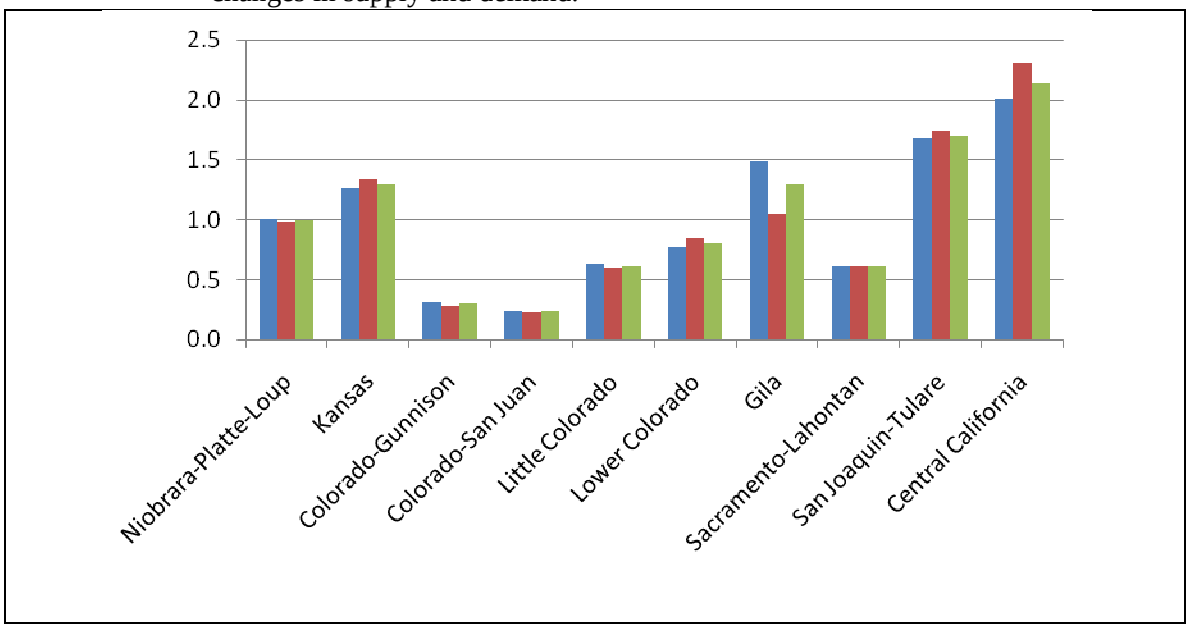


Figure 2.16 Normalized volumes under the response surfaces of vulnerability for selected basins of the High Plains, the Colorado River Basins and California for the current conditions. Blue histograms correspond to the volume under the response surfaces of vulnerability changes as function of mean supply and demand. Red histograms correspond to the volume under the response surfaces of vulnerability changes as function of standard deviation of supply and demand. Green histogram corresponds to the sum of the previous two volumes.

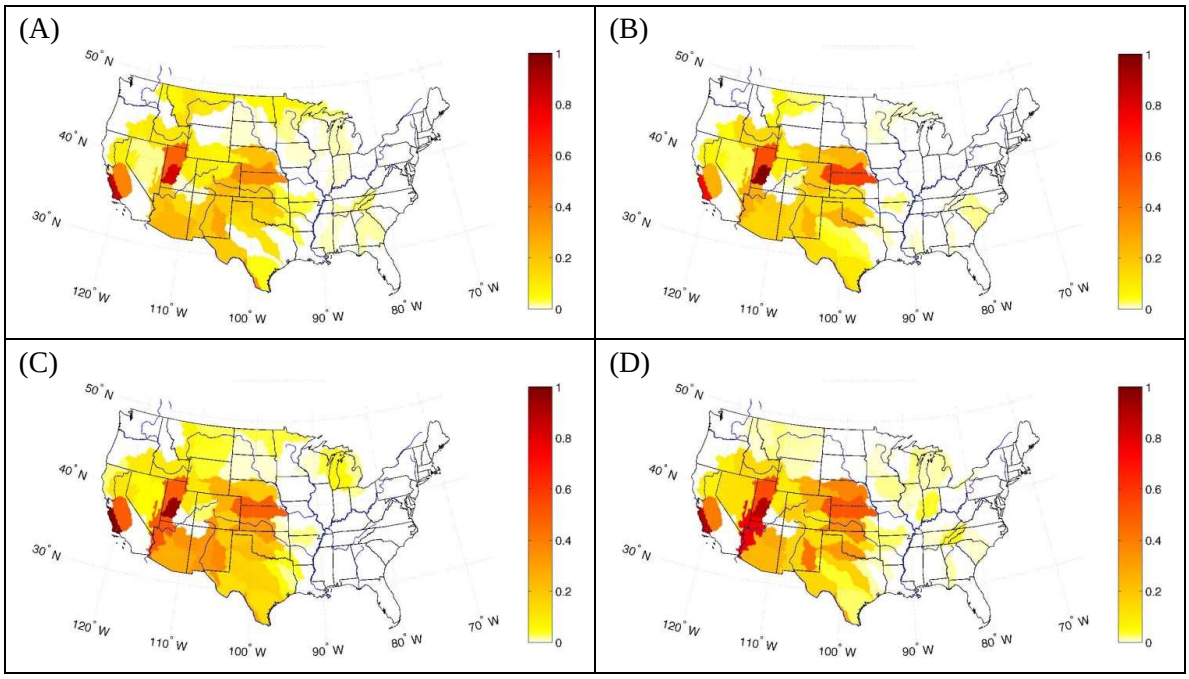


Figure 2.17 Future vulnerability for the CGCM/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

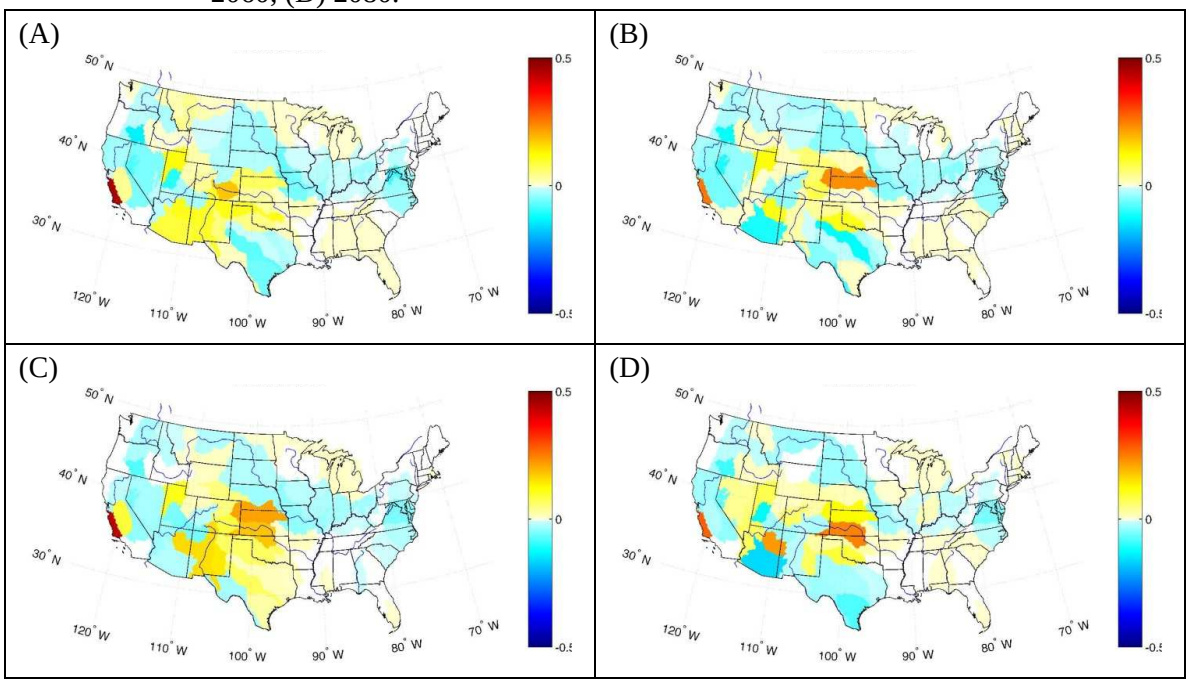


Figure 2.18 Changes in vulnerability for the CGCM/A1B scenario due to changes in average water supply and standard deviation of water supply for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

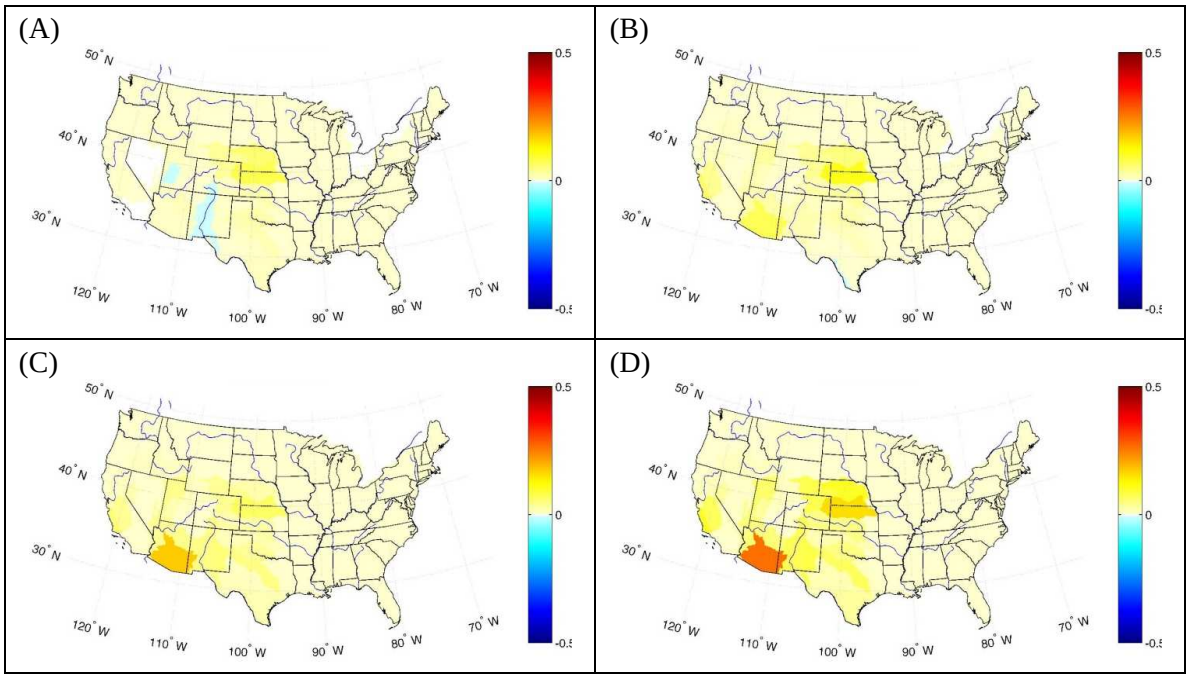


Figure 2.19 Changes in vulnerability for the CGCM/A1B scenario due to changes in average water demand and standard deviation of water demand for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

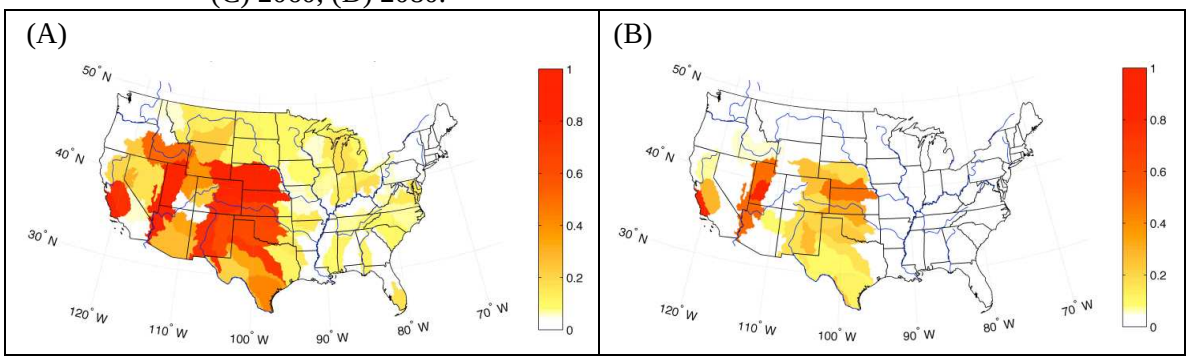
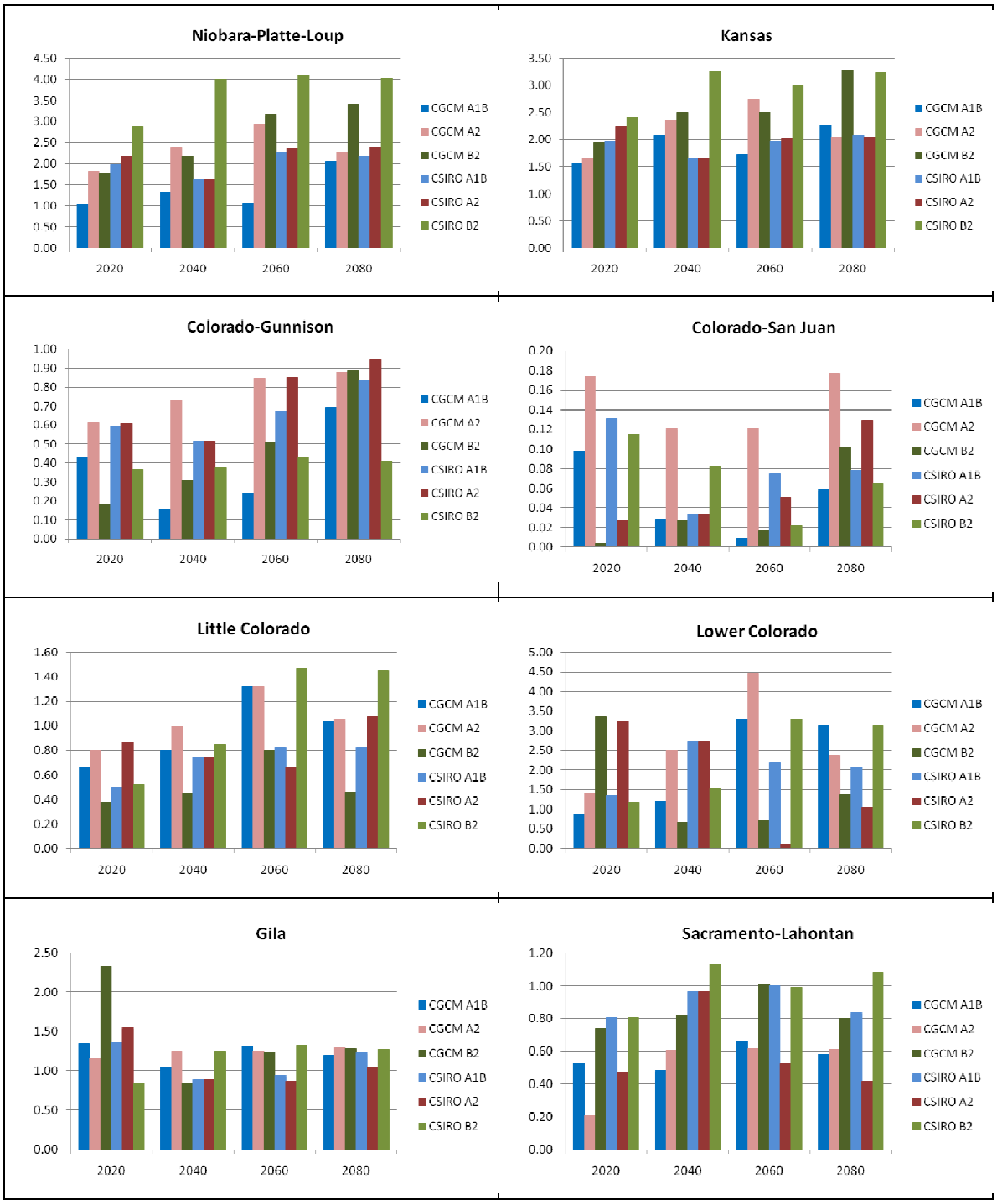


Figure 2.20 Composite for 2060 of maximum (A) and minimum (B) vulnerability projected by six GCM/SRES-scenario combinations.


Niobara-Platte-Loup
Kansas
Colorado-Gunnison
Colorado-San Juan
Little Colorado
Lower Colorado
Gila
Sacramento-Lahontan
CGCM A1B
CGCM A2
CGCM B2
CSIRO A1B
CSIRO A2
CSIRO B2
2020
2040
2060
2080

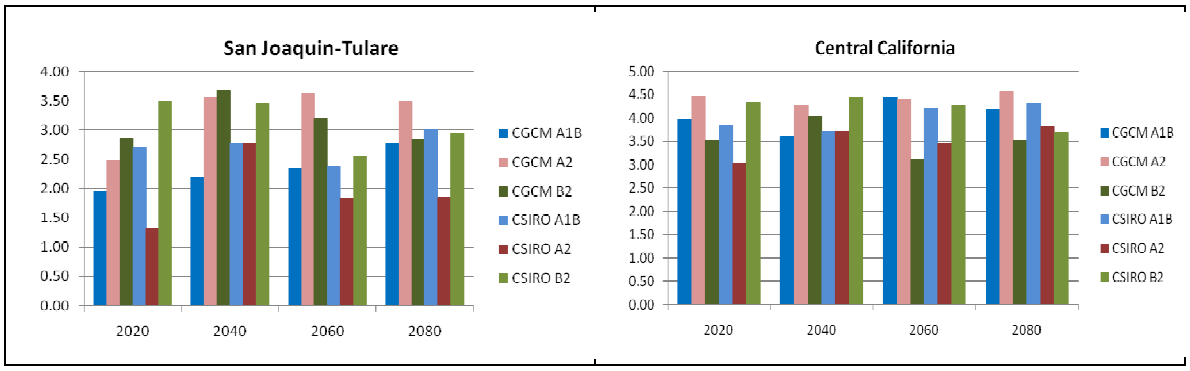


Figure 2.21 Future normalized volumes under the response surfaces for selected basins of the High Plains, the Colorado River Basins and California for six GCM/SRES-scenario combinations.

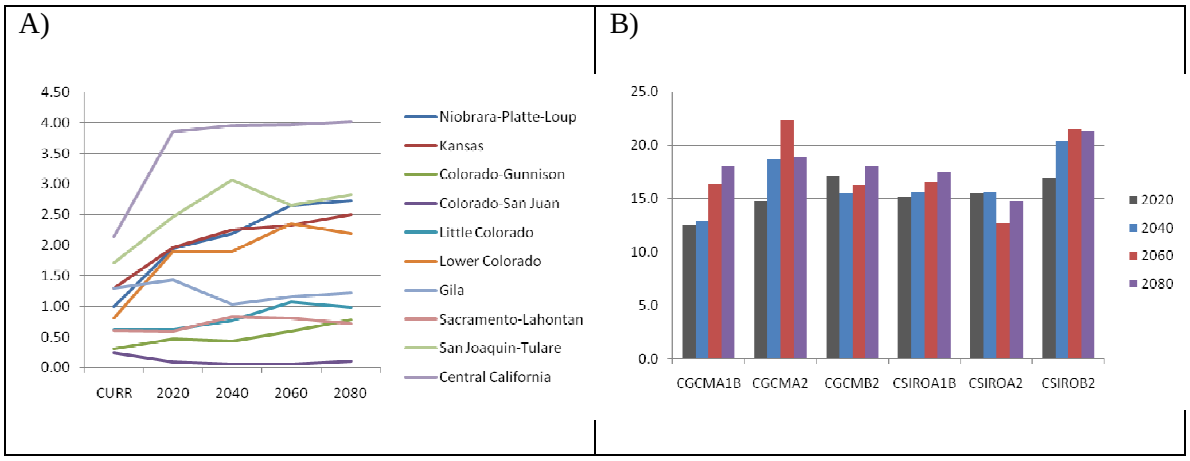


Figure 2.22 A) Future normalized volumes under the response surfaces for selected basins of the High Plains, the Colorado River Basins and California averaged across the six GCM/SRES-scenario combinations. B) Sum of the normalized volumes under the response surfaces across basins and for each target period and each of the six scenarios.

Chapter 3 A mechanistic approach for the description of formation and evolution of vegetation patterns

3.1 Abstract

Vegetation patterns are a common and well-defined characteristic of many arid and semi-arid landscapes. In this paper we explore some of the physical mechanisms responsible for the establishment of self-organized, non-random vegetation patterns that arise, at the hillslope scale, in many areas of the world, especially in arid and semi-arid regions. In doing so, we use a water balance model and provide a fundamental mechanistic understanding of the dynamics of vegetation pattern formation and development. Reciprocal effects of vegetation on the hillslope thermodynamics, runoff production and run-on infiltration, root density, surface albedo and soil moisture content are analyzed. In particular, we: 1) present a physically based mechanistic description of the processes leading to vegetation pattern formation; 2) quantify the relative impact of each process on pattern formation; and 3) describe the relationships between vegetation patterns and the climatic, hydraulic and topographic characteristics of the system.

3.2 Introduction

The presence of self-organized vegetation patterns is a common and well-defined characteristic of many dry landscapes. Indeed, vegetation is in general spatially heterogeneous and its constituent species show spatial distributions that depart from randomness (Greig-Smith 1979), although only in few cases, where this departure is more marked, the pattern structure is easily recognizable, Figure 3.1.

Although in many circumstances the presence of vegetation patterns is detectable from the ground, the advent of aerial photography was needed to give a more comprehensive view of the multitude of shapes and scales that the phenomenon exhibits: banded, spotted or labyrinthine vegetation patterns are not uncommon in many arid to semi-desert areas and can develop at a wide variety of spatial scales. Typical dimensions of vegetation pattern element (i.e. thickness of a band or radius of a patch of vegetation) can span up to two orders of magnitude, ranging from 10^0 to almost 10^2 m (Rietkerk and Van de Koppel 2008).

The identification and characterization of the phenomenon, as well as the individuation of the processes responsible for specific types of these patterns (e.g., *so-called* tiger bushes) were the main focus of numerous studies (Worral 1959; Greig-Smith 1979; Thiéry, d'Herbès et al. 1995; Dunkerley and Brown 1999).

Initially, studies were mainly focused on qualitative descriptions and on identifying and listing the recurrence of certain types of spatial configurations (Worral 1960; Boaler and Hodge 1962). However, during the last decades, research has been directed to a more quantitative characterization (Lefever and Lejeune 1997; D'Odorico, Laio et al. 2006). Most of the studies that attempted a quantitative description of the processes leading to the formation of vegetation patterns agree in taking a "synergy versus competition" approach to the problem (Valentin, d'Herbès et al. 1999; D'Odorico, Laio et al. 2006). In other words, they advocate the idea that the development of non-random self-organized configurations is the result of short-range synergy and long-range competition occurring between plants and groups of plants. According to those studies, therefore, those spatial interactions are responsible for inducing the system to drift away from a spatial configuration characterized by randomly distributed vegetation, promoting the formation of non random vegetation structures.

Plants, especially in arid landscapes, help reduce soil erosion and augment soil permeability; they also protect each other from winds and damage due to animals and extreme temperatures and humidity conditions. In some areas, those factors encourage the formation of bands of vegetation

in mild hillslopes (Bromley, Brouwer et al. 1997) or favor the development of the same kind of pattern in a direction perpendicular to the one of the prevalent winds, in response to their erosive action (Leprun 1999). Although hillslope-scale patterns can arise in a variety of regions and climates, scarcity of water seems to be the common denominator of every landscape characterized by vegetation patterns. For the development of patterns, in fact, it appears to be crucial for the system to be water limited and not able to support a permanent and stable configuration of complete canopy closure, as underlined in the literature mentioned above.

Although considerable efforts were made towards a more detailed representation of the elements playing a role in pattern formation by including both deterministic and stochastic aspects, most of the models proposed to describe the phenomenon belong to three categories: 1) kernel based models (Thiéry, d'Herbès et al. 1995; Lefever and Lejeune 1997; D'Odorico, Laio et al. 2006); 2) advection-diffusion models (HilleRisLambers, Rietkerk et al. 2001; Rietkerk, Boerlijst et al. 2002); 3) differential flow instability models (Klausmeier 1999; Sherrat 2005; Saco, Willgoose et al. 2007). Although they succeed in building a conceptual mathematical framework able to describe the dynamics of vegetation propagation, soil moisture distribution and local vegetation interactions, they do not provide a pure mechanistic representation of those dynamics.

The overarching objective of this study is to explore and identify (some of) the physical mechanisms responsible for the establishment of non-random spatial vegetation patterns that arise, at the hillslope scale, in many areas of the world, especially in arid and semi-arid regions.

The specific objectives are to: 1) develop a physically based mechanistic modeling of the processes leading to vegetation pattern formation; 2) implement such modeling within a framework able to replicate the main physical characteristics of observed vegetation patterns (shapes, dimensions, etc.); 3) individuate the respective impact of each process on pattern formation; and 4) capture the relationships between vegetation patterns and the climatic, hydraulic and topographic characteristic of the system.

3.3 Hypotheses

Vegetation patterns refer to the relative non-random arrangement of vegetated and bare patches of soil on the landscape, where non-randomness indicates any spatial distribution of patches that deviates from a purely spatially random distribution.

Our first hypothesis is that vegetation patterns emerge as a result of reciprocal actions, physical, chemical and physiological, (i.e. feedback processes) between vegetation, hydrologic and climatic processes, and soil properties, and that those feedback processes are amenable to quantitative description and modeling.

Our second hypothesis is that patterns develop because the physical processes in action tend to make certain regions in the neighborhood of an existing clump of vegetation more conducive to the establishment of additional vegetation (or not).

Finally, our third hypothesis is that the spatial distribution of vegetation depends on the spatial distribution of water and energy. Therefore, physiological and hydrological processes conducive to local decreases in the available water and nutrients will tend to inhibit vegetation establishment and those conducive to locally maintaining or increasing the water and energy will tend to promote vegetation establishment.

3.4 Scale of analysis

3.4.1 Spatial scale

Vegetation patterns analysis is scale-dependent. The individuation of spatial structures, as well as their geometric characterization, in fact, depends on the spatial resolution at which we observe the area under analysis. At the finest resolution, a given area of study is either vegetated or bare. As the resolution of the observations increases, however, one will be able to notice that, over the same hillslope, vegetation is indeed distributed unevenly (aside from the extreme cases of fully vegetated or bare hillslopes), showing higher density in some areas and lower density in

others. Thus, the spatial scale of our analysis must be compatible with the scale at which the phenomenon we want to analyze takes place.

In this study we are interested in analyzing macroscopic vegetation agglomerates whose typical dimensions are of the order of magnitude of 10^0 to 10^1 m (Figure 3.2). Hence, we need to be able to characterize spatial structures of that size over a spatial domain representing a hillslope. To this aim, we will further subdivide the study domain (i.e. our hillslope of study) into a study subgrid whose pixels will be small enough to resolve the spatial configuration of the patterns we intend to analyze. The areas of the whole domain and of a single pixel of our study subgrid measure 10^5-10^7 m^2 and 10^0-10^2 m^2, respectively (Figure 3.3).

3.4.2 Temporal scale

The dynamics of the water and energy fluxes and of the spatial interactions between vegetation, soil and climate take place at various time scales. While, on the one hand, water and energy fluxes show large fluctuations on a daily or subdaily basis, on the other hand, the time scale of vegetation pattern formation and dynamics is much longer. In particular, for a given set of environmental forcings (climate and soil), the amount of vegetation that is ultimately established in a certain area and its spatial configuration are the result of a long-term process of mutual adaptation. Vegetation patterns, in fact, show low variability over time (i.e., statistical properties characteristic of patterns do not change drastically within a year and from year to year), suggesting that patterns are themselves more sensitive to the long-term average characteristics of the climate-soil system rather that to short-term disturbances. Under the assumption of stationary climate, therefore, we will use long-term average climatic and hydraulic conditions, in order to determine the spatial configurations of vegetation and associated water and energy fluxes that are in long-term equilibrium with the climate.

3.5 Methods

We simulate the vegetation-soil-climate dynamics on a two-dimensional gridded domain using a model that quantifies the water budget along with the net absorbed solar radiation and latent heat of evapotranspiration and that, at the same time, is able to reproduce the effect of spatial interactions between vegetation in neighboring pixels.

Water and energy fluxes occur mainly in the vertical direction, across the interface between soil and atmosphere. Spatial interactions between vegetation groups, and mutual interactions between soil, water fluxes, and vegetation, on the other hand, mostly occur in the horizontal directions. Therefore, the description of these coupled processes would ideally require a 3D dynamical model. However, in this work, we model these 3D interactions by means of a combination of a 1D water balance model able to describe the vertical fluxes coupled to a 2D model able to capture the spatial interactions. This approach is predicated on the assumption that the iterative sequential application of these two models will be substantially equivalent, for our purposes, to simulating the system using a fully 3D dynamical model.

3.5.1 Procedure schematization

We model the soil-climate-vegetation system by characterizing the spatial configuration of water fluxes and vegetation density of a hillslope. Our spatial domain represents a hillslope of given topography, further subdivided into a subgrid of interconnected NxN pixels in order to capture the spatial variability of both fluxes and vegetation coverage.

Because the long-term average vegetation density at a certain location in space is the long-term response of the climate-soil-vegetation system to a set of environmental forcings, knowledge of the spatial configuration of the environmental forcings over a certain domain can be used to determine the spatial configuration of vegetation density over the same domain. Therefore, the objective of our modeling is to determine a spatial configuration of fluxes and vegetal density that simultaneously satisfies the water budgets both at the global (i.e. for the entire study domain) and

the local (i.e. for each pixel) scales, while taking into account lateral interactions (i.e. between adjacent pixels) between vegetation, climate and soil. However, while the water and energy fluxes may be considered known at the hillslope scale, at the pixel level they are unknown, because they are sensitive to and ultimately depend on the vegetal density at the local scale (i.e., present at each pixel).

In order to determine the pixel-scale fluxes and vegetation density that are in equilibrium with the hillslope-scale conditions, we use an iterative simulation procedure. For a given set of initial climatic conditions and soil properties, long term averages of annual fluxes of water, solar radiation and latent heat of evapotranspiration, as well as of vegetal density, are estimated at the pixel level for each one of the pixels of the study domain by means of a water balance model. Once the long-term annual average vertical fluxes and vegetal density are estimated at the pixel level and for the generic iteration n, the mutual lateral effects of vegetation and fluxes on adjacent pixels of the study grid are evaluated. In turn, those lateral interactions will have the net effect of modifying the forcings and the hydraulic parameters of the system at each pixel. Evaluation of lateral effects, therefore, allows the estimation of an updated set of climatic inputs and soil parameters that are used to perform the subsequent (iteration n+1) water budget at each pixel. A flow chart of the simulation procedure is provided in Figure 3.4.

As shown in the procedure schematization outlined in Figure 3.4, the simulation procedure is preceded by three preliminary steps (qualitatively described in Figure 3.5): 1) evaluation of the hillslope-scale vegetal coverage or vegetation density (obtained by solving the long-term average water balance at each pixel for the same climatological forcings and hydraulic properties); 2) application of an initial random perturbation to the vegetal density of each pixel of the study domain and 3) evaluation of the effect of the lateral interactions from the perturbed vegetation configuration. The reason for these preliminary steps is simple: in accordance with the hypotheses of this work, vegetation patterns are the result of spatial mechanisms of facilitation and inhibition between adjacent groups of vegetation. Therefore, two adjacent pixels of our domain must be

characterized by different vegetal density for their interaction to have a net effect on each other. Thus, the application of a random perturbation is intended to reproduce the small deviance from the uniform spatial configuration of water fluxes and vegetation density that is necessary to trigger the lateral interactions responsible for the formation of patterns. It is worth underlying that this random perturbation is applied only at the initial simulation step.

3.5.2 1D Water budget – Vertical fluxes and forcings

The water budget at any pixel of our study domain is quantified using Eagleson's annual water balance model, a one-dimensional representation of soil moisture dynamics as forced by a stochastic climate (Eagleson, 1978a-g). It describes the relationship between annual amounts of precipitation, runoff, infiltration and evapotranspiration as a function of volumetric soil moisture and soil and vegetation characteristics (see Appendix A for details).

Soil hydraulic properties are characterized by the following parameters: total porosity, pore size distribution index, surface retention capacity, saturated hydraulic conductivity and matric potential at effective saturation.

Climatic characteristics are described in terms of the mean storm duration, mean time between storms, rainy season length, mean and variance of storm depth, mean annual precipitation and mean annual potential evapotranspiration. Vegetation is characterized by the plant transpiration efficiency, that is, the ratio between the potential rate of evaporation and the potential rate of transpiration and by the fractional vegetation cover or vegetation density. The model predicts the long-term averages of the water fluxes as well as the long-term soil moisture content (over a semi-infinite soil column) and vegetation fractional coverage. The model is based on the following main assumptions:

- the soil-vegetation system is in a long term equilibrium with the climate;
- the value of long-term soil moisture content under which the balance is closed maximizes vegetal biomass production under conditions of minimal vegetation stress.

3.5.3 2D Spatial feedbacks characterization - Horizontal fluxes and interactions

Our main hypothesis is that vegetation at a point (i.e., at each pixel) affects water and energy fluxes in its neighborhood (i.e., pixel and surrounding pixels). Plants, in fact, affect the physical structure of the system by perturbing the thermal and aerodynamic properties of the canopy layer as well as the soil structure (i.e., texture, porosity, connectivity, hydraulic conductivity, etc.). As a result of those perturbations, water and energy fluxes change in the neighborhood of the plant, modifying the environmental conditions in a way that can promote or inhibit the establishment of surrounding vegetation.

The climate-soil-vegetation system is very complex and governed by strong feedbacks between all elements of the system. In the context of this paper, most, if not all, of the forcing and physical characteristics governing the dynamics of the climate-soil-vegetation system are in affected by the spatial configuration of the vegetation. However, we focus only on a subset of factors that we hypothesize are the main drivers of the process of vegetation patterns formation and evolution. These factors are: 1) modification of the spatial distribution of soil hydraulic conductivity by vegetation, 2) infiltration of surface runoff, a phenomenon known as run-on infiltration, 3) spatial reconfiguration of soil albedo, 4) spatial soil moisture redistribution due to roots, and 5) redistribution of nutrients and available energy for evapotranspiration.

Although fire, livestock, and other such external forces may be the main cause determining vegetation patterns in some instances, vegetation patterns as those shown in Figure 3.1 are observed even in the absence of such forces. Therefore, our work focuses on the feedbacks and interactions between vegetation, soil, and hydro-climatic processes only.

3.5.3.1 Effect of vegetation on soil hydraulic conductivity

The soil hydraulic characteristics vary depending on the presence or absence of vegetation and on the evolution of vegetation density. The impacts of plants on soil are, in fact, disparate. Plants influence erosion and sediment transport by limiting the effect of wind and slowing the

surface runoff velocity, therefore constituting an area of potential sediment accumulation. In addition, the superficial soil of a vegetated area is much richer in litter and organic debris, therefore it is richer in nutrients and more porous and permeable. Permeability of deeper layers is also affected by the presence of roots and rotting roots, which create preferential routes for infiltrated water (Boaler and Hodge 1962; Bromley, Brouwer et al. 1997).

All these effects have been observed in areas characterized by vegetation patterns, where vegetated soil exhibits higher permeability than adjacent bare soil, which often has a highly impermeable superficial crust (Valentin, d'Herbès et al. 1999). The range of hydraulic conductivity of an area characterized by vegetation patterns can be very wide, often spanning several orders of magnitude and subjected to random variations within very short distances (Bromley, Brouwer et al. 1997). Soil permeability at a site, therefore, is a function of vegetation density (see schematization in Figure 3.7a).

We propose to model the saturated hydraulic conductivity at time iteration *n* as a continuous function of only the fractional vegetation coverage $\mathbf{M_X}$ at the given pixel **X** and iteration *n*-1 as follows:

$$K_{s\,\mathbf{X}}(n) = \left\{ \sum_{i=1}^{i=10} \left[a_i + b_i \cdot \left(\mathbf{M}_X^{n-1} - 0.1 \cdot i \right) \cdot I_{[0.1(i-1),0.1i)} \left(\mathbf{M}_X^{n-1} \right) \right] \right\}_{n-1} \tag{3.1}$$

where $I_{[a,b]}(M)$ is the indicator function such that

$$I_{[a,b)}(M) = \begin{cases} 1, & a \le M < b \\ 0, & otherwise \end{cases} \tag{3.2}$$

The choice of the coefficients a_i and b_i of equation (3.1) is aimed at obtaining a piecewise continuous function spanning a range of saturated hydraulic conductivities compatible with literature values and field measurements which, for the sites where vegetation patterns emerge, are typically found in the interval 10^{-6} to 10^{-3} cm/s over a range of fractional coverage ranging between 0 and 1. A random component is superimposed to the value of $K_{sX}(n)$ obtained with

equation (3.1) in order to incorporate the typical random spatial variability of soil conductivity (Bromley, Brouwer et al. 1997). A graph describing a sample function for the hydraulic conductivity is provided in Figure 3.6.

3.5.3.2 Run-on infiltration of surface runoff

A non-uniform spatial distribution of hydraulic conductivity affects both the vertical water fluxes at the individual pixel as well as the water input of downstream pixels through the run-on process of infiltration of surface runoff. Surface runoff plays a key factor in the development of soil and vegetation. Part of the surface runoff produced uphill can pond in small depressions or be trapped in areas of litter deposition downhill and infiltrate (Bromley, Brouwer et al. 1997) (see also schematization in Figure 3.7b). The amount of surface runoff that infiltrates depends on many factors, such as the soil properties, the topography, the overall water input, the characteristics of the rain event and so on. The runoff produced accumulates along the hillslope, causing erosion and sediment transport.

We will consider the surface runoff produced uphill of a given area as an addition to the total water input for that area. Thus, the water balance at each pixel **X** at iteration *n* considers a water input, $P_X(n)$, given by the sum of the long-term average precipitation m_{PA} at the pixel and the average surface runoff, R_{sY}, from the uphill pixel **Y** at iteration *n*-1:

$$P_{\mathbf{X}}(n) = m_{PA} + R_{s\mathbf{Y}}^{n-1} \tag{3.3}$$

3.5.4 Effect of vegetation on albedo

Albedo is a measure of the reflectance of a certain surface with respect to solar radiation. The higher the fraction of incoming radiation that is reflected, the less the solar energy available at that surface. For soil-vegetation system, less absorbed solar energy means less available energy for sensible heating and for evaporating water. Consequently, all else being equal, higher albedo corresponds to lower potential rate of evapotranspiration.

Albedo is a characteristic of the reflecting surface and, among other things, depends on soil moisture and on vegetation characteristics including vegetation density. Wetter and more densely vegetated soils are usually darker and less reflective (Figure 3.7c) and, therefore, are characterized by lower albedo. While albedo may have both a negative and a positive dependence on vegetation density depending on the color of bare soil that characterizes the region of study (Rechid, Raddatz et al. 2009), its dependence on soil moisture is clearer, wet terrain usually being less reflective than dry ones (Lobell and Asner 2002; Wang, Wang et al. 2005).

Based on this consideration, and following an empirical formulation developed by Lobell and Asner, (2002), we express the total albedo of pixel **X, ρ_{Tx},** of our study domain at iteration *n* as a decreasing function of the soil moisture content at iteration *n*-1:

$$\rho_{T\,\mathbf{X}}(n) = \left\{ \rho_{T1} - A \cdot \left(e^{-\frac{s_{0\,\mathbf{X}}}{B}} - e^{-\frac{1}{B}} \right) \right\}_{n-1} \tag{3.4}$$

where ρ_{T1} is the all-frequency (i.e., total) surface albedo for saturated soil, arbitrarily set to 0.25, while *A* and *B* are two coefficients whose value is chosen to be respectively 0.11 and 0.3, following Lobell and Asner, (2002). The previous equation accounts explicitly for the dependence of albedo on soil moisture and implicitly for its dependence on vegetation through the dependence of soil moisture on vegetation density.

3.5.4.1 Soil moisture redistribution by roots

Roots allow plants to uptake water and nutrients from the soil. The root configuration is unique of each vegetation species and is affected by the plant's age and health, as well as the soil characteristics, water availability, temperature, and other environmental factors. Developing a sophisticated model encoding all of the aforementioned variables is a complex task that goes beyond the scope of this work. Nevertheless, we propose a basic approach to model the role of roots in redistributing soil moisture. Our approach is based on the simplifying assumption that the root characteristics of the vegetation populating the domain are uniform and that soil moisture can

be rerouted out of a pixel into another by the influence of root networks only if they extend across the pixel borders (Figure 3.7d). Therefore, we define a parameter representing the degree to which roots extend over the adjacent pixel as the ratio of the rooted area and the characteristic area of the pixel:

$$\xi_R = \frac{A_{roots}}{A_{cell}} \tag{3.5}$$

This parameter defines a range of root action and allows us to take into account the process of subsurface water transfer between adjacent pixels promoted by root systems. By further assuming that the root distribution is isotropic, we estimate the net contribution of water input of cell **X**(x-1,y) to cell **Y**(x,y), $\Delta p_{X,Y}$, by means of a function of the following kind:

$$\Delta p_{\mathbf{X},\mathbf{Y}}(n) = \left\{ \frac{\xi_R - 1}{4} \cdot \left(M_{\mathbf{Y}} \cdot k_{v\mathbf{Y}} \cdot P_{\mathbf{X}} - M_{\mathbf{X}} \cdot k_{v\mathbf{X}} \cdot P_{\mathbf{Y}} \right) \right\}_{n-1} \tag{3.6}$$

Equation (3.6) implies that roots spreading across the borders of a cell can uptake a fraction of the water input of the neighboring cell in a way that is proportional (proportionality being given by the parameter ξ_R) to the fractional coverage of the contiguous pixels M_Y and M_X and to their transpiration efficiencies k_{vY} and k_{vX} This, in turn, implicitly assumes that there is no hydraulic redistribution of soil moisture due to hydraulic gradients existing between cells and due to different soil characteristics or the process of root uptake itself. Those effects, however, can be implicitly included in the coefficient ξ_R.

The cumulative effect of roots at pixel **X** is obtained by applying equation (3.6) to each one of its four adjacent cells.

The net result of using this approach is to modify the water input at two contiguous cells based on the difference between the spatial distributions of their root systems, which in turn depends on their fractional coverage.

3.5.4.2 Effect of vegetation on local soil nutrients and transpiration efficiency

This entire study is based on the hypothesis that individual plants act differently from a community of plants (Callaway, Brooker et al. 2002). The interactions between individual plants are multifold and may lead to positive and negative feedbacks on vegetation density. While, on one hand, plants compete for water and nutrients through roots and for light through foliage (Holmgren, Sheffer et al. 1997; Barbier, Couteron et al. 2008), they can also protect each other from extreme fluctuations of temperature and humidity, from mechanical or herbivore damages, and can improve soil properties through litter formation and nutrient replenishment (Holmgren, Sheffer et al. 1997; Borgogno, D'Odorico et al. 2009) (see also Figure 3.7e).

In this study we do not describe each of those interactions individually, but rather model their cumulative effect on the plants transpiration efficiency, k_v, following the reasoning that the net result of facilitation/competition should be to improve/worsen plants water use by increasing/decreasing the quantity of biomass that can be produced out of a certain amount of transpired water.

Therefore, we model the plant transpiration efficiency $k_{vX}(n)$, at a certain pixel **X** and iteration n, as a function of a base value for the transpiration efficiency, $\overline{k_v}$, and the values, at iteration n-1, of the vegetation density at point **X**, $M_{\mathbf{X}}$, the vegetation density of its adjacent cells, M_U, M_D, M_L, M_R, the hillslope-scale vegetation density at the initial time step, $\overline{M}$, the surface runoff at cell **X**, $R_{s\mathbf{X}}$, the surface runoff of the cells upstream and downstream, R_{sU}, R_{sD} and of the average surface runoff for uniform vegetation density at the initial step, $\overline{R_s}$. This function has the following form:

$$k_{v\mathbf{X}}(n) = \max\left\{0.5, \min\left\{1, \overline{k_v} + \left[\left(\Delta k_{v\mathbf{X}}\right)_1 + \left(\Delta k_{v\mathbf{X}}\right)_2 + \left(\Delta k_{v\mathbf{X}}\right)_3 + \left(\Delta k_{v\mathbf{X}}\right)_4 + \left(\Delta k_{v\mathbf{X}}\right)_5\right]_{n-1}\right\}\right\} \quad (3.7)$$

where:

$$\left(\Delta k_{v\mathbf{X}}\right)_1 = \alpha_1 \cdot \frac{M_{\mathbf{X}} - \overline{M}}{\overline{M}} \tag{3.8}$$

$$\left(\Delta k_{v\mathbf{X}}\right)_2 = \alpha_2 \cdot \frac{M_U + M_D + M_L + M_R - 4 \cdot \overline{M}}{\overline{M}} \tag{3.9}$$

$$\left(\Delta k_{v\mathbf{X}}\right)_3 = \alpha_3 \cdot \frac{\overline{R_s} - R_{s\mathbf{X}}}{m_{PA}} \tag{3.10}$$

$$\left(\Delta k_{v\mathbf{X}}\right)_4 = \alpha_4 \cdot \frac{R_{sU} - R_{s\mathbf{X}}}{m_{PA}} \tag{3.11}$$

$$\left(\Delta k_{v\mathbf{X}}\right)_5 = \alpha_5 \cdot \frac{R_{sD} - R_{s\mathbf{X}}}{m_{PA}} \tag{3.12}$$

Equation (3.8) and (3.9) are intended to describe the way that transpiration efficiency at point **X** and iteration *n* is affected by vegetation point **X** and iteration *n*-1 and vegetation at its contiguous cells at iteration *n*-1.

Since the presence of vegetation in the given pixel and in its neighborhoods is assumed to produce a facilitation effect for further establishment of vegetation, we choose both coefficients in equations (3.8) and (3.9) to be negative.

Equations (3.10), (3.11) and (3.12), on the other hand, reflect the effect of surface runoff through the processes of erosion and sedimentation of both soil particles and nutrients. In particular, the coefficients in equations (3.10) and (3.11) are chosen to be negative, while the coefficient in equation (3.12) is chosen to be positive. Those choices are driven by the following considerations: 1) when a given location (i.e. pixel) is subjected to a surface runoff, $R_{s\mathbf{X}}$, larger than the average surface runoff corresponding to uniform coverage, $\overline{R_s}$, it is concurrently subjected to soil erosion and nutrients deprivation; 2) when a given pixel is subjected to a surface runoff, $R_{s\mathbf{X}}$, lower than the surface runoff of the upstream pixel, R_{sU}, it benefits from the partial deposition of incoming soil particles and nutrients and 3) when a given pixel is subjected to a surface runoff, $R_{s\mathbf{X}}$, lower than the downstream pixel, R_{sU}, it is exposed to nutrients and soil loss.

In order to prevent $k_{v\mathbf{x}}(n)$ from taking unfeasible values, we bounded it between 0.5 and 1.

3.6 Simulation of the system

The climate-soil-vegetation system was simulated under various combinations of climatic forcing, soil parameters and characteristics of the spatial interaction functions in order to explore the conditions controlling the mechanism of pattern emergence and evolution. We present results of the simulation of the system on a study domain of 50x50 pixels, representing a hillslope of about 10^5 m^2. Boundary conditions of the system are: 1) fractional vegetation coverage set to be equal to the uniform solution (obtained by using the domain-averaged inputs and equal to the vegetal coverage of each pixel at the preliminary step of simulation) for all the borders; 2) water fluxes equal to the uniform solution ones for the upstream border and the lateral ones; free flow condition in the downstream boundary, allowing complete drainage downhill.

Given the extremely large number of combinations of feasible climatic, hydraulic and topographic conditions, several properties of the system had to be fixed. In particular, unless differently stated, simulations were carried out on a domain whose hydraulic properties and climatic forcing are reported in the "base conditions" column of Table 3.1. The domain is

Simulations were carried out with the following two objectives:

- reproduction of vegetation patterns similar to those present in nature;
- identification of the individual roles that the system forcings and the spatial interactions have on the spatial distribution of vegetation.

3.7 Results and discussion

In accordance to our hypotheses, a satisfactory modeling of the system should be able to: 1) simulate the emergence of patterns showing the same spatial characteristics (i.e. stripes, spots, labyrinths) as the patterns observable in nature; 2) predict the emergence of patterns only for some combination of climate forcings, soil parameters, and characteristics of the spatial

interactions, while predicting no pattern in the remaining cases; the combination of factors leading to the emergence of patterns should also be the same as that observed in nature; and 3) identify the individual role of the pattern-inducing dynamics.

3.7.1 Spatial analysis

In order to be able to compare different typologies of patterns and objectively measure the individual impact of the climatic and hydraulic properties of the system on pattern emergence and characteristics, we perform a thorough spatial analysis of both simulated and natural vegetation fields. In particular, we explore the following spatial characteristics of vegetation fields:

- PDFs and conditional PDFs of vegetation coverage at the pixel level.
- Spectral analysis of the vegetation fields.
- Analysis of the number, size and shape of the vegetation clusters.

3.7.1.1 PDFs and conditional PDFs of vegetation coverage

PDFs of the vegetation coverage at the pixel level can be used to evaluate how different a given vegetation field is from the typical field that would be produced if the number of plants at each pixel followed a Poisson distribution with rate λ. Assuming that each plant covers a given area, in fact, the vegetation fractional coverage of each pixel is a one to one function of the number of plants present at the pixel itself. If plants were present at each point in space with the same probability, each pixel of the domain would have a number of plants (and, thus a fractional coverage) distributed following a Poisson(λ) distribution. Throughout the whole domain, that would lead to a normally distributed PDF of the fractional coverage[1].

In the presence of self-organized structures, the presence of a clump of vegetation at a certain point in space has an impact on the vegetation establishment in its neighborhood. Therefore,

[1] This is a consequence of the Central Limit Theorem.

presence of self-organized structures can be inferred from the analysis of conditional PDFs of the fractional coverage. We do so by evaluating the PDF of vegetation coverage for all those pixels having at least one neighbor characterized by vegetation coverage higher than the overall domain average of the fractional coverage. The same is done conditioning on being in the neighborhood of a pixel characterized by vegetation coverage lower than the domain average. In order to detect spatial anisotropy, conditional PDFs are evaluated for the x-direction (by looking at the vegetation coverage of the two adjacent pixels on the x-direction), on the y-direction and for all directions.

3.7.1.2 Spectral analysis of vegetation fields

Spectral analysis of a signal is a mathematical operation that decomposes a signal into its constituent frequencies. In this context, we look at the spatial field of fractional coverage M and use the Fourier Transform technique to individuate the presence of frequencies in the vegetation field along the x and y direction of our study domain.

3.7.1.3 Analysis of the clusters

In order to examine the characteristics of patterns, we examine a set of spatial features of the vegetation clusters. We arbitrarily define a cluster of vegetation as a clump of adjacent pixels characterized by having fractional coverage larger than the domain average fractional coverage (a qualitative example is provided in Figure 3.8). In defining a cluster, we connect pixels through any shared edge (that is, von Neumann neighborhood; four immediate neighbors, no diagonals), as done by Scanlon (Scanlon, Caylor et al. 2007). For each cluster defined as above, we calculate the size (that is, the number of pixels that compose the cluster), the span along the x and y directions, the shape ratio (as the ratio between the span along the x direction and the span along the y direction) and the fraction of area filled (as the ratio between the cluster size and the product

between the span along the two directions). We also calculate the total number of clusters present in the whole domain.

This definition of clusters allows us to compare our vegetation clumps with the clumps resulting from a homogeneous binomial process. Thus, once the clusters of vegetation are individuated from the original field, we evaluate the coefficient *p* as:

$$p = \frac{\sum_{i=1}^{N_{CLUSTERS}} Size_i}{Size_{DOMAIN}} \tag{3.13}$$

We then compare the cluster statistics of a binomial process with probability *p*, occurring homogeneously within our domain.

We arbitrarily define a pattern as a clustered vegetative configuration whose average cluster size is higher than the 0.975 quantile of the cluster size distribution of the corresponding uniform binomial process. In addition, based on the statistical characteristics of the clumps of vegetation, we distinguish three types of pattern as follows:

- Spots: a pattern whose shape ratio is within the range 0.6-1.6.
- Bands: a pattern whose shape ratio is lower than 0.6 or higher than 1.6.
- Labyrinths: a pattern whose largest cluster spans at least the 75% of the domain and whose fraction of area filled is less than 0.75.

The latter subdivision is motivated by the fact that spots are intuitively defined as structures whose dimensions in the x and y directions are similar (and, thus, characterized by shape ratios close to 1), while bands are characterized by having a dominant dimension (either on x or y). As Labyrinthine fields of vegetation, on the other hand, are characterized by few big clumps of vegetation (thus the reason why we look at the areal span of the bigger cluster) embedding several patches or stripes of bare soil (thus the reason why we look at the area of the cluster effectively filled with vegetation).

3.7.2 Simulated patterns versus natural patterns

A set of sample patterns emerging from our simulations is provided in Figure 3.9. As shown, our modeling of the system is able to predict the emergence of several types of vegetation patterns.

Figure 3.10 shows the cluster analysis of the sample patterns in Figure 3.9. The definition provided in the previous section supports the qualitative observation of Figure 3.9, confirming that only the fields in Figure 3.9a, Figure 3.9b and Figure 3.9c satisfy our definition of pattern, while the field in Figure 3.9d has an average cluster size that is not distinguishable from a homogeneous binomial process. Moreover, the pattern in Figure 3.9a is spotted and has a shape ratio of 1.18, while the one in Figure 3.9b is banded with shape ratio equal to 4.2. Figure 3.9c, instead, corresponds to a labyrinthine pattern because 1) the smallest rectangle that circumscribes the largest cluster occupies an area equal to 91% of the domain and 2) 55% of its area is vegetated.

Below, we present a quantitative comparison between simulated vegetation patterns and natural patterns observed in two African locations, namely an area of Niger near Niamey and a region of Somalia near Garoowe.

3.7.2.1 Niger

The area situated about 45 km south of Niamey, the capital of Niger, is known for the characteristic vegetation patterns known as tiger bushes. The area is characterized by an average annual precipitation of 56 cm, half of which falls at an intensity higher than 35 mm/h and a third above 50 mm/h (Bromley, Brouwer et al. 1997). Soil is gravely sandy loam and is highly prone to crusting in bare areas, while vegetation is concentrated in strips of a few tens of meters wide and a few hundred meters long (Bromley, Brouwer et al. 1997).

The model parameters for the simulations correspond to the climatic and hydraulic characteristics of the area reported in the above literature and are shown in Table 3.1. In addition,

soil hydraulic conductivity at each pixel was calculated as a function of the fractional vegetation cover at that pixel and was set to span a range of $3 \cdot 10^{-7}$ to $9.5 \cdot 10^{-6}$ m/s respectively for the crusted bare soil and a full canopy coverage, as suggested by field measurement (Bromley, Brouwer et al. 1997) and estimations from grain analysis (Casenave and Valentin 1992). Values for the parameters of equations (3.4) through (3.12) are provided in Table 3.2.

Google Earth aerial photographs of this region of Niger characterized by vegetation patterns were used to infer field observations. A few random study areas were sampled from the vast region characterized by the presence of tiger bushes, all of which have a surface area of about 10^5 square meters (see Figure 3.11) and present a mild semi-regular slope. Each picture was then processed in order to estimate the vegetation fractional coverage at every pixel. This was done by examining the color levels of the pixels of the digitized pictures using as intervals the null vegetation coverage, that is, M=0, for those pixels characterized by bare soil and M=1 for those ones characterized by full coverage. Photos, whose original resolution was of about 400x400 pixels, were then further processed in order to match the resolution of our study grid. This was done by superimposing our study grid on the original photo and averaging the fractional coverage of the set of pixels of the original photo that fell within the bounds of each pixel of our 50x50 grid.

A preliminary qualitative comparison between simulations and observations is presented in Figure 3.11, where three original aerial photos are shown together with their digitized vegetation coverage maps and three sample results from our simulations. The figure shows a good qualitative agreement between natural and simulated patterns in terms of typical shape and dimension of the vegetation structures and in the overall spatial configuration of the patterns within the study domain.

Results of several simulations exhibited a noteworthy sensitivity of the emerging patterns to changes in the spatial interaction functions and in particular to the dependence of k_v and hydraulic conductivity on vegetation. Differences between patterns in Figure 3.11g and Figure 3.11f, for

example, are due to changes of about 5% in the coefficients of the equation (3.9), while Figure 3.11h was obtained by increasing the soil conductivity in the interval corresponding to a fractional cover only in the range of 0.3 to 0.5 by about 10%, while keeping the overall span of the range fixed between $3 \cdot 10^{-7}$ to $9.5 \cdot 10^{-6}$ m/s. In addition, results from simulations carried out with the same input parameters showed appreciable difference simply as a consequence of the random perturbation to which they were subjected at the preliminary step of iteration. However, differences were mainly noticed in the placement of the vegetation structures within the study domain rather than in their typical shape or dimension.

An analysis of the spatial distribution of water fluxes and soil moisture content is provided in Figure 3.12. In particular, Figure 3.12a shows the value of the average effective precipitation at each pixel, computed as average precipitation plus the run-on. As shown, many areas receive an amount of water several times higher than the actual mean precipitation from the low permeable pixels located upstream, in accordance with field observation of factors of concentration (ratio between the effective amount of water received and the actual precipitation) as high as 3 and 4 (Bromley, Brouwer et al. 1997). In addition, the higher values of groundwater runoff observable in correspondence of the vegetated patches (shown in Figure 3.12b) confirm that vegetation favors the infiltration of the hillslope run-on. Taken together, the extra water input from upstream and the enhanced permeability of the more vegetated soil trigger a positive feedback for further vegetal biomass establishment, confirming that surface water redistribution due to ponding and subsequent infiltration of run-on is one of the main drivers of pattern formation (Boaler and Hodge 1962; Casenave and Valentin 1992; Valentin, d'Herbès et al. 1999). Moreover, in accordance with published results (Borgogno, D'Odorico et al. 2009), higher values of average soil moisture content are also observed under vegetated patches (especially in their uphill side) than in the surrounding bare soil (Figure 3.12c), which, in turn, create suitable conditions for the sustainment and/or further development of vegetation.

In order to characterize quantitatively some of the spatial features of the Niger tiger bushes and compare them with the results of the simulations, we estimated the PDFs of vegetation coverage. In particular, we evaluated the PDF of the vegetation coverage at the pixel level for both the simulated and the digitized observed patterns. In addition, PDFs of the vegetation coverage were estimated conditioned on the vegetation coverage of the neighbor pixel. The latter was done by calculating the PDF of vegetation coverage for all those pixels having at least one neighbor characterized by vegetation coverage higher than the mean. The same was done conditioning on vegetation coverage lower than the domain average. Conditional PDFs were calculated for the x-direction (by looking at the vegetation coverage of the two adjacent pixels on the x-direction), on the y-direction and for all directions.

As shown in Figure 3.13, both observed and simulated PDFs are bimodal, supporting the observation that vegetation coverage is not normally distributed around the mean, as it would be expected if plants were spatially distributed as a homogeneous point process across the domain. Bimodality is more evident when the PDF of vegetal cover is conditioned on the neighbor pixel having a vegetation cover higher than the domain average. The latter suggests that the vegetal coverage of each pixel is more likely to be higher than the average providing that it is in the neighborhood of a pixel whose cover is also higher than the spatial average. The opposite is true when the condition is on the neighbor pixel having a coverage which is lower than the average. A slight prevalence of structures in the x-direction (perpendicular to the domain slope) is apparent from the analysis of the conditional PDFs for the pattern under analysis. This is apparent from both the qualitative observation of natural and simulated structures (Figure 3.11a, Figure 3.11d and Figure 3.11g) and from the analysis of the directional conditional PDF of Figure 3.13. Looking at the interval $0<M<0.1$, in fact, the PDF of the vegetal coverage of pixels being in the x-direction neighborhood of a pixel with vegetal coverage higher than the domain average shows a lower density than the y-directional conditional PDF. Conversely, for the same interval of

fractional coverage, higher density is shown for the PDF of the vegetal coverage of pixels being in the x-direction neighborhood of a pixel with vegetal coverage lower than the domain average.

Spectral analysis of the vegetation density field M(x,y) is provided for both natural and simulated patterns in Figure 3.14. Average of the one-dimensional spectral densities, evaluated for both the x-direction and the y-direction are provided along with the two-dimensional power spectrum. Peak spectral densities appear in correspondence with the lower frequencies, capturing the presence of the large scale structures. In agreement with the analysis of the directional PDFs, spectral analysis confirms the presence of a slight anisotropy in the pattern shape for both the natural and the simulated case, by detecting a prevalence of structures recurring at frequencies between 2 and 6 (that is, 2 to 6 structures per domain length) along the y-direction.

The analyses of the characteristics of the vegetation clusters of both natural and simulated vegetation fields, whose results are reported in Table 3.3, show that two out of three natural patterns (Figure 3.11d and Figure 3.11e) fall within our definition of spots, while the pattern in Figure 3.11f is labyrinthine. Simulated patterns, however, are all spotted, even though the one shown in Figure 3.11i, whose largest cluster covers the 68% of the domain, almost meets the requirement for being classified as labyrinthine.

3.7.2.2 Somalia

Vegetation stripes are a widespread occurrence on the Somaliland plateau (Boaler and Hodge 1962) and in the Puntland area (Borgogno, D'Odorico et al. 2009). The area is semi-desert and characterized by an arid to semi-arid climate with precipitation highly variable in space and time (Muchiri 2007). The area selected is located about 30 km west of Garoowe, the administrative capital of the Puntland region of Somalia and is characterized by annual precipitation ranging between 10 and 20 cm/year mostly occurring in the period of May-September (Muchiri 2007). Dominant soils are Gypsisol and Calcisol (Venema 2007) according to the FAO definition and are characterized by hydraulic conductivity that can be as low as a few cm/day but can span two

orders of magnitude. These kinds of soil are also very susceptible to crusting and cracking (FAO 2001).

A study area of about $5 \cdot 10^5$ m^2 located at 7°43´N, 48°02´E and characterized by the presence of vegetation strips was selected from the Puntland area near Garoowe. In comparison to the Niger case, overall climate is drier (precipitation being less than one third that of the Niger case) and patterns – in this case well-defined strips – occur at a slightly larger scale. In accordance with the climatic and soil information reported above, the simulations of the area used the parameters shown in Table 3.1 and Table 3.2. Results of the simulations were compared to the observed patterns, as shown in Figure 3.15 through Figure 3.18.

As done for the case of the Niger tiger bushes, a sample aerial photograph was processed in order to obtain estimates of the vegetal coverage at a 50x50 study grid, and compared with the simulated results (Figure 3.15). The comparison shows a good qualitative match between observed and simulated vegetal spatial configurations. Dimensions and shape of the bands and of the inter-band gaps are similar, although the vegetation structures emerging from the simulations look slightly sharper then the observed ones. In addition, orientation of the bands in the direction perpendicular to the slope is clearer in the simulations than in the observations. This is due to the fact that, although the direction of the natural and simulated slope was set to coincide, the natural topography presented some irregularities, while the simulations were performed on a regular slope. Nevertheless, both in the observed and in the simulated case, the configuration of vegetal coverage is characterized by strips of vegetation a few tens of meters wide and extending for the whole length of the study domain.

Compared to the previously analyzed case of the Niger tiger bushes, the directionality of the Somaliland plateau vegetation patterns is more noticeable, both in the observations and in the simulations. Moreover, the overall amount of biomass (integrated across the domain) is lower than in the case of the Niger tiger bushes, as easily noticeable from the comparison of the PDFs of the vegetal coverage shown in Figure 3.13 and Figure 3.16. Analysis of the PDFs of Figure

3.16 shows also that the presence of multiple modes is less evident here than in the case of tiger bushes when the unconditioned PDF of the vegetation coverage is considered. However, multiple modes become apparent in the conditional PDFs of both the observed and the simulated vegetation fields, especially with respect to the PDF in the x-direction conditioned on neighbor coverage higher than the domain average. The presence of multiple modes in the conditional PDFs and, in general, the fact that the conditional PDFs look different than the unconditional one, implies that the vegetation at a given pixel has an impact on the vegetation distribution of its neighborhood. In particular, the fact that the probability of finding a pixel whose fractional coverage is higher than the average is higher in the neighborhood of pixels that are themselves characterized by coverage higher than the average supports the observation that vegetation tends to form clumps rather than being distributed completely randomly in space. In addition, the fact that the PDF conditioned on x-direction neighbor coverage higher than the domain average shows higher density for values of vegetation in the upper end of the domain interval than the other (conditional and unconditional) PDFs confirms the prevalence of strips of vegetation in the x-direction itself, that is, perpendicular to the domain slope.

Spectral analysis (Figure 3.17) supports the conclusions drawn from both the qualitative analysis and the analysis of the PDFs of the vegetation coverage, showing the presence of frequencies between 5 and 10 cycles along the domain only along one direction. However, comparisons between the spectral densities of the observed and simulated vegetation field also imply that the prevalence of stripes perpendicular to the main slope is higher in the simulated field than the observed. Since the only anisotropic effect present in our model is topographic (through the surface runoff production and run-on infiltration), and the shape of the natural pattern seems to indicate that the governing mechanism of pattern formation is topographic and gravitational, we attribute the discrepancy between observations and simulations to the to the irregularities of the natural topography of the observed area with respect to the regular slope used in the simulated domain.

Statistical analysis of the vegetation clusters, whose results are reported in Table 3.3, show that both natural and simulated patterns are banded.

3.7.3 Analysis of the hypothesized pattern-promoting dynamics

3.7.3.1 Impact of climate forcing

In order to explore the effect of the climate forcing on pattern formation and spatial characteristics of the vegetation distribution, the system was simulated by varying the climatic forcing, while keeping everything else fixed. All the inputs for these simulations are reported in the column "base conditions" of Table 3.1 and Table 3.2. We investigated the impact of two climatic components: mean annual precipitation and mean annual solar radiation.

Simulations showed that the shape and the presence of patterns at the hillslope scale depend not only on the mean annual precipitation but also on the parameters of the rainy events (mean storm duration, mean time between storms, mean storm intensity, etc.). Below, we focus our analysis only on the dependence on the mean annual precipitation.

Figure 3.18 shows the result of the statistical analysis of the vegetation fields obtained by varying annual precipitation in the range 26 to 85 cm. The variation in mean annual precipitation was achieved by varying the storm duration only and leaving the mean storm intensity unchanged. All the other parameters of the model, both climatic and hydraulic were kept at the values set for the "base conditions". Figure 3.18a shows that, according to our classification, patterns of vegetation start to emerge for annual precipitation higher than 32 cm and cease to exist when precipitation approaches 60 cm per year. All the patterned fields are banded, as shown in Figure 3.18b, and none presents labyrinthine characteristics, as shown in Figure 3.18c and Figure 3.18d. Three sample fields obtained for annual precipitation of 70, 48 and 32 cm per year are presented in Figure 3.19 for a visual interpretation of the transition from undistinguishable patterns to self-organized structures.

The effect of solar radiation is the opposite, that is, patterns start to emerge as the incoming solar radiation increases, everything else being equal (Figure 3.20 and Figure 3.21). According to our classification, patterns emerge as solar radiation exceeds 230 W/m^2. As solar radiation increases, and, therefore, climatic conditions become more arid (since annual precipitation is kept fixed), clumps of vegetation decrease in number and increase in average size.

Neither precipitation nor solar radiation is, by itself, sufficient to characterize the aridity of a given climate. In order to incorporate both parameters, therefore, as Eagleson (1978) we define potential humidity, PH, as the ratio between the annual precipitation and the domain-averaged annual potential evapotranspiration and estimate the relationship between PH and the spatial characteristics of the simulated vegetation fields.

In order to characterize the relationship between PH and the spatial characteristics of the vegetation fields, we use the same fields simulated before by varying annual precipitation and solar radiation. Results of the statistical analysis of the vegetation clusters are reported in Figure 3.22. We notice that vegetation patterns arise for values of PH between 0.2 and 0.3 and that, for the other climatic and hydraulic conditions characterizing this set of simulations, all patterns are banded.

In general, the analysis of Figure 3.18, Figure 3.20 and Figure 3.22, supports the observation that patterns arise in arid and semi-arid areas, that is, in water limited environments, agreeing with all the available literature on the topic. For wetter climates, in fact, (as here occurs for PH higher than 0.3) the vertical water input is enough to support a substantial amount of vegetation even in absence of surface water redistribution or mechanisms of facilitation/competition. When this occurs, the impact of those dynamics of surface water redistribution and facilitation/competition becomes comparatively less important (that is, the lateral interactions are overpowered by the vertical water and energy fluxes) and does not induce the emergence of recognizable patterns. As the conditions become more arid, that is, for lower values of PH (in our simulations for $0.2<PH<0.3$), the water input from lateral redistribution becomes determinant for

the amount of vegetation that establishes at each pixel; moreover, the benefits of the facilitation mechanisms existing in the neighborhood of vegetated pixels become (comparatively to more humid conditions) more significant and, thus, promote vegetation rearrangement and pattern formation. For the lowest values of PH, that is, for the most arid conditions explored in this analysis, on the other hand, the climate conditions are so adverse to vegetation establishment that the study hillslope tends to be too scarcely vegetated for the facilitation/competition dynamics to take place, ultimately resulting in the absence of patterns.

3.7.3.2 Impact of slope

As reported in the literature, some observed vegetation patterns tend to migrate uphill (Worral 1959; Worral 1960; Valentin, d'Herbès et al. 1999; Sherrat 2005). Although we do not explicitly include a time description of the system evolution, our iterative modeling approach has an implied time evolution. Therefore, we can infer information about the development of the system over time from its evolution through the numerical iterative process. Once a pattern is established, in fact, the vegetal density at each pixel can either remain fixed (or not change significantly in between simulation steps), or it can undergo changes that – although significant at the pixel level - do not alter the macroscopic structure of the pattern itself, as in the case of vegetation structures that migrate across the domain.

In order to analyze this effect, we compared the vegetation field corresponding to the base conditions at different iteration steps, reporting in Figure 3.23 the vegetation coverage at the 45^{th}, 50^{th}, 55^{th} and 60^{th} iteration of our simulation procedure. It is evident that the patterns have been already established by the 45^{th} step of the iteration, shown in Figure 3.23a, and that the stripes migrate uphill as the simulation progresses. This migration is induced by the effect of surface run-on ponding and infiltration in the uphill part of the stripes, which in turn creates a favorable opportunity for the uphill expansion or migration of the vegetation. This claim is supported by the analysis of the spatial distribution of soil moisture and water fluxes (not shown), which confirms

the presence of wetter soils and higher infiltration in the uphill portion of the vegetation bands, which as the simulation progresses, creates a favorable environment for further vegetation establishment and a positive feedback for the uphill expansion of the vegetation clumps. The drier conditions observed in the downhill portion of the bands and created by the fact that most of the available run-on has already infiltrated, on the other hand, result in the creation of adverse conditions that inhibit vegetation establishment.

However, pattern migration was not observed in all the simulated cases. Several simulations (not shown) developed patterns that, once established, did not exhibit any tendency to migrate from their original location. This is attributed to the predominance of the local inhibition dynamics present in the uphill portion of vegetation clusters. In those cases, in fact, it has been observed for the pixels immediately uphill of a clump of vegetation that the inhibition effect due to the terms in equations (3.10) and (3.11) (which reflect the effect of soil erosion due to the surface runoff) overpower the facilitation due to the presence of vegetation immediately downhill.

3.7.3.3 Impact of hydraulic conductivity

Expressing hydraulic conductivity as a function of vegetal coverage is a way to incorporate the effect of plants on the permeability of the soil. Regions where vegetation patterns occur are characterized by a soil permeability that is highly variable in space, being higher in presence of vegetated soil and lower where the soil is bare (Boaler and Hodge 1962; Bromley, Brouwer et al. 1997; Valentin, d'Herbès et al. 1999; HilleRisLambers, Rietkerk et al. 2001; Saco, Willgoose et al. 2007).

Here we investigate the impact that a vegetation-dependent hydraulic conductivity has on pattern formation. To this aim, we first compare the following three situations:

- hydraulic conductivity determined at each pixel as a function of vegetation according to equation (3.1) (base conditions);

- hydraulic conductivity fixed in space and set equal to the spatial average of the hydraulic conductivity corresponding to the base conditions;
- hydraulic conductivity at each pixel randomly sampled from a uniform distribution spanning the same range of hydraulic conductivities as in the first case.

As shown in Figure 3.24a, no pattern emerged as the result of the system simulation in case of constant hydraulic conductivity. In the case of hydraulic conductivity randomly distributed in space we also observe the absence of well-defined structures. However, in this latter case, the spatial variability of the vegetation coverage is higher than the case with constant conductivity (as apparent in Figure 3.24b) but with a distribution that does not present the characteristic (bimodality, asymmetry) found in the ones where well-defined patterns were evident (Figure 3.31). In particular, the vegetation that arises in the case of Figure 3.24b traces the spatial distribution of the hydraulic conductivity itself (pixels with higher conductivity soils are more vegetated than those where the soil is less permeable). In both cases of Figure 3.24a and Figure 3.24b, the absence of feedback between vegetation and the hydraulic properties of the soil prevents well-defined patterns from emerging, even when all the other spatial effects are in play. This suggests that the primary driver of the mechanism of pattern formation (for a given set of climatic conditions) is the capacity of vegetation to affect the soil properties and, thus, the spatial distribution of water fluxes.

Although not shown, no patterns were observed in the hypothetical case of perfectly horizontal hillslope, simulating the extreme condition of absence of redistribution of surface run-on. This, along with the previous findings, individuates in the mechanisms of surface runoff production and surface run-on infiltration the primary drivers of the phenomenon for a given set of climatic conditions.

Once it has been established that dependence of hydraulic conductivity on vegetation density is necessary for the formation of patterns, we investigate the role played by the shape of the function in equation (3.1) in the ultimate vegetation configuration. To this purpose, we simulate

the system with a set of alternative functions of the type in equation (3.1). Those equations were obtained from the base condition equation (whose parameters are reported in Table 3.2) by using the following transformation:

$$\{K_s(M)\}_i = [1 + A_i \cdot \sin(\pi \cdot M)] \cdot \{K_s(M)\}_{BaseConditions} \tag{3.14}$$

where $\{K_s(M)\}_i$ is the hydraulic conductivity function of simulation *i*, *lb* and *ub* are respectively the lower and upper bound of the base condition function and A_i represents a scaling factor. Such formulation allows us to simulate the system using hydraulic conductivity functions that span the same range as that of the base condition, while varying the shape of those functions, as shown in Figure 3.25.

Eight simulations were performed using hydraulic conductivity functions obtained through equation (3.14), with coefficients *A* equal to: -0.4, -0.3, -0.2, -0.1, 0.1, 0.2, 0.3 and 0.4. The statistical analysis of the vegetation clusters, shown in Figure 3.26, shows the impact that the shape of the hydraulic conductivity function has on the spatial configuration of the vegetation. As noticeable in Figure 3.26a, patterns emerge only for $A \geq -0.3$, and only for $A \geq -0.2$are banded ($A=0.3$ produces a spotted configuration according to our cluster classification criteria). No labyrinthine configurations were found with the analyzed set of conditions. As it is noticeable from Figure 3.25, the functions obtained through equation (3.14) are not dramatically different from that of the base conditions, even in the two limit cases of $A=-0.4$ and $A=0.4$. Nevertheless, even little differences may have a strong impact on the system response. This happens, for example, in the case of $A=-0.4$, where the shape of the hydraulic conductivity function prevents patterns from emerging. To explain this behavior we observe (from Figure 3.31) that the PDF of the fractional coverage has modes at $M \approx 0.1$ (bare soil patches) and $M \approx 0.45$ (vegetated clumps), and that those two values of *M* correspond to the range for which the slope of the hydraulic conductivity function of the base conditions is higher. In the base conditions simulation this large function gradient allows areas with $M \approx 0.4$ to be sufficiently permeable to favor water infiltration

and further vegetation establishment, triggering the positive feedback which ultimately promotes pattern formation. In the case of A=-0.4, instead, the soil permeability required to favor run-on infiltration would be reached in areas with vegetal cover M>0.7, which is too high to be sustainable, given the climatic and the hydraulic properties of the system. Other simulations performed on different sets of climatic and hydraulic conditions support this finding and indicate that each set of climatic and hydraulic conditions require the hydraulic conductivity function to have a particular shape for the vegetation configuration to be patterned. Specifically, the hydraulic conductivity function must be such that: 1) the permeability for low vegetated areas (e.g. M<0.2) promotes the formation of surface runoff without allowing further establishment of vegetation (and, thus, positive feedback on permeability) and 2) the vegetated areas (e.g. M>0.4) are permeable enough to allow run-on infiltration and sustain (for the given climatic conditions) their vegetal coverage and/or promote further vegetation establishment.

3.7.3.4 Impact of local vegetation interactions

The effect of interactions between adjacent clumps of vegetation was modeled, as indicated before, by allowing the transpiration efficiency of the vegetation at a certain pixel to depend on the vegetal density of the nearby pixels (see equation (3.7)). We explored the effect on pattern formation of this spatial interaction by examining the evolution of the system in its absence and comparing it with the results obtained for the base conditions, which includes it. Figure 3.27 shows that even in the absence of our modeling of the spatial interactions between plants in adjacent pixels (equation (3.7)), the system evolves towards a patterned configuration. However, the shape of the pattern and the total amount of vegetation arising in the domain are different in the two cases, as also evident in Figure 3.31. This suggests that the effect of all those spatial interactions encoded in equations (3.8) through (3.12) (protection from temperature and humidity fluctuations, protection from soil erosion, protection from the mechanical damages due to winds and animals, enhancement of soil fertility through litter formation and nutrient replenishment and

so on) is important for the ultimate configuration of patterns and total amount of biomass produced, although not essential for the emergence of the patterns themselves.

In order to investigate the individual effect of the coefficients of equations (3.8) through (3.12), we carried out five simulations of the system, each one performed by setting one of the coefficients α_i of equations (3.8) through (3.12) equal to zero. Statistical analysis of the vegetation clusters obtained in those five cases is compared to the base conditions, as illustrated in Figure 3.28. It is noticeable that the simulation obtained with α_2=0, which corresponds to the absence of facilitation due to surrounding vegetation, does not lead to the formation of patterns, whereas the other cases simply affect the shape and dimension of the emerging patterns, although never in a measure that results in the transition to spots or labyrinths.

3.7.3.5 Impact of soil moisture redistribution due to roots

In order to investigate the role of the soil moisture redistribution due to the presence of roots, we compare the spatial distribution of vegetation obtained by allowing vegetation from each pixel to extract water from up to 15% of the area of each of the four adjacent pixels with the one resulting from neglecting this effect. Results of this comparison are presented in Figure 3.29. It is noticeable that the patterns that arise when this effect is neglected are nearly indistinguishable from the ones arising in the base conditions case. However, a deeper analysis (Figure 3.31) shows that the PDF of the vegetation density obtained in the case where no competition for available soil moisture due to roots is considered has a lower mode for the bare soil interval than the general case, with an average vegetation (average of the pixel vegetation density across the domain) being roughly equal (0.209 against 0.208). Since the magnitude of this effect is way lower than the cases previously analyzed, the overall conclusion is that roots effect has a minor promoting effect on the formation of patterns.

3.7.3.6 Impact of albedo

Modeling the soil reflectance as a function of soil moisture content was aimed at incorporating the mutual effect that spatially variable water fluxes and vegetation have on the local energy budget through the local potential rate of evapotranspiration. Figure 3.30 presents a comparison between a case in which this effect is modeled as proposed in equation (3.4) and the case in which this effect is totally neglected and the reflectance is set constant in space and equal to the average of the values of reflectance of each pixel of the general case. Although a spatially variable reflectance determines a spatially variable absorbed incident radiation and, thus, affects the energy budget of the entire hillslope, the patterns obtained in the two cases described above are qualitatively indistinguishable. The analysis of the PDF of vegetal coverage shown in Figure 3.31, however, suggests that the vegetal configuration over the hillslope deviates from the base condition case. In particular, the fact that the PDF obtained with a constant albedo shows higher modes (for both low vegetation and high vegetation) than the general case, indicates that a solar reflectance that increases as soil moisture decreases has a minor negative impact on the formation of patterns.

3.8 Summary and conclusion

In this study we proposed a mechanistic modeling of those hydro-thermo-dynamics that, at the hillslope scale, are responsible for the phenomenon of vegetation pattern formation and evolution. The model performs the water and energy balance of a hillslope and accounts for water and energy fluxes routing over the study domain. The dynamics inducing the emergence of vegetation patterns are explicitly identified and modeled and their individual impact on the phenomenon is quantified.

Our results show that the proposed model is able to qualitatively reproduce the types of patterns commonly referred in literature as bands, spots and labyrinths. The model was satisfactorily validated by comparing simulations with observed natural patterns in the areas of

Niger near Niamey and Somalia near Garoowe. Our model was able to capture the local dynamics inducing the formation of patterns and generated results that are qualitatively and quantitatively compatible with the observations and the literature information.

The analyses of the processes involved in the formation of patterns suggest that the phenomenon is primarily driven by run-on infiltration and local mechanisms of facilitation/competition existing among adjacent vegetation groups. Nonetheless, even in presence of those mechanisms, patterns arise only when the climatic conditions, particularly annual precipitation and solar radiation, are favorable. In particular, we found that, with decreasing precipitation or, conversely, increasing incoming radiation, the system drifts from fully vegetated with undistinguishable vegetation structures to self-organized patterns due to the equilibrium between short-range facilitation mechanism, that tend to aggregate vegetation in clusters, and fight for limited resources that impede the system to sustain more than a given amount of vegetal biomass. When the behavior of the system is analyzed as a function of the potential humidity, which incorporates both the effect of precipitation and solar radiation, we found that patterns emerge when the potential humidity index is within a certain range (that for our simulated condition was found to be between 0.2 and 0.3), while no distinguishable patterns arise for climatic conditions too arid or too humid.

In the range of climatic conditions favorable to the formation of self-organized vegetation structures, the peculiar spatial features of patterns are determined by the characteristics of the spatial interactions induced by run-on infiltration, facilitation/inhibition dynamics between adjacent vegetation groups, effects of nutrient and litter transport and deposition, competition for soil moisture through roots and effect of spatially inhomogeneous surface reflectance. Nonetheless, our study indicates that the surface run-on infiltration is the dominant dynamics. No self organized structures, in fact, were observed in the absence of any surface runoff production and subsequent run-on infiltration; moreover, the system was found to be extremely sensitive to the relationship between vegetation density and soil permeability. As for the impact of the other

dynamics, we found that the effect of facilitation/competition due to vegetation interactions, soil erosion and nutrients transport has the highest impact on the phenomenon and has an influence on the ultimate shape of patterns; among those dynamics, the effect of facilitation due to the presence of surrounding vegetation was found to be the more significant. Effects of roots and albedo are comparatively less important but still have an impact on pattern definition, evolution and on the total biomass that establishes on the domain.

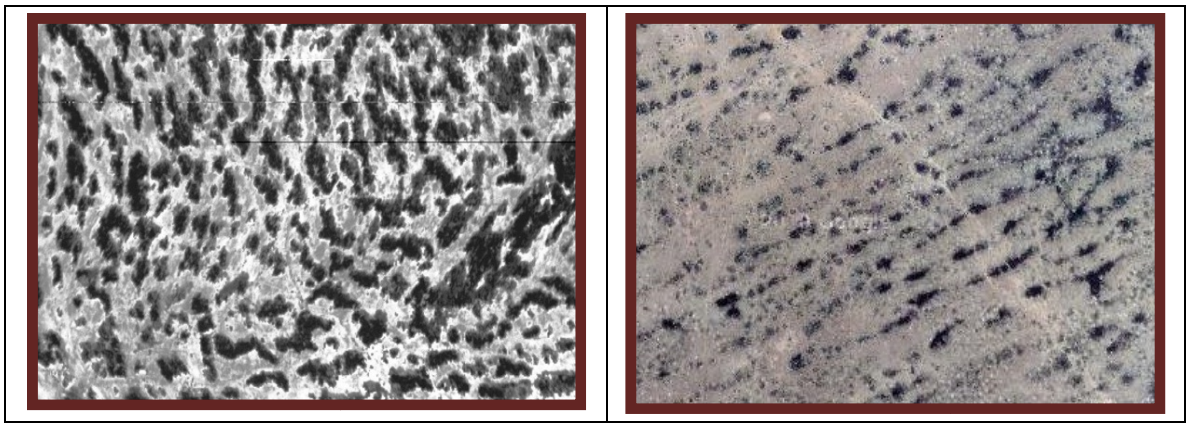


Figure 3.1 Vegetation patterns in Niger (from D'Odorico et al, 2006). b) Vegetation stripes in Senegal (from google maps).

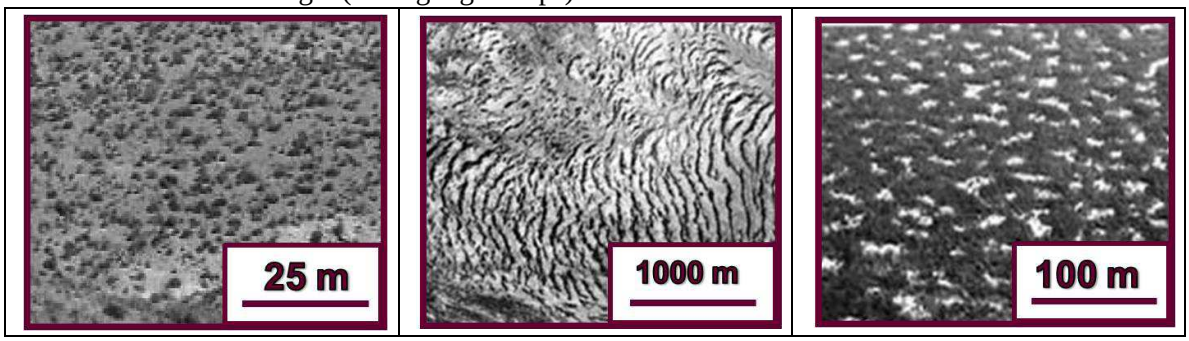


Figure 3.2 Vegetation patterns typical shapes and dimensions.

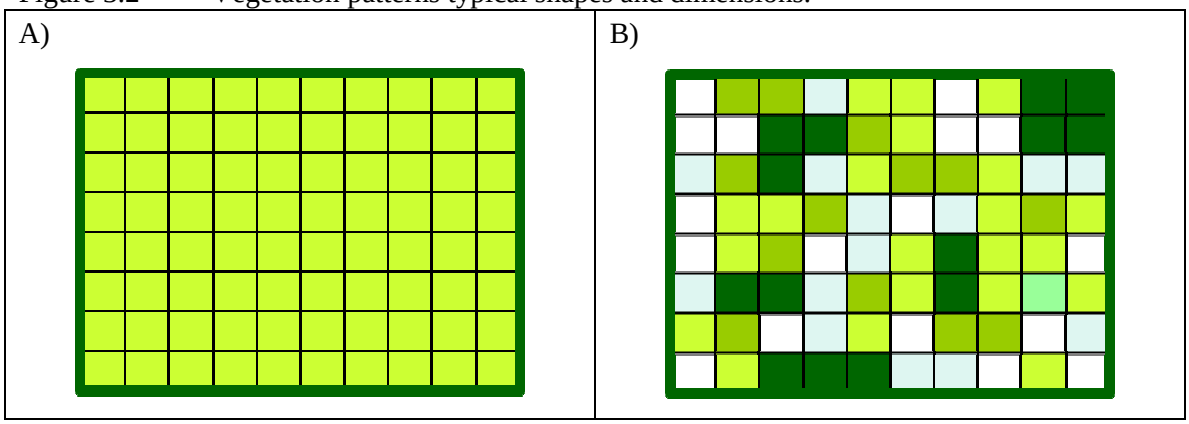


Figure 3.3 A) schematization of the study domain and study subgrid, B) schematization of uneven spatial distribution of vegetation across the study domain. The study domain is representative of a hillslope and has size of about 10^5-10^7 m^2. Each pixel of the study subgrid will be representative of an area of about 10^0-10^2 m^2.

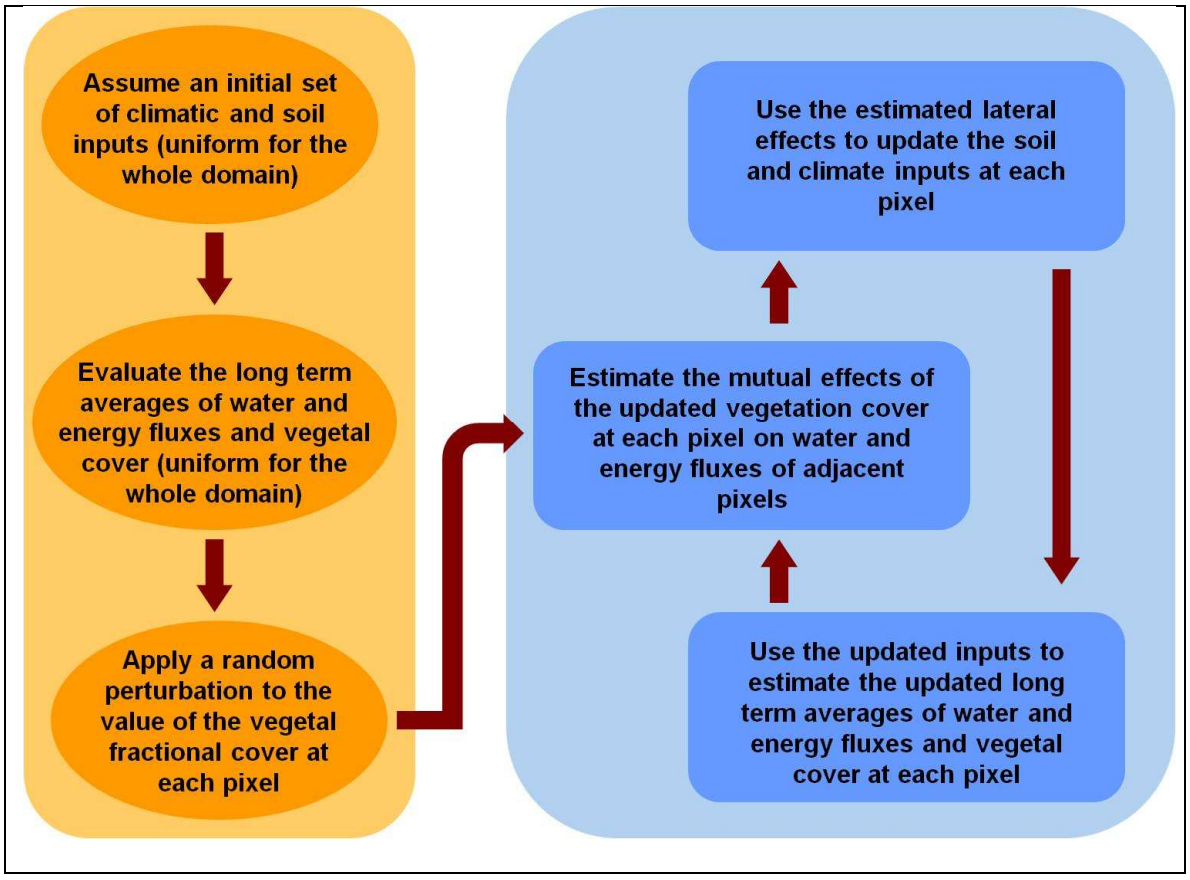


Figure 3.4Flowchart of the procedure of simulation of the soil-climate-vegetation system. The orange blocks represent the first steps of simulation, while the loop in the blue blocks represents the standard iterative steps.

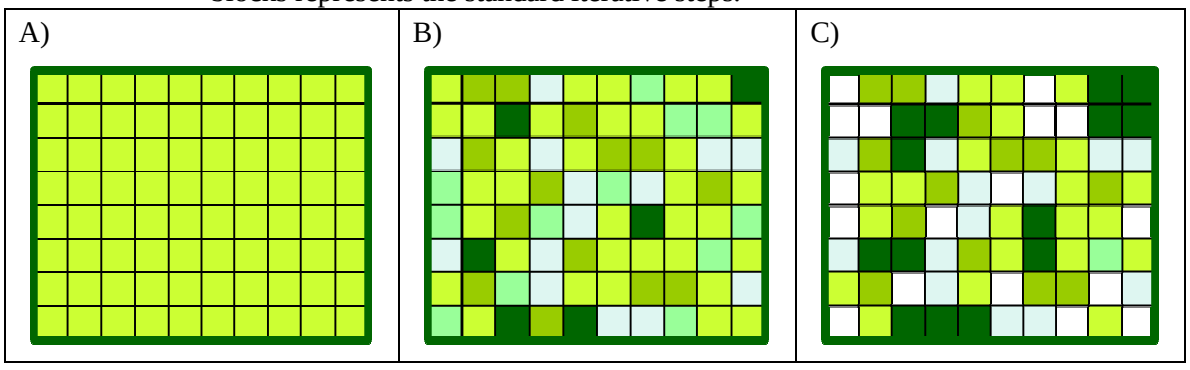


Figure 3.5 Qualitative schematization of the preliminary steps of simulation procedure: a) uniform vegetal coverage obtained by applying the water and energy budget with the given set of soil and climate inputs; b) vegetation coverage after randomly perturbing the fractional coverage of each pixel; c) vegetation coverage after evaluating the spatial effects, updating the soil and climatic parameters at each pixel and performing the water and energy budget.

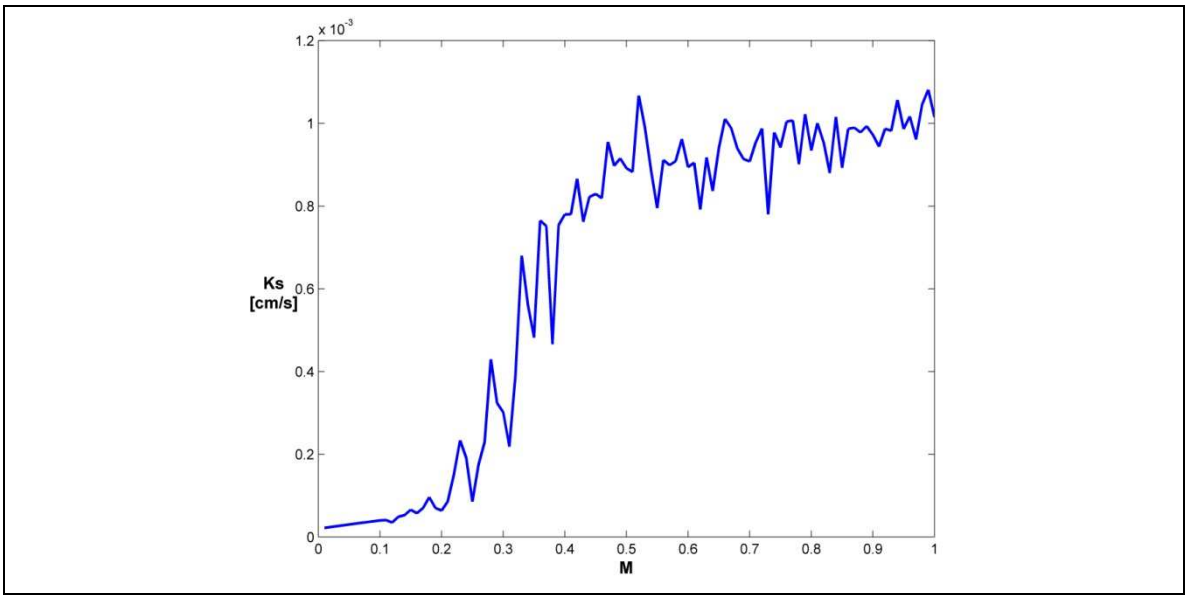


Figure 3.6 Sample function showing the relationship between hydraulic conductivity and vegetation fractional cover.

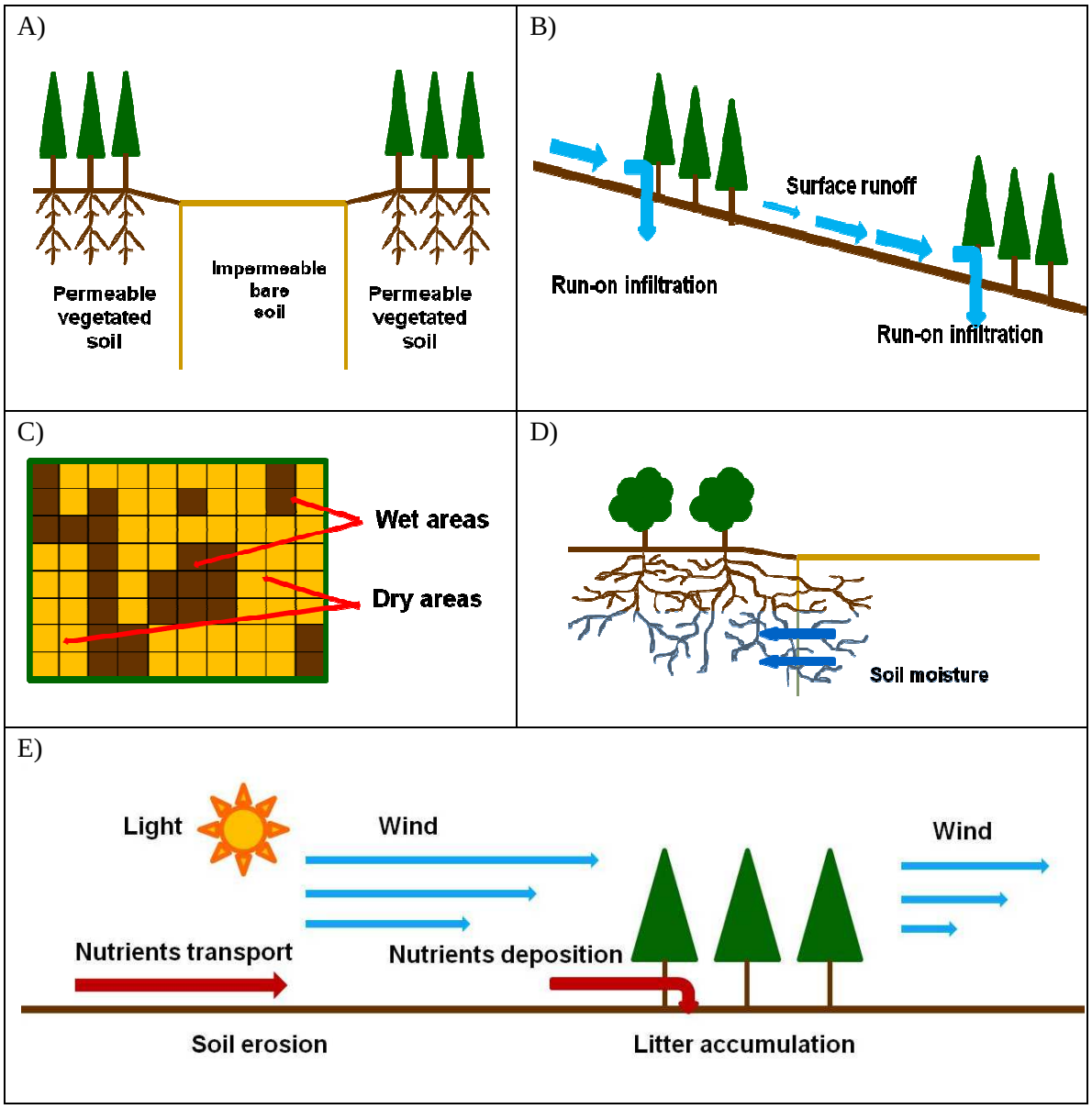


Figure 3.7 Qualitative schematization of the interactions between soil, climate and vegetation. A) Effect of vegetation on soil permeability. B) Effect of vegetation on run-on infiltration. C) Effect of soil moisture on soil reflectance. D) Effect of roots on soil moisture. E) Other spatial interactions (effect of vegetation on wind, light, nutrients, etc.).

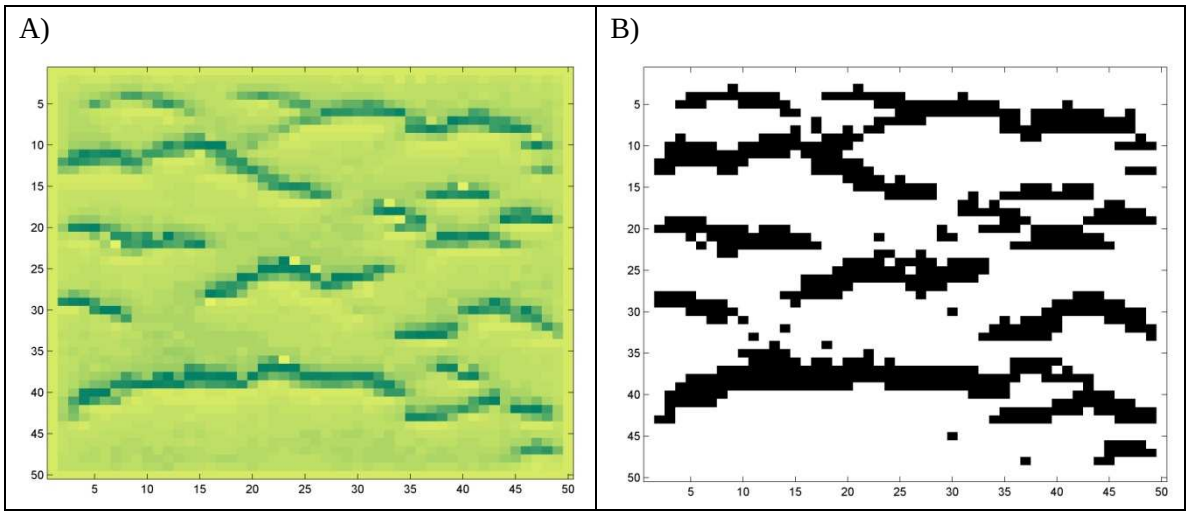


Figure 3.8 A) A sample vegetation field. B) Clustered field corresponding to the vegetation field in A.

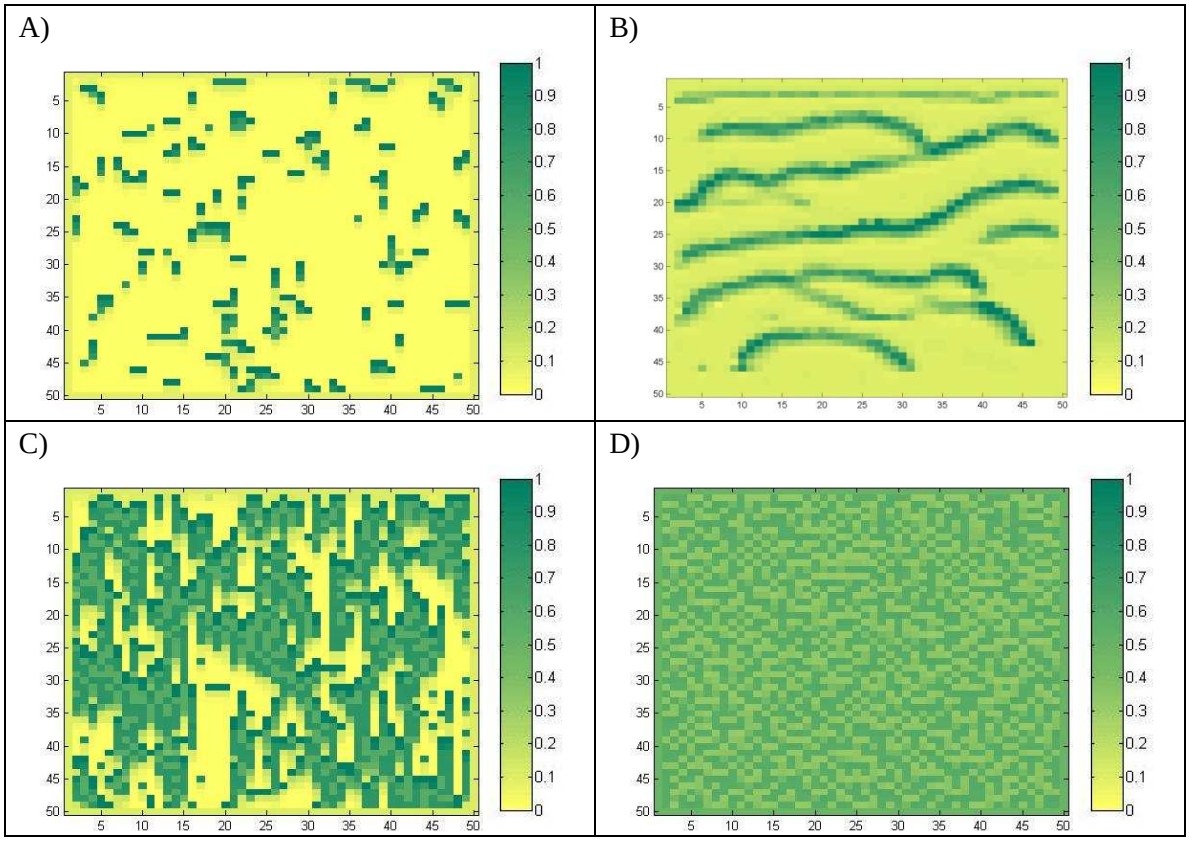


Figure 3.9 Sample results of the system simulations, expressed in terms of fractional vegetation coverage, *M*. A) A spotted pattern. B) A banded pattern. C) A labyrinthine pattern. D) No pattern.

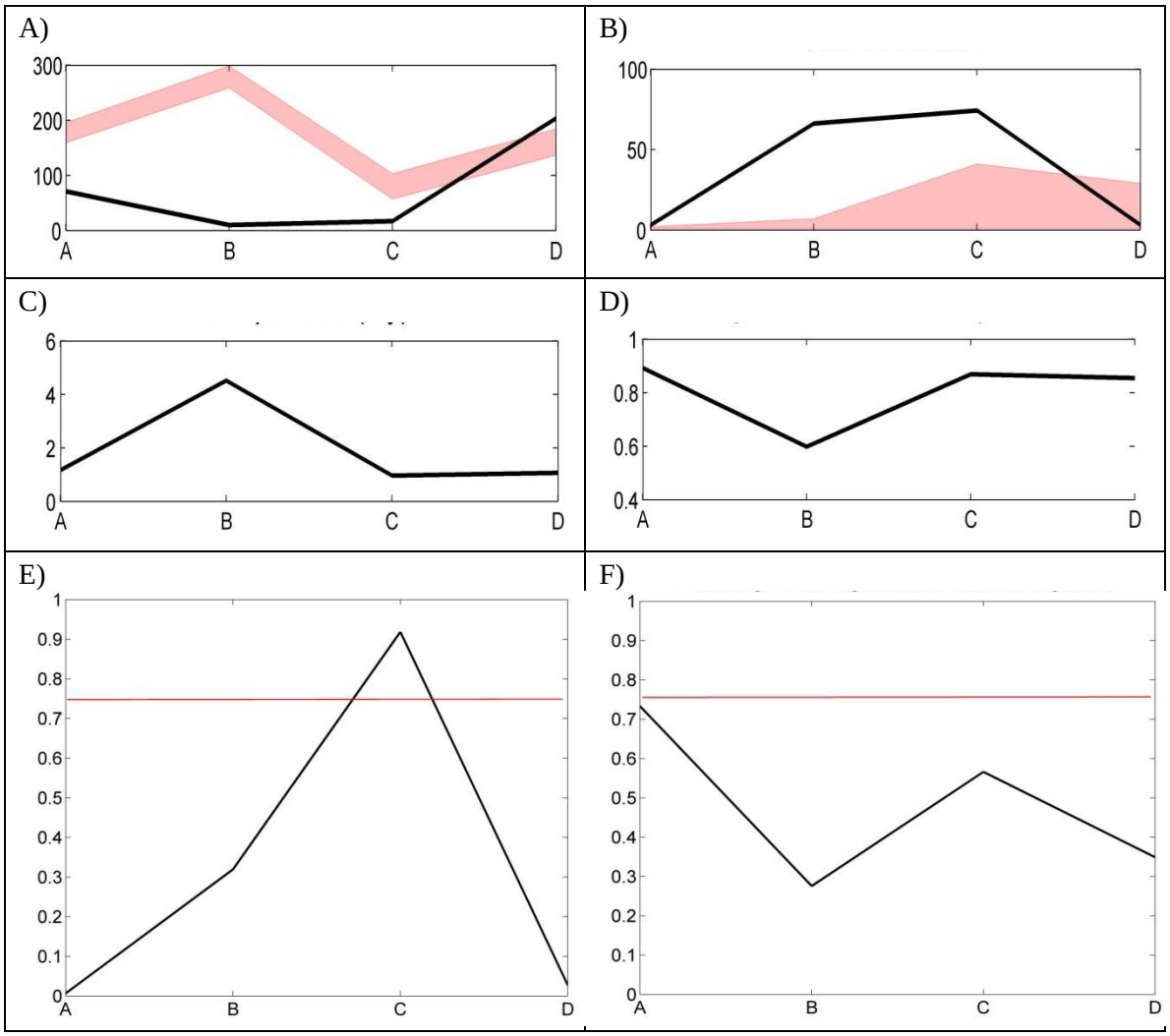


Figure 3.10 Statistical analysis of the vegetation clusters for the sample patterns shown in Figure 3.9. A) Number of clusters and range 0.025-0.975 quantile of the number of clusters corresponding to a binomial process with the same percentage cover. B) Mean size of clusters and range 0.025-0.975 quantile of the mean cluster size corresponding to a binomial process with the same percentage cover. C) Mean ratio between the span in the x-direction and the span in the y-direction of each cluster; red line represents the boundary between spots and bands. D) Mean fraction of area filled by each cluster. E) Percentage of the domain filled by the largest cluster; red line represents the minimum fraction of area filled for the field to be considered labyrinthine. F) Fraction of area of the largest cluster filled by vegetation; red line represents the maximum fraction of the largest cluster that can be filled for the field to be considered labyrinthine.

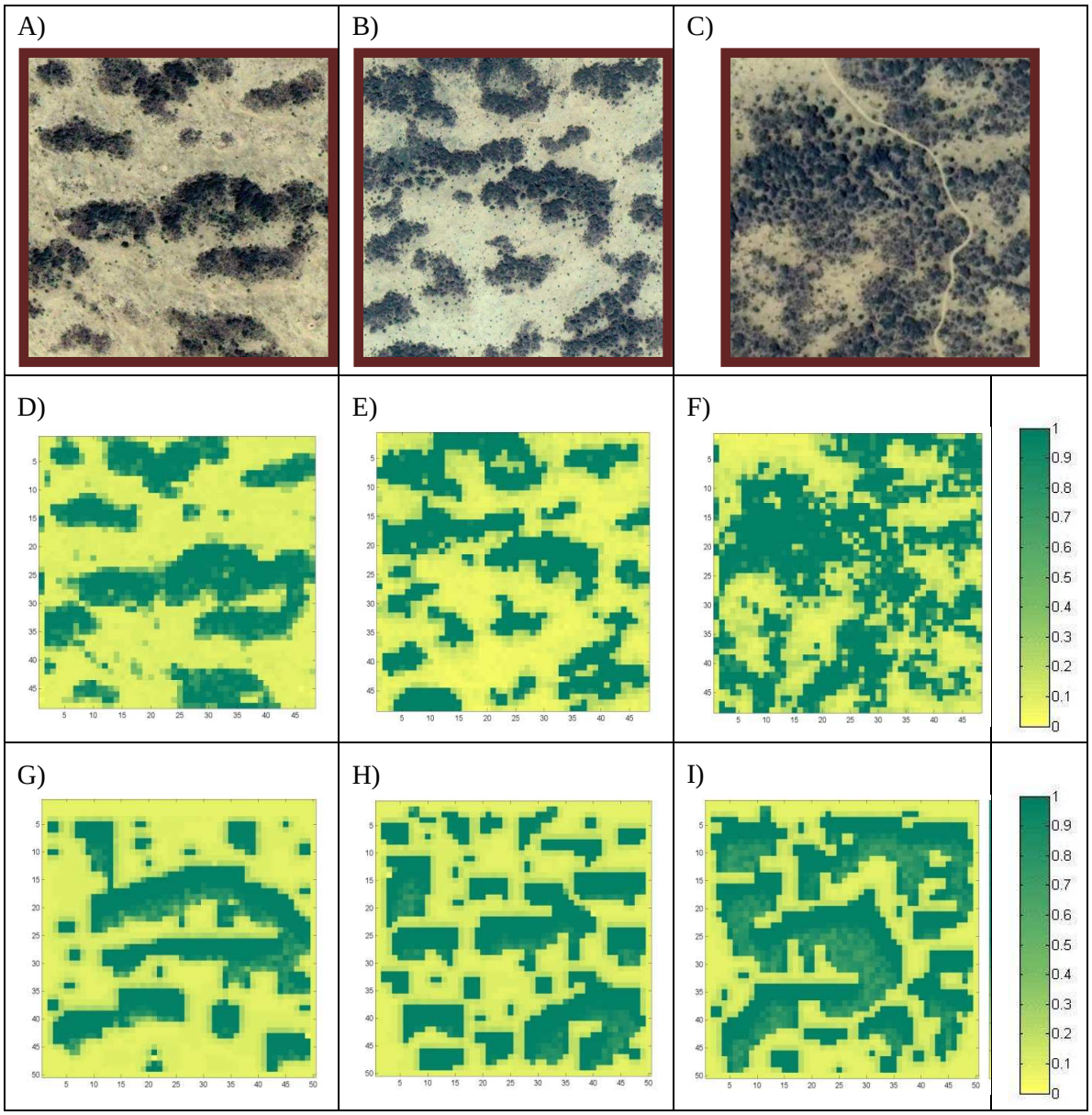


Figure 3.11 A), B), C) Aerial photographs of natural patterns (Tiger Bushes) in Niger (13°20´N, 2°04´E). D), E), F) Digitized representation of the vegetation covers of the natural patterns corresponding respectively to A), B) and C). G), H), I) Simulated patterns.

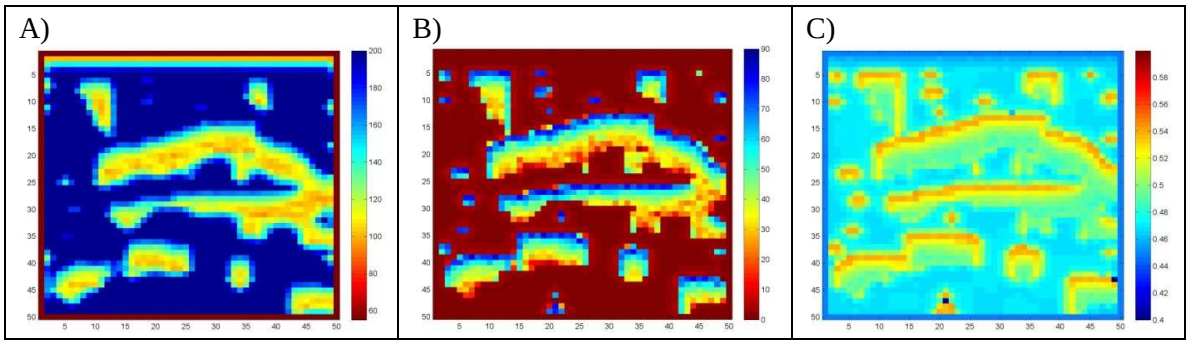


Figure 3.12 A) Ratio between the effective amount of water received and annual precipitation for the simulated pattern in Figure 3.11g. B) Long term groundwater runoff on the study domain for the simulated pattern in Figure 3.11g. C) Long term relative soil saturation during the rainy season on the study domain for the simulated pattern in Figure 3.11g.

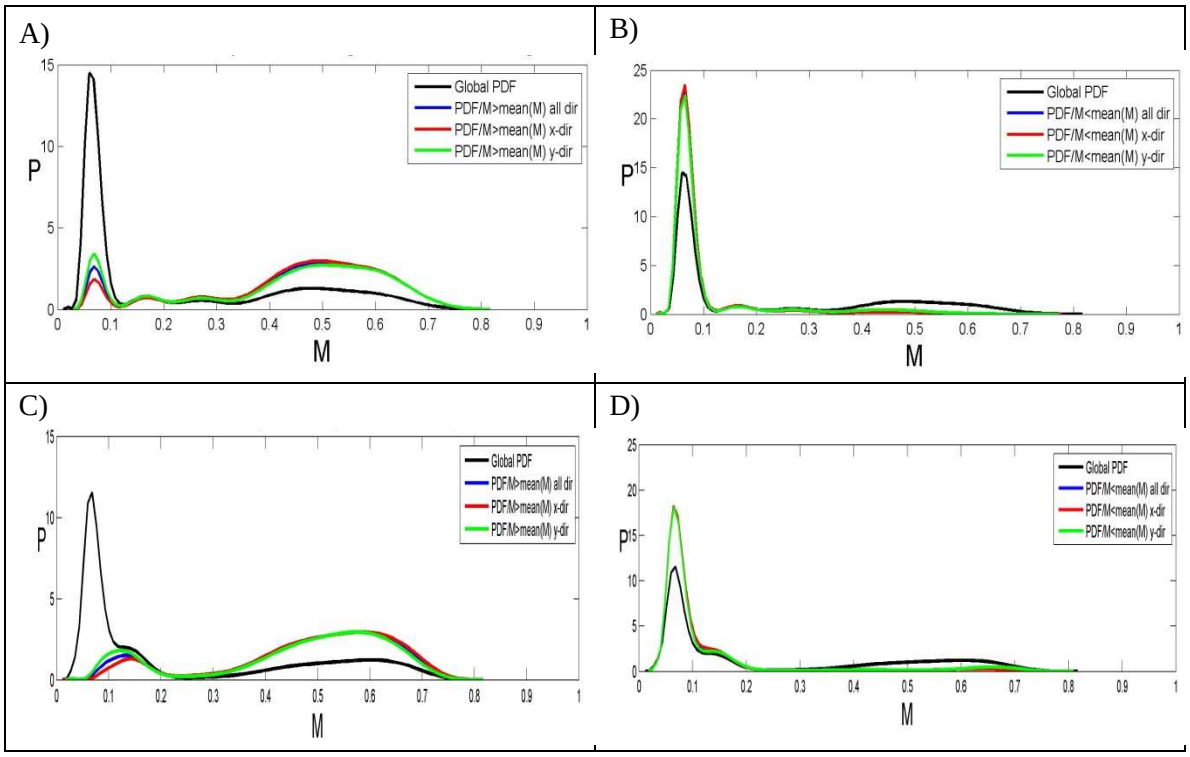


Figure 3.13 A) PDFs of vegetation coverage for the natural pattern shown in Figure 3.11a (as digitized in Figure 3.11d): global PDF (black line) and PDFs of vegetation coverage conditioned on having a neighbor pixel with vegetation coverage higher than the global average for all directions (blue line), x-direction (red line) and y-direction (green line). B) PDFs of vegetation coverage for the natural pattern shown in Figure 3.11a (as digitized in Figure 3.11d): global PDF (black line) and PDFs of vegetation coverage conditioned on having a neighbor pixel with vegetation coverage lower than the global average for all directions (blue line), x-direction (red line) and y-direction (green line). C) PDFs of vegetation coverage for the simulated pattern shown in Figure 3.11g: global PDF (black line) and PDFs of vegetation coverage conditioned on having a neighbor pixel with vegetation coverage higher than the global average for all directions (blue line), x-direction (red line) and y-direction (green line). B) PDFs of vegetation coverage for the simulated pattern shown in Figure 3.11g: global PDF (black line) and PDFs of vegetation coverage conditioned on having a neighbor pixel with vegetation coverage lower than the global average for all directions (blue line), x-direction (red line) and y-direction (green line).

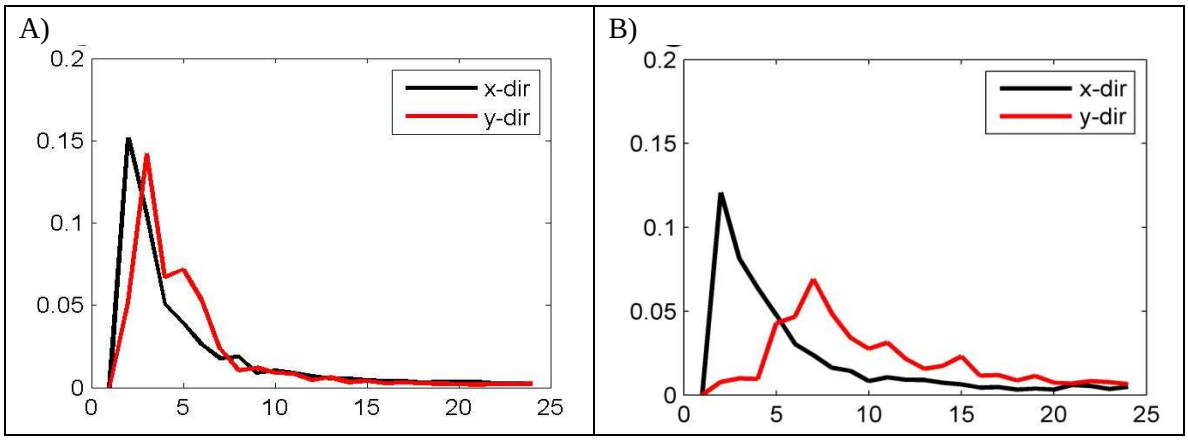


Figure 3.14 Power spectral densities of the vegetation density field; frequencies (in the horizontal axis) are expressed in terms of number of wavelengths present within the domain length and width. A) Average of the 1D power spectral densities of the vegetation coverage for the natural pattern shown in Figure 3.11a along the x-direction (black line) and y-direction (red line). B) Average of the 1D power spectral densities of the vegetation coverage for the simulated pattern shown in Figure 3.11g along the x-direction (black line) and y-direction (red line).

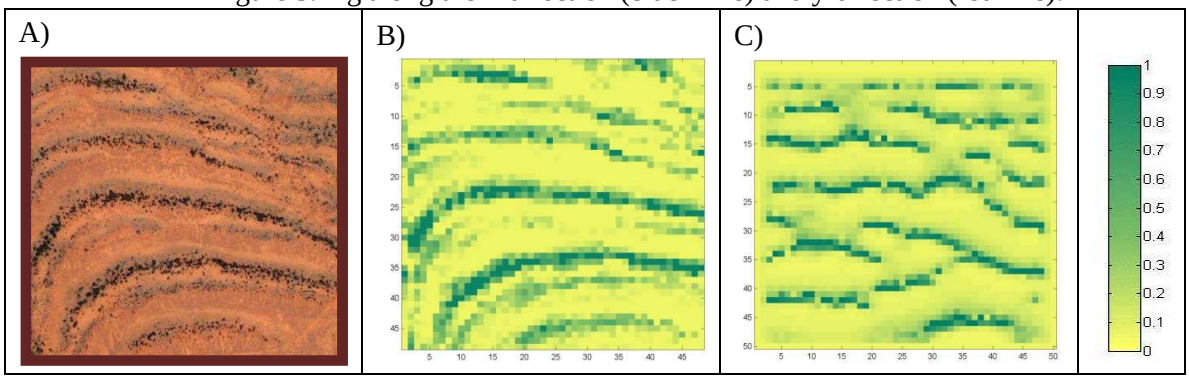


Figure 3.15 A) Aerial photograph of a natural vegetation pattern in Somalia (7°43´N, 48°02´E); B) Digitized representation of the vegetation cover of the natural pattern in A). C) Sample simulated pattern.

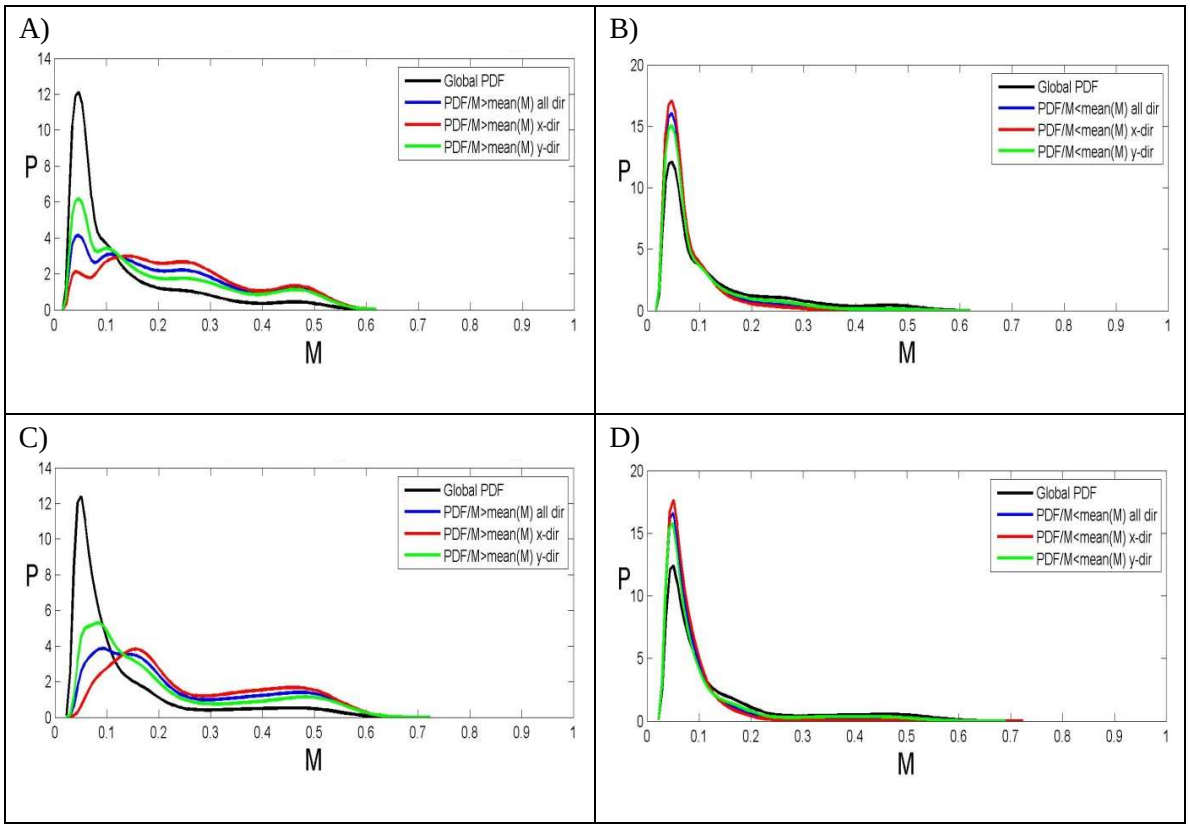


Figure 3.16 A) PDFs of vegetation coverage for the natural pattern shown in Figure 3.15a (as digitized in Figure 3.15b): global PDF (black line) and PDFs of vegetation coverage conditioned on having a neighbor pixel with vegetation coverage higher than the global average for all directions (blue line), x-direction (red line) and y-direction (green line). B) PDFs of vegetation coverage for the natural pattern shown in Figure 3.15a (as digitized in Figure 3.15b): global PDF (black line) and PDFs of vegetation coverage conditioned on having a neighbour pixel with vegetation coverage lower than the global average for all directions (blue line), x-direction (red line) and y-direction (green line). C) PDFs of vegetation coverage for the simulated pattern shown in Figure 3.15a: global PDF (black line) and PDFs of vegetation coverage conditioned on having a neighbour pixel with vegetation coverage higher than the global average for all directions (blue line), x-direction (red line) and y-direction (green line). B) PDFs of vegetation coverage for the simulated pattern shown in Figure 3.15a: global PDF (black line) and PDFs of vegetation coverage conditioned on having a neighbour pixel with vegetation coverage lower than the global average for all directions (blue line), x-direction (red line) and y-direction (green line).

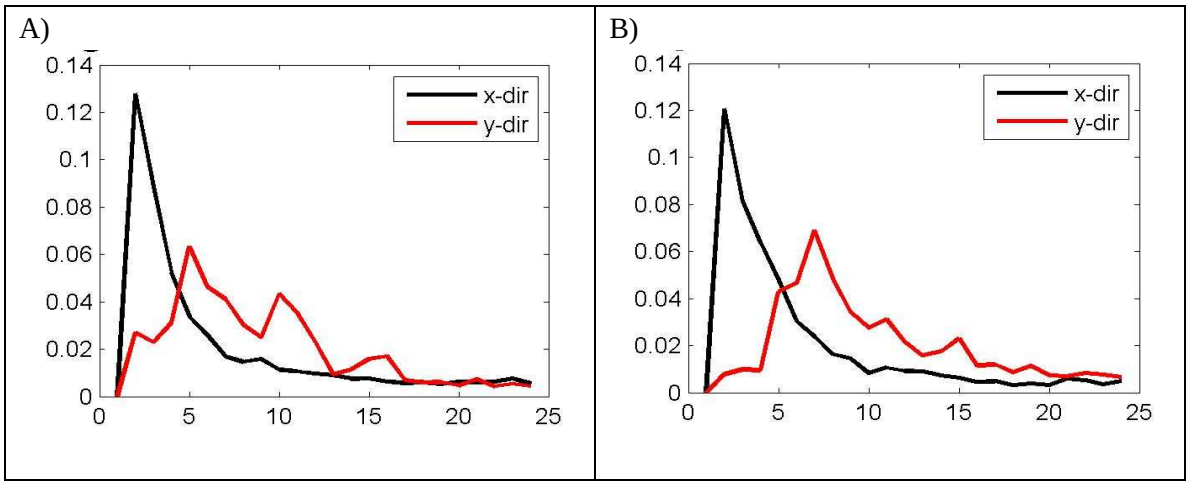


Figure 3.17 Power spectral densities of the vegetation density field; frequencies (in the horizontal axis) are expressed in terms of number of wavelengths present within the domain length and width. A) Average of the 1D power spectral densities of the vegetation coverage for the natural pattern shown in Figure 3.15a along the x-direction (black line) and y-direction (red line). B) Average of the 1D power spectral densities of the vegetation coverage for the simulated pattern shown in Figure 3.15c along the x-direction (black line) and y-direction (red line).

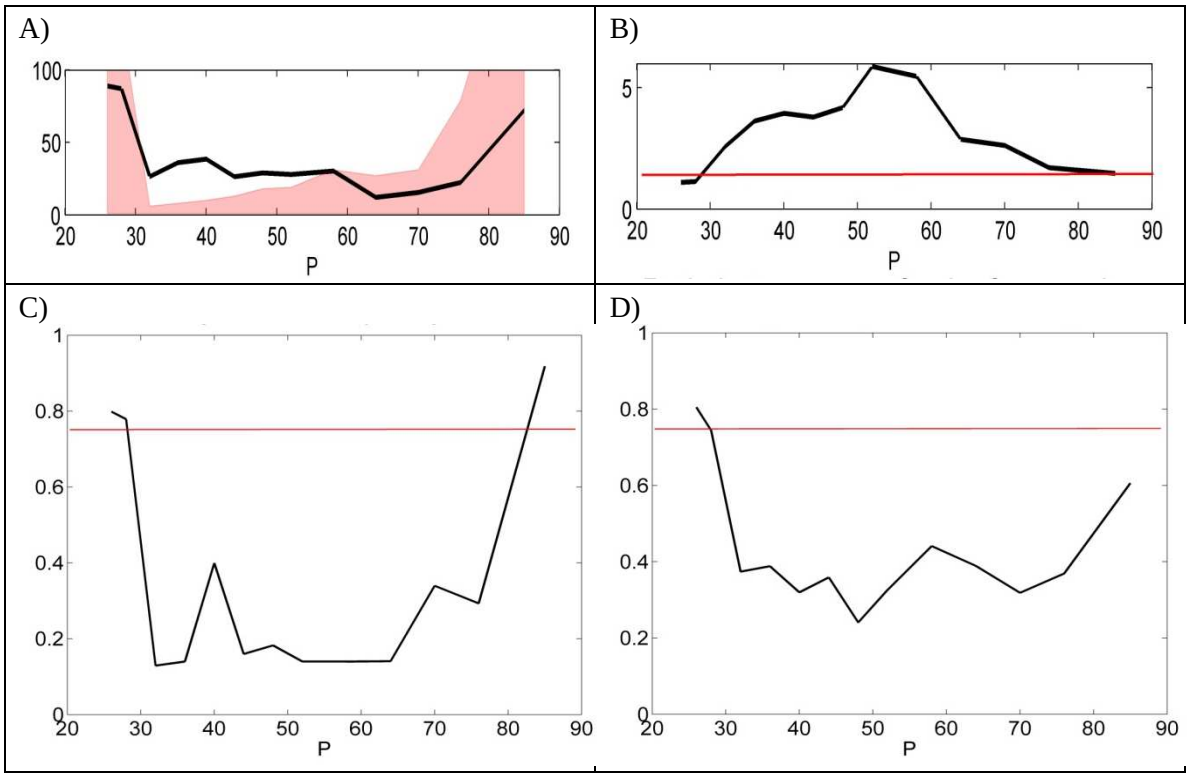


Figure 3.18 Statistical analysis of the vegetation clusters for the fields of vegetation obtained by varying annual precipitation. A) Mean size of clusters and range 0.025-0.975 quantile of the mean cluster size corresponding to a binomial process with the same percentage cover. B) Mean ratio between the span in the x-direction and the span in the y-direction of each cluster; red line represents the boundary between spots and bands. C) Percentage of the domain filled by the largest cluster; red line represents the minimum fraction of area filled for the field to be considered labyrinthine. D) Fraction of area of the largest cluster filled by vegetation; red line represents the maximum fraction of the largest cluster that can be filled for the field to be considered labyrinthine.

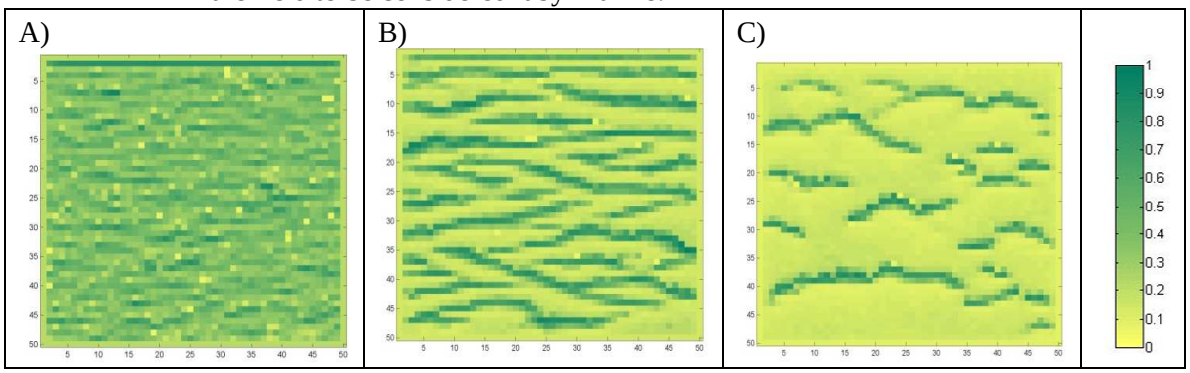


Figure 3.19 Effect of precipitation on pattern formation. A) P=70 cm/yr; B) P=48cm/yr; C) P=32 cm/yr.

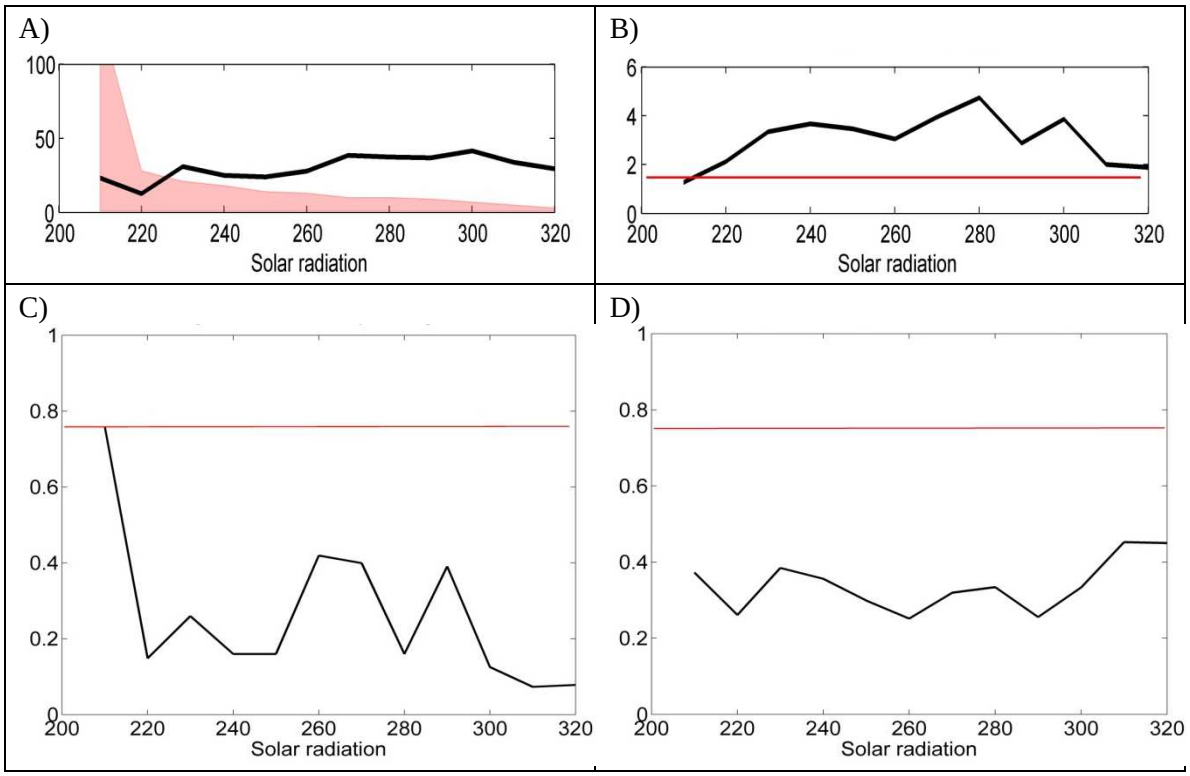


Figure 3.20 Statistical analysis of the vegetation clusters for the fields of vegetation obtained by varying average solar radiation. A) Mean size of clusters and range 0.025-0.975 quantile of the mean cluster size corresponding to a binomial process with the same percentage cover. B) Mean ratio between the span in the x-direction and the span in the y-direction of each cluster; red line represents the boundary between spots and bands. C) Percentage of the domain filled by the largest cluster; red line represents the minimum fraction of area filled for the field to be considered labyrinthine. D) Fraction of area of the largest cluster filled by vegetation; red line represents the maximum fraction of the largest cluster that can be filled for the field to be considered labyrinthine.

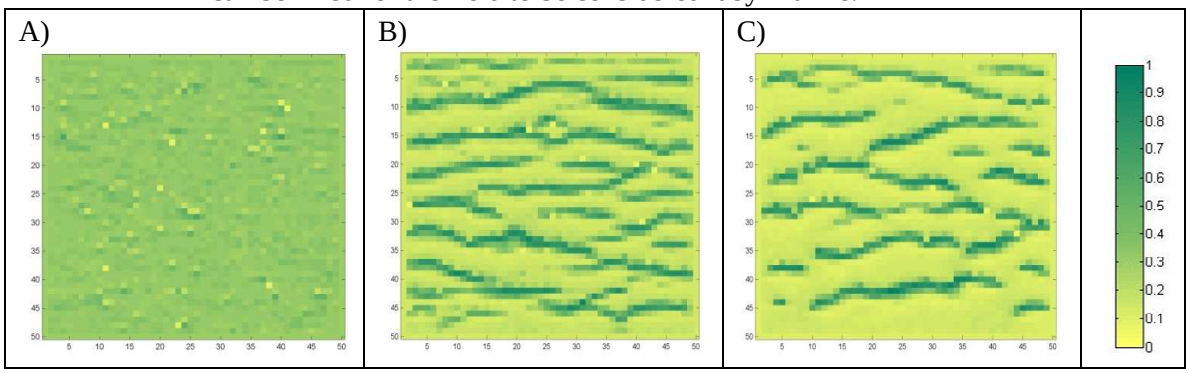


Figure 3.21 Effect of solar radiation on pattern formation. A) Solar radiation=210 W/m^2; B) Solar radiation=250 W/m^2; C) Solar radiation=290 W/m^2.

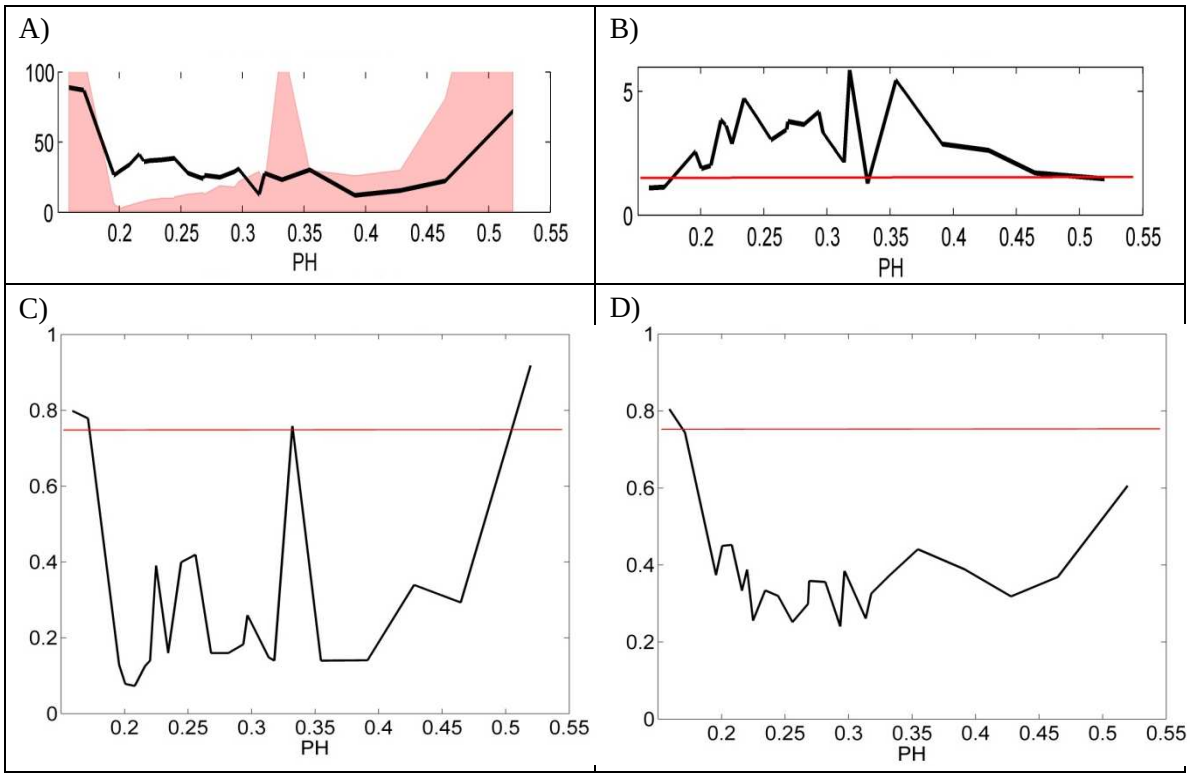


Figure 3.22 Statistical analysis of the vegetation clusters for the fields of vegetation as a function of the ratio between annual precipitation and annual potential evapotranspiration. A) Mean size of clusters and range 0.025-0.975 quantile of the mean cluster size corresponding to a binomial process with the same percentage cover. B) Mean ratio between the span in the x-direction and the span in the y-direction of each cluster; red line represents the boundary between spots and bands. C) Percentage of the domain filled by the largest cluster; red line represents the minimum fraction of area filled for the field to be considered labyrinthine. D) Fraction of area of the largest cluster filled by vegetation; red line represents the maximum fraction of the largest cluster that can be filled for the field to be considered labyrinthine.

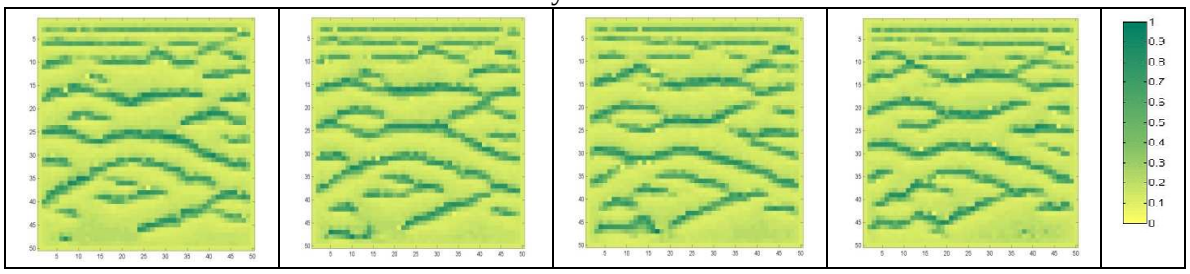


Figure 3.23 Pattern migration. The evolution of a pattern is tracked through iteration steps: A) 45; B) 50; C) 55; D) 60.

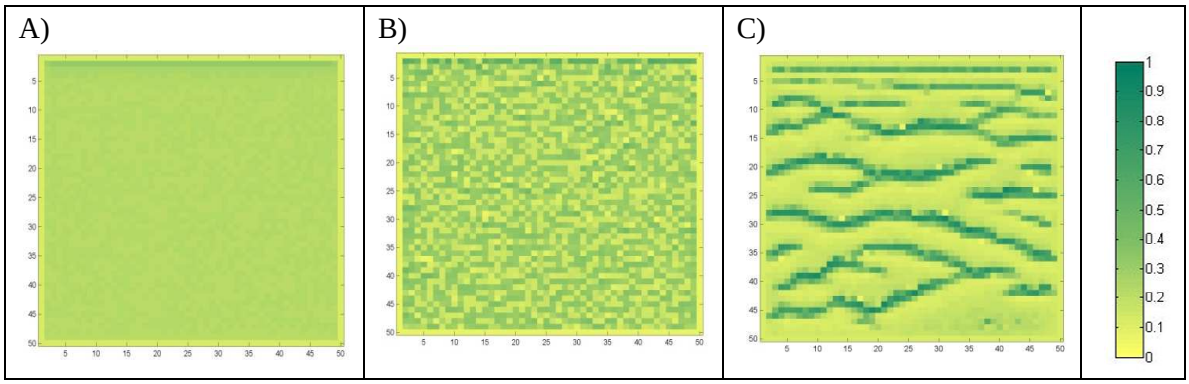


Figure 3.24 Effect of hydraulic conductivity on pattern formation. A) Uniform hydraulic conductivity. B) Hydraulic conductivity randomly variable in space. C) Hydraulic conductivity variable in space as a function of vegetation coverage.

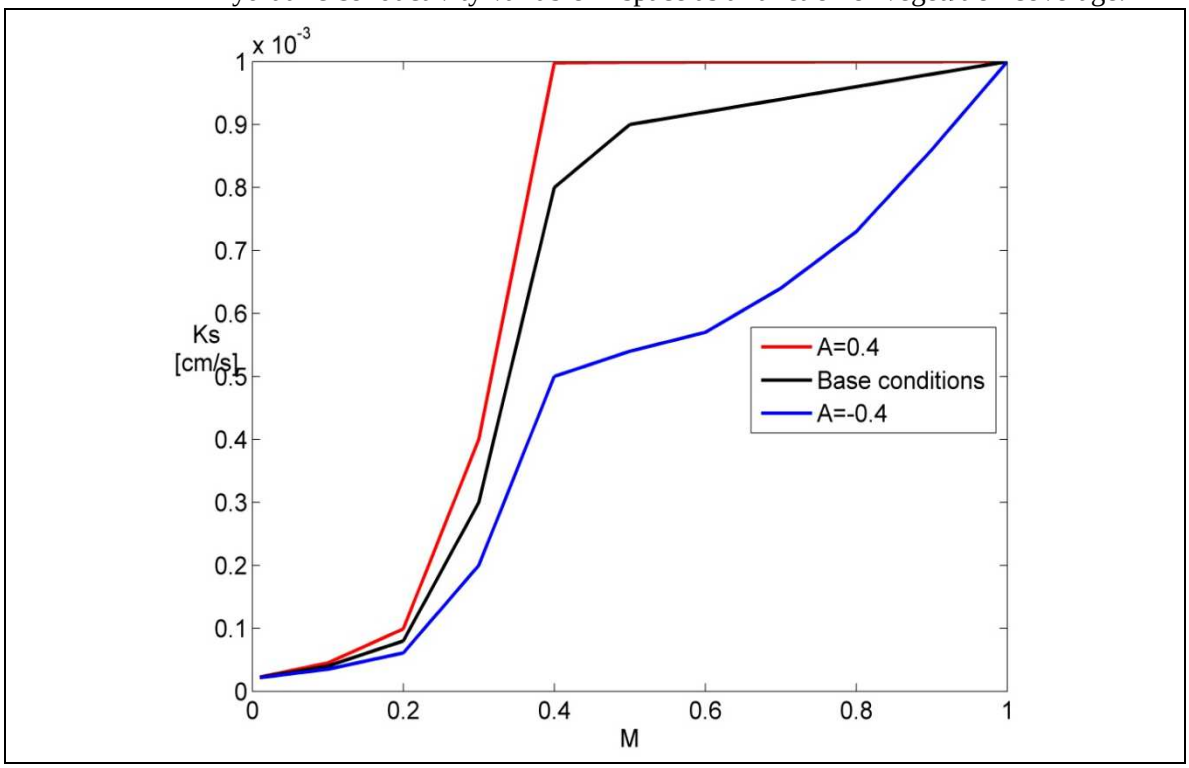


Figure 3.25 Hydraulic conductivity functions (without the random component) for the base conditions (black), compared with the functions obtained through equation (3.14) with coefficients A=0.4 (red) and A=-0.4 (blue).

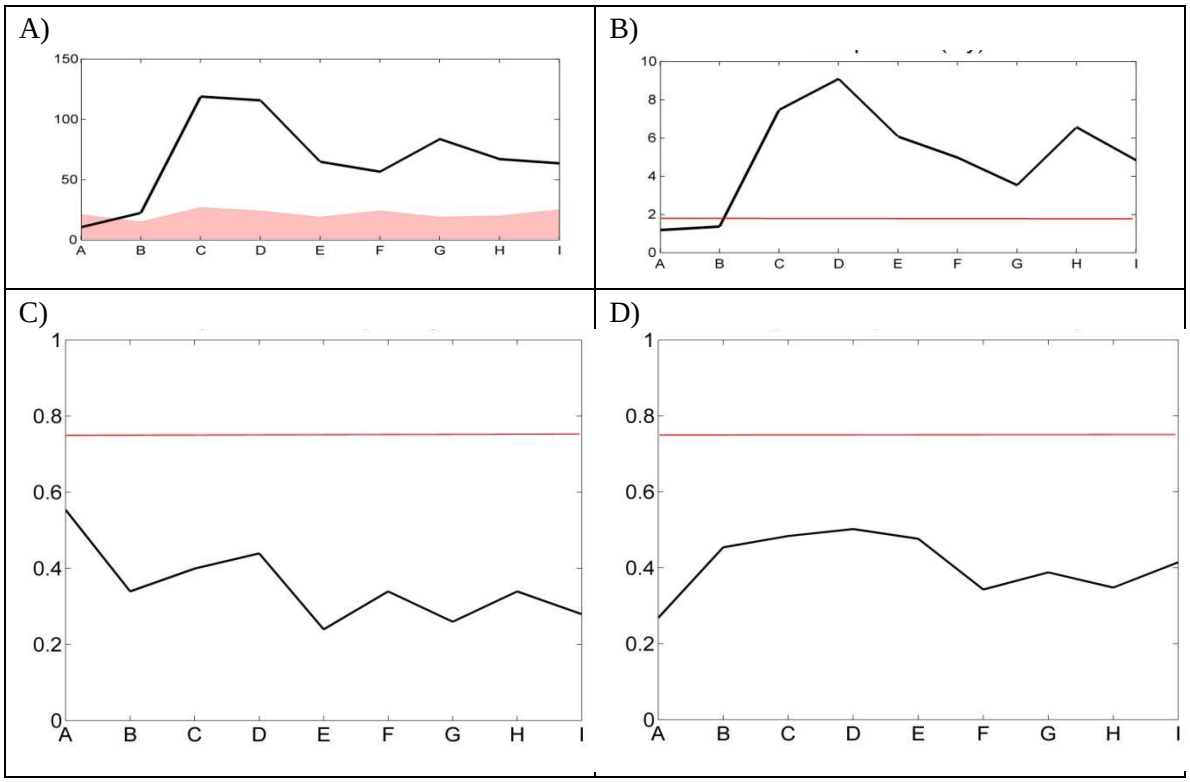


Figure 3.26 Statistical analysis of the vegetation clusters for the fields of vegetation obtained with hydraulic conductivity functions modified from the base conditions through equation (3.14). Labels A through I on the horizontal axis correspond to values for coefficient A of equation (3.14) of respectively: -0.4, -0.3, -0.2, -0.1, 0 (that is, base conditions), 0.1, 0.2, 0.3 and 0.4. A) Mean size of clusters and range 0.025-0.975 quantile of the mean cluster size corresponding to a binomial process with the same percentage cover. B) Mean ratio between the span in the x-direction and the span in the y-direction of each cluster; red line represents the boundary between spots and bands. C) Percentage of the domain filled by the largest cluster; red line represents the minimum fraction of area filled for the field to be considered labyrinthine. D) Fraction of area of the largest cluster filled by vegetation; red line represents the maximum fraction of the largest cluster that can be filled for the field to be considered labyrinthine.

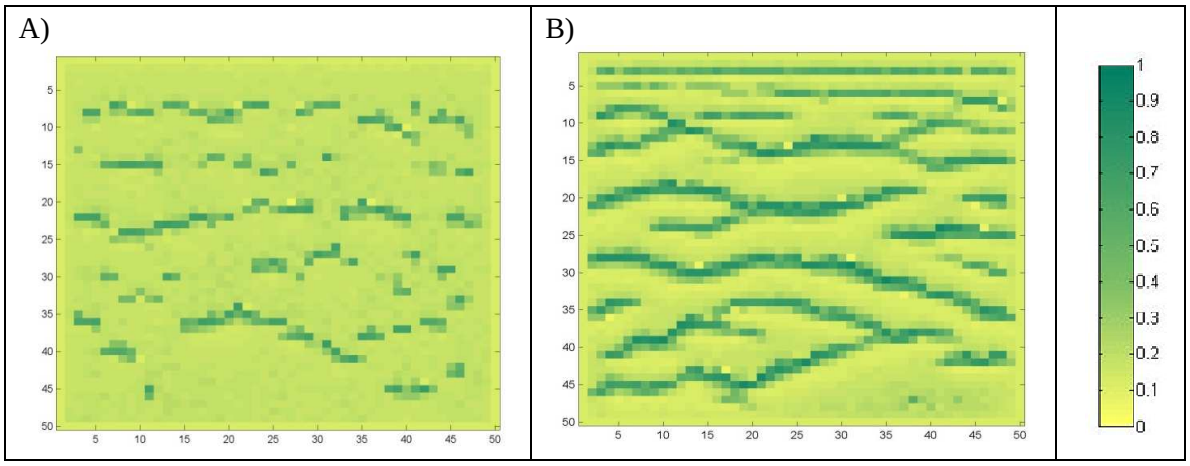


Figure 3.27 Effect of local plants interactions on pattern formation. A) No interactions. B) Spatially variable (function of neighbours' vegetation and surface runoff) interactions.

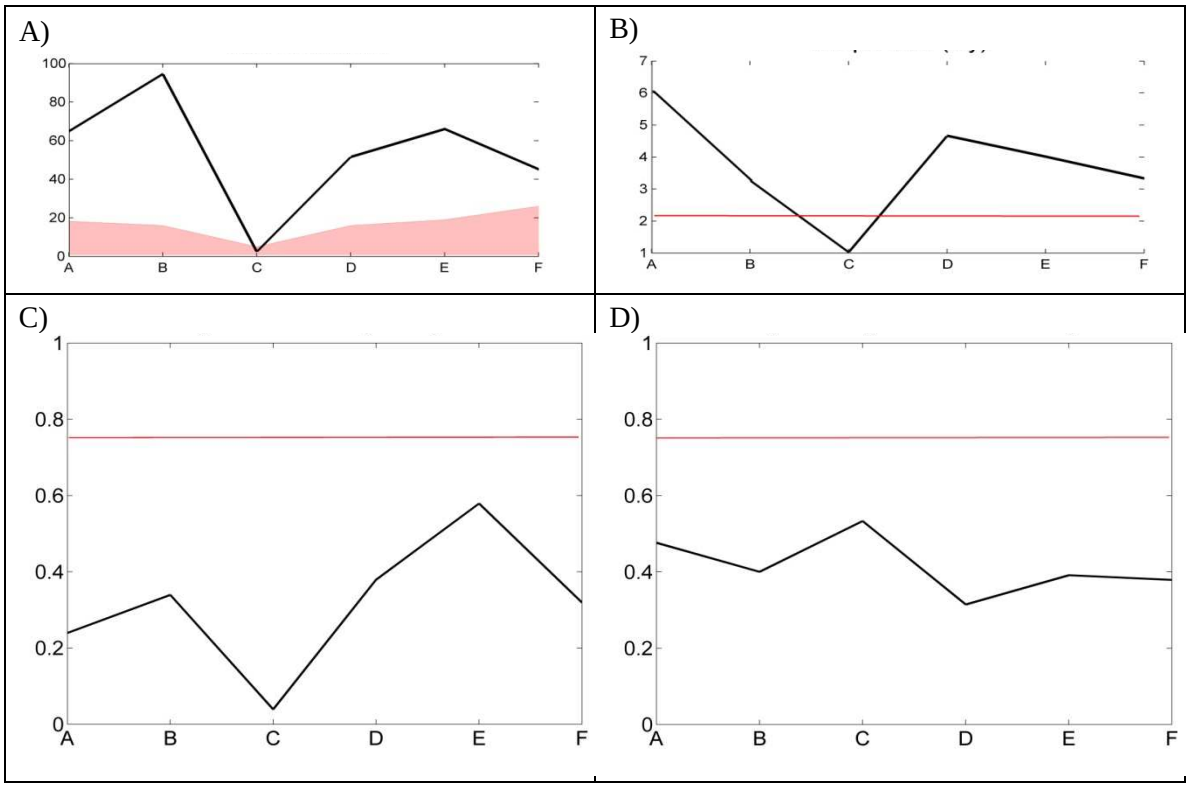


Figure 3.28 Statistical analysis of the vegetation clusters for the fields of vegetation obtained by setting one of the coefficients α_i of equations (3.8) through (3.12) equal to zero. Labels A through F on the horizontal axis correspond, respectively to: base conditions, α_1=0, α_2=0, α_3=0, α_4=0 and α_5=0. A) Mean size of clusters and range 0.025-0.975 quantile of the mean cluster size corresponding to a binomial process with the same percentage cover. B) Mean ratio between the span in the x-direction and the span in the y-direction of each cluster; red line represents the boundary between spots and bands. C) Percentage of the domain filled by the largest cluster; red line represents the minimum fraction of area filled for the field to be considered labyrinthine. D) Fraction of area of the largest cluster filled by vegetation; red line represents the maximum fraction of the largest cluster that can be filled for the field to be considered labyrinthine.

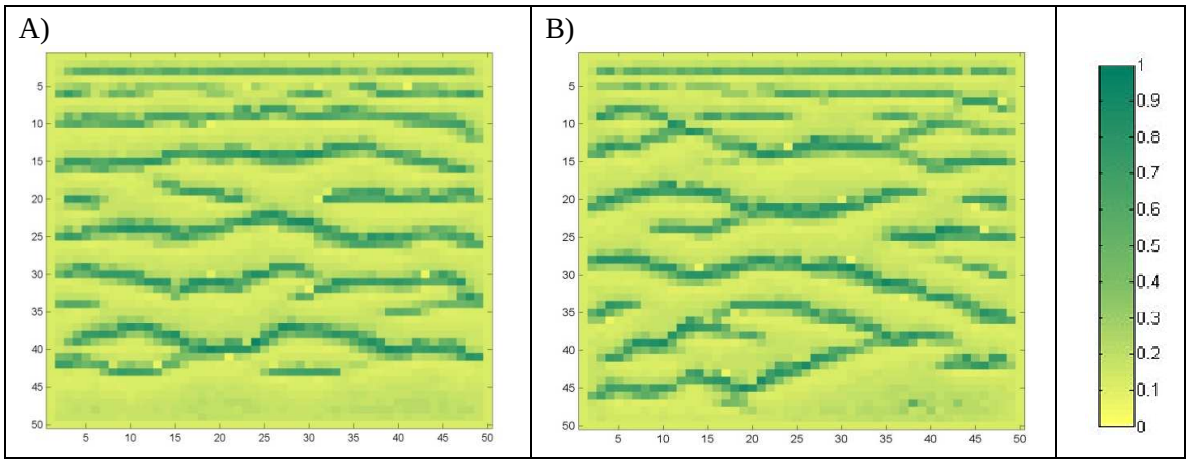


Figure 3.29 Effect of soil moisture redistribution due to roots on pattern formation. A) No roots redistribution. B) Roots are able to reroute the soil moisture from adjacent less vegetated areas.

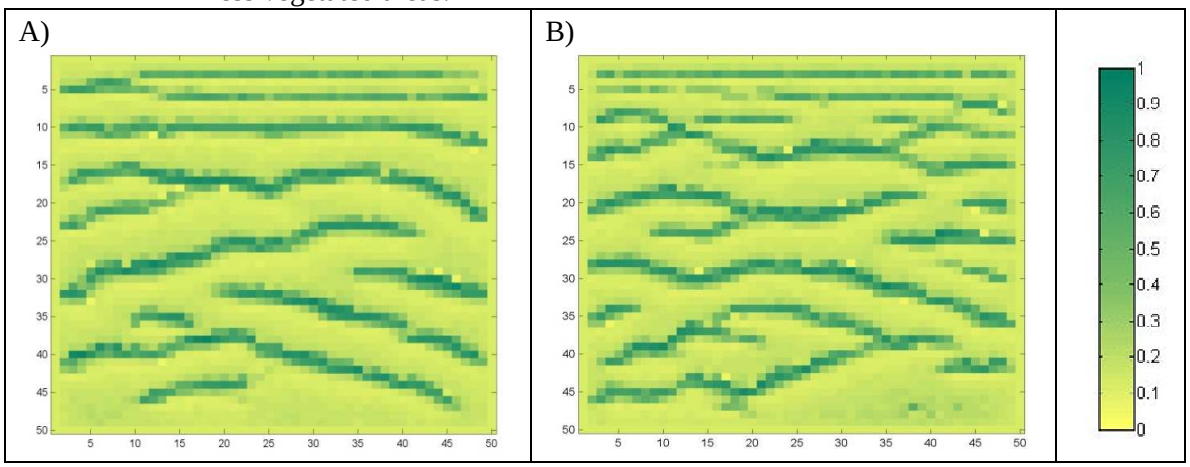


Figure 3.30 Effect of spatially variable reflectance on pattern formation. A) Uniform reflectance, that is, constant albedo. B) Spatially variable albedo.

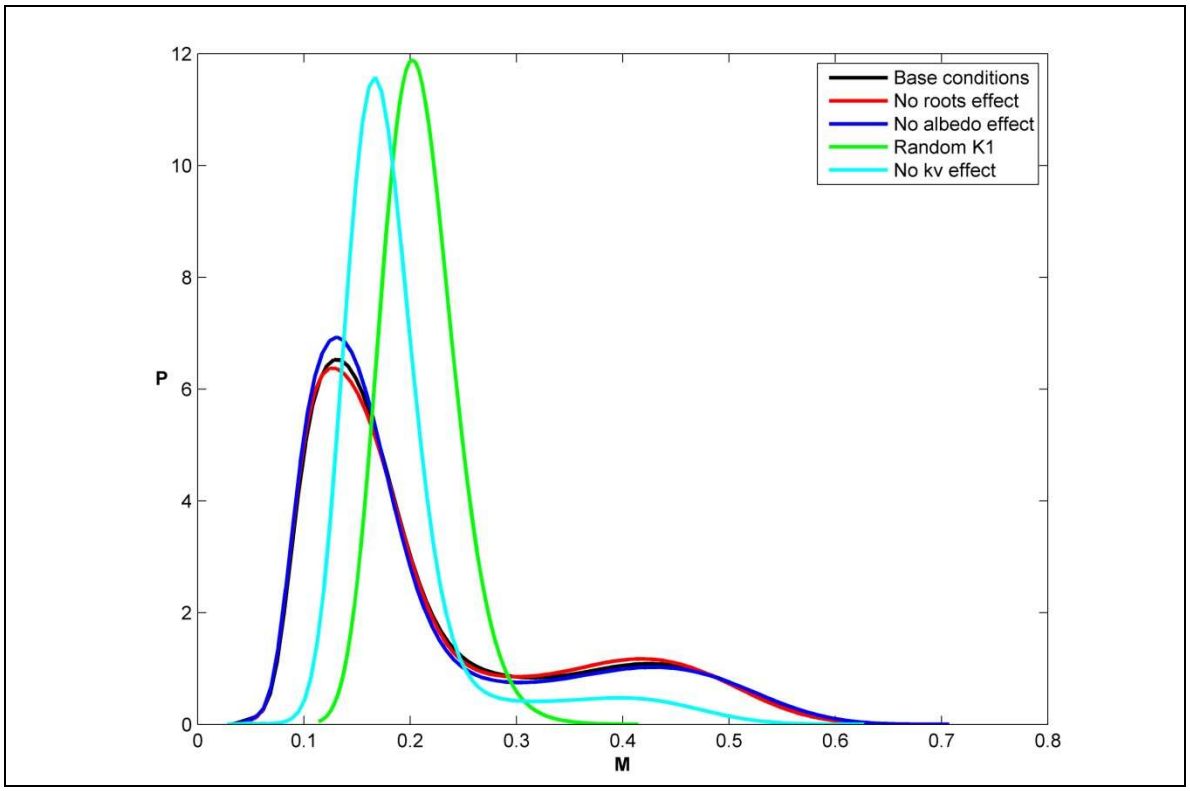


Figure 3.31 PDFs of the vegetation coverage for the cases of: base conditions (black); no water redistribution due to roots (red); no spatially variable albedo (blue); hydraulic conductivity randomly distributed in space and independent from the vegetal coverage (green); no lateral facilitation and competition effects through transpiration efficiency (cyan).

Table 3.1 Climate, soil and vegetation properties of the system.

	Niger tiger bushes	Somalia bands	Base conditions
Soil			
Total soil porosity [-]	0.4	0.42	0.4
Pore size distribution index [-]	0.4	0.42	0.4
Surface retention capacity [cm]	0.1	0.1	0.1
Matric potential at effective saturation [cm]	20	14	25
Climate			
Mean storm duration [days]	0.05	0.2	0.2
Mean time between storms [days]	10	10	10
Length of the rainy season [days]	150	150	200
Parameter of the gamma distribution of storm depth [-]	0.6	0.6	0.6
Surface temperature [°K]	300	300	300
Screen height temperature [°K]	300	300	300
Specific humidity [-]	0.01	0.01	0.01
Cloud fractional coverage [-]	0.0	0.1	0.1
Surface solar radiation [W/m2]	280	260	270
Mean precipitation [cm]	56	16	40
Vegetation			
Base value of transpiration efficiency, $\overline{k_v}$	0.75	0.75	0.75

Table 3.2 Model parameters.

	Niger tiger bushes			Somalia bands	Base conditions
	A	B	C		
a_1 [cm/s]	$3\cdot10^{-5}$	$3\cdot10^{-5}$	$3\cdot10^{-5}$	$9\cdot10^{-6}$	$2\cdot10^{-5}$
a_2 [cm/s]	$5\cdot10^{-5}$	$5\cdot10^{-5}$	$5\cdot10^{-5}$	$1\cdot10^{-5}$	$4\cdot10^{-5}$
a_3 [cm/s]	$8\cdot10^{-5}$	$7\cdot10^{-5}$	$7\cdot10^{-5}$	$3\cdot10^{-5}$	$8\cdot10^{-5}$
a_4 [cm/s]	$3\cdot10^{-4}$	$4\cdot10^{-4}$	$4.5\cdot10^{-4}$	$1\cdot10^{-4}$	$3\cdot10^{-4}$
a_5 [cm/s]	$6\cdot10^{-4}$	$6.5\cdot10^{-4}$	$6.8\cdot10^{-4}$	$7\cdot10^{-4}$	$8\cdot10^{-4}$
a_6 [cm/s]	$8\cdot10^{-4}$	$8\cdot10^{-4}$	$8.5\cdot10^{-4}$	$9\cdot10^{-4}$	$9\cdot10^{-4}$
a_7 [cm/s]	$8.2\cdot10^{-4}$	$8.2\cdot10^{-4}$	$8.7\cdot10^{-4}$	$9.1\cdot10^{-4}$	$9.2\cdot10^{-4}$
a_8 [cm/s]	$8.4\cdot10^{-4}$	$8.4\cdot10^{-4}$	$8.9\cdot10^{-4}$	$9.2\cdot10^{-4}$	$9.4\cdot10^{-4}$
a_9 [cm/s]	$8.6\cdot10^{-4}$	$8.6\cdot10^{-4}$	$9.1\cdot10^{-4}$	$9.3\cdot10^{-4}$	$9.6\cdot10^{-4}$
a_{10} [cm/s]	$8.8\cdot10^{-4}$	$8.8\cdot10^{-4}$	$9.3\cdot10^{-4}$	$9.4\cdot10^{-4}$	$9.8\cdot10^{-4}$
b_1 [cm/s]	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$1\cdot10^{-5}$	$2\cdot10^{-4}$
b_2 [cm/s]	$3\cdot10^{-4}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$4\cdot10^{-4}$
b_3 [cm/s]	$2.8\cdot10^{-3}$	$3.3\cdot10^{-3}$	$3.8\cdot10^{-3}$	$7\cdot10^{-4}$	$2.2\cdot10^{-3}$
b_4 [cm/s]	$3\cdot10^{-3}$	$2.5\cdot10^{-3}$	$2.3\cdot10^{-3}$	$6\cdot10^{-3}$	$5\cdot10^{-3}$
b_5 [cm/s]	$2\cdot10^{-4}$	$1.5\cdot10^{-3}$	$1.7\cdot10^{-3}$	$2\cdot10^{-3}$	$1\cdot10^{-3}$
b_6 [cm/s]	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$1\cdot10^{-4}$	$2\cdot10^{-4}$
b_7 [cm/s]	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$1\cdot10^{-4}$	$2\cdot10^{-4}$
b_8 [cm/s]	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$1\cdot10^{-4}$	$2\cdot10^{-4}$
b_9 [cm/s]	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$1\cdot10^{-4}$	$2\cdot10^{-4}$
b_{10} [cm/s]	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$2\cdot10^{-4}$	$1\cdot10^{-4}$	$2\cdot10^{-4}$
α_1	-0.01	-0.01	-0.01	-0.03	-0.01
α_2	-0.019	-0.02	-0.018	-0.05	-0.02
α_3	-0.03	-0.03	-0.03	-0.04	-0.03
α_4	-0.03	-0.03	-0.03	-0.04	-0.03
α_5	-0.09	-0.09	-0.09	-0.08	-0.09
ρ_R	1.5	1.5	1.5	1.25	1.5

Table 3.3 Cluster analysis for natural and simulated patterns in Niger and Somalia.

	Niger						Somalia	
	Natural			Simulated			Natural	Simulated
FIGURE	9D	9E	9F	9G	9H	9I	12B	12C
Number of clusters	21	20	29	22	30	14	49	24
Average cluster size	42	38	38	37	32	90	15	28
Range 0.025-0.975 quantile of cluster size for binomial process	[1-18]	[1-12]	[1-35]	[1-14]	[1-24]	[1-74]	[1-12]	[1-9]
Shape ratio	1.4	1.5	1.1	1.2	1.4	1.2	1.9	2.6
Percentage of domain filled by the largest cluster	0.35	0.34	0.78	0.57	0.18	0.68	0.36	0.15
Fraction of area of the largest cluster filled with vegetation	0.51	0.35	0.37	0.42	0.63	0.44	0.21	0.23

Appendix A Water yield

A.1 Overview

Water supply begins with water yield. To estimate water yield we adopted a water balance model proposed by Eagleson (1978a). Eagleson's model is a mechanistic representation of the water dynamics occurring across the soil-atmosphere interface as a result of a stochastic climatic input. Input to the model includes soil hydraulic properties and statistics describing the stochastic climate (storm statistics), with climate represented by the probability distributions functions of precipitation and potential evapotranspiration. The model generates probability distribution functions of water fluxes (actual evapotranspiration, surface runoff and groundwater runoff) as output. The water balance model was calibrated in order to match historical streamflow records (for years 1953-2005) and then applied to all locations in the U.S. under current and potential future climatic conditions, providing estimates for future years 2006-2090. Since the model is a lumped representation of the annual water balance, we subdivided the U.S. territory into a study grid of 5x5 km cells, and estimated water fluxes in each cell.

This chapter is organized as follows: section A.2 describes the water balance model; section A.4 is dedicated to the description of the area of analysis and parameters of the model; section A.5 describes the model calibration procedure; section A.6 deals with the extension of the calibrated parameters to the whole conterminous U.S.; section A.7 describes the model input for future climatic and socio-economic scenarios.

A.2 Problem Formulation

Eagleson's model is a one-dimensional representation of soil moisture dynamics as forced by a stochastic climate (Eagleson, 1978a-g). The model describes the relationships between annual amounts of precipitation, runoff, infiltration and evapotranspiration as a function of volumetric soil moisture and soil and vegetation characteristics. The description is physically based and only accounts for processes operating in the vertical direction, across the soil-atmosphere interface.

The water balance equation for the control volume (Fig. A.1) is written as follows:

$$\int_0^t \left\{ i(t) - e_T(t) - \frac{\partial}{\partial t}\left[V_{ss}(t) - V_{rg}(t)\right] \right\} \cdot dt = \int_0^t \left[r_s(t) + r_g(t)\right] \cdot dt \qquad \text{(A.1)}$$

where $i(t)$, $e_T(t)$, $V_{ss}(t)$, $V_{rg}(t)$, $r_s(t)$ and $r_g(t)$ are respectively: storm intensity, the evapotranspiration rate, volume of water storage in the surface, volume of water storage in the ground, surface runoff rate and the groundwater runoff rate. Snow, ice and movements of soil moisture as vapor are not considered.

The integration of (A.1) is very difficult for an arbitrary period of time, because of several reasons including: 1) the climatic forcing (i.e., precipitation and potential evapotranspiration) is stochastic, 2) all terms in (A.1) depend on the soil moisture content, which is difficult to evaluate or measure, and 3) the integration requires that carry-over storage be evaluated. In order to obtain an analytical solution to the water balance equation, it is then assumed that the system is in equilibrium with the climate in its mean value. This implies that the long-term mean of moisture storage (above and below the surface) is constant, that is, the long-term average of the change in total soil moisture storage is zero, thereby avoiding the need to compute carry-over storage. Therefore, taking the expectation of (A.1) under the assumption of system in equilibrium with the climate leads to:

$$E\left[P_A\right] - E\left[E_{T_A}\right] = E\left[R_{s_A}\right] + E\left[R_{g_A}\right] \qquad \text{(A.2)}$$

where $E[P_A]$, $E[E_{TA}]$, $E[R_{SA}]$ and $E[R_{gA}]$ are respectively: the expected annual precipitation, the expected annual actual evapotranspiration, the expected annual surface runoff and the expected annual groundwater runoff.

Due to the previous assumption, the terms relative both to the change in surface and in ground storage do not longer appear in equation (A.2).

Each of the water balance terms in (A.2) is a function of soil moisture, the characteristics of the stochastic precipitation input, the rate of potential evapotranspiration, the physical properties of the soil (e.g., porosity, intrinsic permeability, pore disconnectedness), and the properties of the vegetation (transpiration potential and fractional vegetation cover).

A.3 Model structure and assumptions

The physical system is represented as a soil moisture dynamic process where the precipitation input is the stochastic variable, while the output is a set of average annual value of the other components of the water balance.

As far as precipitation is concerned, the arriving storms are assumed to be manifestation of a Poisson process of rectangular pulses (Fig. A.2). This simplification enables a satisfying representation of the precipitation process through few easily treatable parameters. As further assumptions, storm intensity and the storm duration are supposed to be independent and exponentially distributed, while storm depth is considered gamma distributed.

The soil is assumed to be homogeneous and characterized by a vegetative coverage operating in equilibrium with its environment in an unstressed state. Soil moisture dynamics are captured through a simplified version of the concentration-dependent diffusion equation (Phillip 1969), while the soil properties are based on Brooks and Corey model (Brooks and Corey 1966).

The solution of the water balance equation in the form (A.2) or in an equivalent formulation implies the knowledge of all the water fluxes. As mentioned before, water fluxes can be

expressed as analytical functions the soil moisture content, here defined as the relative soil saturation (equation (A.3)), and a small number of climate, vegetation and soil parameters.

$$s = \frac{\theta_t - \theta_r}{n_t - \theta_r} \tag{A.3}$$

where θ_t, θ_r and n_t are respectively: the total volumetric water content of the soil, the residual volumetric soil water content and the porosity.

Under the assumption of stationary system, the solution of equation (A.2) is obtained under the constraint of uniform soil moisture content in a semi-infinite soil column. To solve the balance equation, then, Eagleson (1978a-g) proposes to use a single value of the soil moisture concentration, s_0, that can be defined as a "temporal mean of the spatial average", Fig. A.3.

With soil and climate parameters fixed for the given control volume, the water balance equation is essentially a function of the following two unknowns: the average soil moisture concentration, s_0, and the vegetation fractional coverage, M. However, for each set of soil and climate parameters, there is more than a unique couple of (s_0, M) that satisfies (A.2) and closes the water balance. In this framework, we further assume that vegetation operates, on the long term, under conditions of minimum stress. This implies, as suggested by Eagleson (1978g), that equation (A.2) will be solved for the couple of soil moisture content and vegetation coverage under which soil moisture content is maximized.

A.4 Area of analysis and data

The U.S. was subdivided into a 5x5 km grid and mapped using the Albers 1966 projection, leading to a study grid of 630 rows and 994 columns. The water balance model requires the input of soil, precipitation, and vegetation properties for each cell of the study grid, as well as the climatic forcing estimates.

In particular, for spatial unit where the model needs to be applied, it is necessary to evaluate the following parameters and climatic inputs:

Soil hydraulic parameters:

- Total porosity, n
- Pore size distribution index, m
- Matric potential at effective saturation, $\Psi(1)$
- Saturated hydraulic conductivity, $K(1)$
- Pore disconnectedness index, c
- Diffusivity index, d
- Surface retention capacity, h_0

Climatic parameters (long term storm statistics):

- Mean storm duration, t_r
- Mean time between storms, t_b
- Mean interarrival time, t_a
- Mean storm intensity, m_i
- Mean storm depth, m_h
- Mean number of storms per year, m_n
- Mean annual precipitation, m_{Pa}
- Mean rainy season duration, τ
- Parameter of the gamma distribution of storm depth, k

Vegeration parameters:

- Plants transpiration efficiency, k_v

Climatic input (forcing of the water balance):

- Joint probability distribution function of annual precipitation and annual potential evapotranspiration.

Collection of available soil and climatic datasets was required and in many case adaptation of the available datasets to the selected grid of study was needed.

A.4.1 Soil Hydraulic Parameters

The VEMAP soil dataset (Kittel, Rosenbloom et al. 1995; Kittel, Rosenbloom et al. 1996) was used as the preferred source of parameters describing the hydraulic characteristics of the soil. It contains 18 parameters for the 0-50 cm and the 50-150 cm layers. The VEMAP coverage uses a 0.5x0.5 degree grid covering the conterminous U.S. and inlcudes bulk density and texture (i.e. percentages of sand, silt and clay).

Using the standard assumption that mineral density is 2.65 g/cm^3, bulk density ρ can be readily converted to total porosity as follows:

$$n = 1 - \frac{\rho}{2.65} \tag{A.4}$$

Pore disconnectedness index, *m*, residual water content, ξ, and matric potential at effective saturation, $\Psi(1)$, are evaluated using a multivariate linear regression relating them to the percentages of clay, silt and sand and based on the 11 USDA textural classes (Kochendorfer 2005), whose parameters are listed in Table 3.I.

Hydraulic conductivity was estimated following the equation derived by Brutsaert (Brutsaert 1967):

$$K(1) = a \cdot \left(\frac{n - \xi}{\psi(1)} \right)^2 \cdot \frac{m^2}{(m+1) \cdot (m+2)} \tag{A.5}$$

where the coefficient *a* was set equal to 35 cm^3/s (Kochendorfer 2005).

Pore disconnectedness index, *c*, and diffusivity index, *d*, are instead evaluated as follows:

$$c = \frac{(2 + 3m)}{m} \tag{A.6}$$

$$d = c - \frac{1}{m} - 1 \tag{A.7}$$

The VEMAP data are provided on a 0.5x0.5 degree grid covering the U.S. That grid contains 5520 cells, with 115 columns and 48 rows, whereas our study grid contains 994 columns and 630

rows. In order to extend the results obtained at the VEMAP database resolution to our spatial resolution, a simple inverse distance squared procedure was developed. The procedure consisted of the following steps:

- The values of latitude and longitude were determined for the center point of each of the 5520 cells.
- Each cell of the 630x994 grid was assigned the values of the corresponding cell from the 48x115 grid.
- Remaining blank cells at the 630x994 resolution were filled by weighting the 9 nearest cells of the VEMAP databases according to their inverse distance squared.

A.4.2 Long Term Storm Statistics

Long-term means of hourly storm statistics, characterizing the Poisson arrival precipitation model, were estimated for stations available in the National Climatic Data Center (NCDC) hourly dataset. Estimates for the stations were then extended to the U.S. at the 5x5-km spatial level of resolution.

A total of 5264 hourly data gages were available from the NCDC dataset. The NCDC gages, however, are spatially and temporally heterogeneous, being more numerous in the densely populated regions of the U.S. and scarce in desert and mountainous areas (Fig. A.5). Record length was also extremely variable, ranging from 1 to 53 years (Fig. A.4). Furthermore, many of the NCDC records were characterized by large amounts of missing or unreliable data (the latter being records that, according to the NCDC, did not pass an extreme value threshold test and were, therefore, flagged at the source). Because we evaluated the storm statistics by analyzing the sequences of storms as they occurred in time, a large amount of missing data could not be tolerated, for it alters dramatically the shape of such sequence, potentially leading to large errors.

Of the 5264 stations available, 2088 were eventually selected for further use. We included only those stations with at least 30 years of hourly precipitation records with no more than 25% of the data missing. No correction was performed on the extreme values, principally because of the lack of specific information on which to base a correction.

Using the complete record of each included station, the following long-term storm statistics were evaluated: mean storm duration, t_r, mean time between storms, t_b, mean interarrival time, t_a, mean storm intensity, m_i, mean storm depth, m_h, mean number of storms per year, m_n , mean annual precipitation, m_{Pa}, mean rainy season duration, τ, and parameter of the gamma distribution of storm depth, k.

In agreement with Eagleson's model (Eagleson, 1978a-b), storm sequences are treated as a series of rectangular pulses. Thus, the total precipitation of a given event was obtained by summing hourly precipitation amounts over the duration of the event. Rainfall intensity was computed as the total precipitation during an event divided by the event's duration. Characterizing storms as rectangular pulses allows precipitation to be represented, in statistical terms, by a few easily measurable parameters. Each single storm and inter-storm period may be completely described by the time of arrival of the storm, the storm intensity, the inter-storm time, and the storm duration. In this approach, storm intensity and storm duration are assumed to be independent and exponentially distributed (Eagleson, 1978b), whereas the sequence of storms is assumed to be Poisson distributed.

If storm events are extracted from the precipitation records under the assumption that a single hour with no precipitation is sufficient to separate events, the resulting sequence of events typically is not Poisson-distributed. The lack of a Poisson distribution indicates that some raw storms in the data are not really independent of each other, suggesting that the time period being used to separate discrete rain events is too short and some contiguous rain periods should be considered as part of the same rainy event rather than as independent events (implying that a short

rainless period between two rainfall events may represent a mere interruption in the continuity of the event).

To address this issue, the raw sequences of precipitation data were processed following the procedure outlined by Restrepo-Posada and Eagleson (1982). This procedure requires determining the minimum rainless time span between two rain events that needs to elapse for the events to be considered as separate storms. To this aim, each raw sequence of rain events was subjected to the condition of being first-order Poissonian, that is, to have the mean of inter-arrival time equal to its standard deviation. The procedure uses the ratio of the standard deviation to the mean (that is, the coefficient of variation) of storm event inter-arrival time as the criterion for determining when a first-order Poisson distribution is achieved. The coefficient of variation is computed first from the original precipitation sequence, where a single rainless hour is used to separate rainfall events. If the resulting coefficient of variation is greater than 1, the minimum rainless time span between storms is increased to two hours, such that events separated by only a rainless period of one hour are merged together. The process continues, increasing the time between storms by one hour each iteration, until the coefficient of variation of the resulting sequence is as close as possible to 1. The final sequence of storms is then used to evaluate all of the long-term storm statistics for the given station.

A.4.2.1 Spatial Extension of Storm Statistics

To extend the station storm statistics to the full 5x5 km grid of the U.S., we used on a regionalization procedure that relies on regressing storm statistics on total precipitation; ordinary kriging and simple inverse distance methods were not considered because they do not take into account factors that may deeply influence storm statistics, such as elevation or total precipitation. Grid cells containing a station were assigned the storm statistics (storm depth, duration, inter-arrival time, time between storms, etc.) of that station. For each cell without a station, separately for each storm statistic, the statistic was regressed on average precipitation, with the cases for the

regression being the stations falling within a circular region of 100-km radius centered at the cell. The regressions were then used to estimate the values of the storm statistics of the cells (Fig. A.6). Weighted linear regression was used, with the weights being proportional to the square of the inverse of the distance between the station and the cell of interest. The values of total precipitation at the station points were taken from the PRISM dataset (Daly, Neilson et al. 1994).

A.4.3 Vegetation parameters

Application of Eagleson's water balance model requires specification of only one vegetation parameter, plant transpiration efficiency, k_v. k_v is defined as the ratio between potential evaporation from bare soil and potential transpiration from vegetated soil, under a condition of unlimited water supply. As described below, we use this parameter as a primary knob for model calibration. Therefore, for each cell of the study grid, the value of k_v was estimated in a way that allows modeled fluxes to match observed ones, rather than being calculated *a priori*. Details about the estimation of k_v are provided in the model calibration section below.

A.4.4 Climatic Variables (forcing of the water balance)

For the annual implementation of Eagleson's model, annual joint probability distribution functions of precipitation and potential evapotranspiration were needed; it is from these that it is possible to determine the probability distribution functions of the water fluxes. Annual historical values for precipitation and for minimum, maximum, and dew point temperature were taken from the PRISM database (Daly, Neilson et al. 1994) at the 5x5 km resolution. Those datasets were mapped using Albers 1966 projection in order to match our existing datasets format.

Historical values for potential evapotranspiration were derived according to a modification of Penman's equation by Linacre (1977):

$$ETp = \left[500\left(T' + 0.0006\,h\right)/\left(100 - A\right) + 15\left(T' - Td\right)\right]/\left(80 - T'\right) \qquad \text{(A.8)}$$

where *T'*, *h*, *Td* and *A* represent respectively: the monthly mean temperature in degrees Celsius, the elevation in meters, the mean monthly dew point temperature in degrees Celsius and the latitude in degrees.

A.5 Model Calibration

The goal of water balance model calibration was to have the model reproduce observed annual natural water yield as closely as possible, not only in terms of the long-term mean annual yield but also in terms of replicating annual observed traces of streamflow. The model was calibrated by minimizing the mean squared error between modeled water yield and estimates of historical natural streamflow. Three different historical streamflow datasets were used:

- A 42-year series of annual streamflow records for 655 relatively unmodified test basins across the U.S. (Hobbins et al. 2001; Slack and Landwehr 1992).
- Reconstructed natural streamflow estimates for years 1906-2006 for a set of watersheds in the Colorado River Basin, provided by the U.S. Bureau of Reclamation (USBR).
- 30-year average reconstructed natural streamflow for the 8-digit basins of the U.S. estimated by the U.S. Geological Survey (USGS) (Krug et al. 1989).

Calibration was performed at the basin (either a test basin or 8-digit basin) level, which required running the water yield model at the basin level. The estimation of the parameters needed to run the model at the basin level was performed by averaging parameter values across all 5x5-km cells within a basin.

Recall that the water yield model provides estimates of total natural water yield, equal to the sum of surface and subsurface yield. The calibration process is subject to errors if measured or reconstructed natural flows used for calibration do not accurately capture the sum of surface and subsurface flow that would naturally leave the basin, be it a test basin or 8-digit basin. Such error can occur where some of the water yield leaves the basin beneath surface, so that it is not captured at the stream gauge measuring basin outflow. It may also happen if withdrawals within

the basin are not accounted for. Further, at the annual time step additional error can be caused by annual fluctuation in the amount of water stored as groundwater.

A.5.1 Calibration of the model over the 655 Test Basins.

The 655 test basins were given first priority for model calibration; 8-digit basin data were used for calibration only outside of the boundaries of the test basins. The test basins were preferred because they are relatively unaffected by human intervention, such that streamflow is a fairly accurate estimate of natural water yield, thereby avoiding the need for natural flow reconstruction.

For each of the test basins a 42-year (1953-1994) sequence of annual streamflow data was used to calibrate the model, allowing us to compare predictions and observations on a year-by-year basis[2].

Given that the water yield model represents a first order expansion around an equilibrium solution (Eagleson, 1978f), calibrating the model on a mean annual basis (that is, comparing observed and predicted observed mean annual yield) theoretically would be sufficient to obtain an acceptable year-by-year fit. However, because of the limitations discussed above, especially the possibility of annual changes in stored water that go unnoticed, a year-by-year calibration may produce an improved fit. As seen below, where a mean annual calibration did not produce acceptable results, we tested a year-by-year calibration. Two approaches described in the next two sections, one based on the plant transpiration efficiency factor of the Eagleson model and the other based on a comparison of modeled and measured water yield, were developed to calibrate the model for areas within the 655 test basins. Each basin was individually calibrated using the procedure that produced the better result.

---

[2] Restricting the test basin calibration data to years before 1995 was a matter of convenience, as it allowed us to take advantage of our previous research (e.g., Hobbins et al. 2001)

A.5.1.1 Calibration through the plant transpiration efficiency

Adjustment of plant transpiration efficiency (k_v in the water balance model) was selected a priori as the principal calibration approach. Although in principle the model could have been calibrated by adjusting any other model parameter or set of parameters, plant transpiration efficiency was the only parameter for which we had neither a direct measurement nor any reliable approximation available at the large scale. Further, we sought to avoid use of sophisticated multivariate methods of calibration because of the complexities involved with using such methods over large spatial scales with many calibration sites, and because we hoped to keep the approach as easily tractable as possible. A simple bisection method - which is a mathematical solution-finding method that repeatedly bisects an interval and then selects a subinterval in which the solution must lie for further processing - was implemented in order to calculate, at the basin level, the single value of k_v that allowed a perfect match between average modeled water yield and observed average streamflow.

For a large majority of the test basins (see Fig. A.7) a value for k_v within the range 0.4 – 1.5 was found that allowed mean predicted annual yield to equal the 42-year average measured streamflow. In these cases, annual water yields were calculated by running the model with k_v set at the determined level and using annual values of precipitation and evapotranspiration as model inputs. Results were then compared with the observed traces of streamflow, as shown in Fig. A.8. The year-by-year calibration option was examined when an acceptable mean annual value of k_v was not obtained. This option allows k_v to vary year-by-year in relation to precipitation fluctuations. Allowing k_v to vary annually with precipitation (essentially, allowing plant transpiration to vary) reflects short-term plant adaptability to climatic conditions. The procedure consisted of the following steps:

- find a set of 42 annual values of k_v that allows the modeled water yield to match the observed annual streamflow at each year;
- linearly regress the annual values of k_v on the annual precipitation;
- calculate water fluxes at a given year i, using the value of $(k_v)_i = a + b \cdot (P_A)_i$, where $(P_A)_i$ is the precipitation at year i and a and b are the coefficients of the linear regression between P_A and k_v illustrated in Fig. A.10.

We adopted the year-by-year k_v calibration procedure if it improved the fit of the model. For an example of the improvement achieved using an annually-varying k_v for calibration, compare Fig. A.9 with Fig. A.8. Using the year-by-year calibration, a better match between observed streamflow and modeled yield was found for a considerable number of basins (Fig. A.11).

A.5.1.2 Calibration through the effective water yield

For those basins where use of the k_v calibration procedures described above did not allow modeled fluxes to match historical observations, an alternative approach was used that relied on computation of the ratio of modeled to observed streamflow. For this approach, the transpiration efficiency factor (k_v) was set to 1 (such that plant potential transpiration equals potential evaporation). With k_v set to 1, all fluxes were computed and the ratio between modeled mean annual yield and mean measured streamflow, *YS*, was calculated, as follows:

$$YS = \frac{\left(\overline{Y_A}\right)_{MODELED}}{\left(\overline{Stramflow}\right)_{OBSERVED}} \tag{A.9}$$

YS then was used to scale each year's modeled total yield. The effective water yield for year i, $(Y_A)_{iEFF}$, was calculated as:

$$\left(Y_A\right)_{i\,EFF} = \left(Y_A\right)_{i\,MODELED} \cdot \frac{1}{YS} \tag{A.10}$$

See Fig. A.12 for an example of the application of this procedure.

As was done with the k_v approach to calibration, to improve the year-by-year fit of the yield traces and better capture the annual variability of the observed streamflow, we investigated the possibility of allowing *YS* to vary annually in relation to precipitation. An optimal scaling value was determined by making sure that both the long-term mean yield (that is, the yield predicted by the model if precipitation and potential evapotranspiration are set at their mean annual values) and the 42-year average yield (that is, the average of the 42 annual estimates of yield, each obtained with the annual values of precipitation and potential evapotranspiration) converged to the mean observed streamflow. In summary, this procedure consisted of finding the value of η in equation (A.11) that leads to the sequence of scaled values of *YS* at year *i*, YS^*_i, that satisfies the aforementioned condition.

$$\left(YS^*\right)_i = \frac{\left(\left(P_A\right)_i - m_{PA}\right)\cdot\eta + m_{PA}}{m_{PA}}\cdot YS \tag{A.11}$$

When this procedure was successful, that is, when it improved the fit between observations and simulations, effective water yield in the given basin in year *i* was calculated as:

$$\left(Y_A\right)_{i\,EFF} = \left(Y_A\right)_{i\,MODELED}\cdot\frac{1}{\left(YS^*\right)_i} \tag{A.12}$$

For an example of the improvement achieved by allowing YS to vary with precipitation, that is, using in each year *i* the value of that is calculated based on actual precipitation, compare Fig. A.13 with Fig. A.12. This annualized procedure improved the fit for many test basins (Fig. A.14).

Results for spatial error distribution are shown in Fig. A.15 and Fig. A.16. As apparent from these maps, the largest absolute errors between estimated water yield and observed streamflow occur in the eastern and western U.S., with smaller errors in the central U.S. However, examining relative as opposed to absolute errors reveals a different pattern, with many areas of the central and southern U.S. having errors comparable in magnitude to the average yield.

A.5.2 Calibration over the 8-digit Basins and Colorado River Basins Watersheds

In areas where no test basins were present, the model was calibrated using 30-year (1951-1980) average reconstructed natural flows estimated by the USGS for the 8-digit basins of the U.S. (Krug et al. 1989). In addition, for the Colorado River Basin the 8-digit basin flows were used in combination with reconstructed natural flows estimated by the USBR. Of the USBR flow estimates, we used data for years 1953-2004.

The natural flow estimates from the USBR were considered more reliable than the 8-digit basin averages from the USGS for the same area. Some of the flows from the USBR refer to stations with drainage areas considerably larger than the typical 8-digit basin (Fig. A.17). Where USBR flow estimates were available for a catchment that included more than one 8-digit basin, the interior 8-digit basin averages were scaled in such a way that the aggregate streamflow across the interior 8-digit basins matched the USBR average for the catchment. The scale factor, ξ, was obtained as follows:

$$\xi = \frac{\overline{Streamflow}_{USBR} \cdot A_{USBR}}{\sum \left(\overline{Streamflow}_{8DB} \cdot A_{8DB} \right)_i} \tag{A.13}$$

where $\overline{Streamflow}_{USBR}$, A_{USBR}, $\overline{Streamflow}_{8DB}$ and A_{8DB} are respectively the average observed streamflow for the USBR catchment, the area of the USBR catchment, the average streamflow, and the area of each of the 8-digit basins contained within the USBR catchment. Calibration was then performed at the 8-digit basin scale by matching the modeled yield with a corrected 8-digit basin streamflow, $\overline{Streamflow}^{*}_{8DB}$, obtained as follows:

$$\overline{Streamflow}^{*}_{8DB} = \xi \cdot \overline{Streamflow}_{8DB} \tag{A.14}$$

This procedure guarantees that global simulated yields match the observed streamflow at the scale of the USBR catchments and, simultaneously, that the lower scale variability at the 8-digit basin is preserved. Calibration over the 8-digit basins, or over the 8-digit basins in combination

with the USBR streamflow records, was then performed following the approach described above for the test basins, except for the year-by-year regressions, which were not used with the 8-digit basins.

This procedure guarantees that global simulated yields match the observed streamflow at the scale of the USBR catchments and, simultaneously, that the lower scale variability at the 8-digit basin is preserved. Calibration over the 8-digit basins, or over the 8-digit basins in combination with the USBR streamflow records, was then performed following the approach described above for the test basins, except for the year-by-year regressions, which were not used with the 8-digit basins.

A.6 Extension of Calibrated Parameters to the Conterminous U.S.

Once calibration was performed for the entire study area, the model could be applied at the 5x5 km or any larger spatial scale. To apply the model at the 5x5 km scale, each cell was assigned the values of k_v or *YS* that allowed convergence between average observed streamflow and mean modeled water yield for the basin (either test basin or 8-digit basin) to which the cell belonged. Ultimately, we found that applying the model at the 5x5 km resolution for all of the different scenarios-GCM combinations was computationally too time-consuming to be practicable. Therefore, we used the model at the grid cell level only to estimate the average mean annual yield (Fig. A.19), for comparison with the USGS mean annual runoff estimates (Fig. A.18). Estimation of annual historical water yield and future water yield based on climatic and socio-economic projections was instead performed at the 8-digit basin scale.

A.7 Model Input Parameters for Future Climatic Scenarios

Application of the model to predict future water yield required estimates of model input parameters for future years. As explained above, input parameters of the model include the soil hydraulic properties, vegetation properties (essentially vegetation transpiration efficiency, k_v), and

storm statistics. With varying levels of confidence, we assumed that future levels of these parameters would equal past levels. Regarding the soil hydraulic properties, this assumption is easily accepted, as the parameters are representative of soil texture and composition, which are unlikely to change over this century. Somewhat less easily accepted, but still in our judgment reasonable, is the assumption that the plant coefficient k_v will remain constant over the time horizon, as it ideally represents the result of the plant evolutionary adaptation to the environment. On the other hand, storm characteristics may be expected to change as the climate changes. Our estimation of storm statistics based on past weather data relied on the assumption of a stationary climate, which is an assumption that will be increasingly untenable if the climate changes as indicated by current global climate models. Estimation of storm statistics for the future, however, is constrained by the fact that their estimation relies on hourly precipitation data. As described above, hourly data were available for past years, but estimates of future precipitation were available only at the monthly time step. Therefore, to reflect future climatic conditions, changes in the values of storm statistics would need to be inferred from changes in monthly climatic data. Because of the magnitude of uncertainty that would be introduced by such inference, it was decided to apply the current storm statistics to future water yield estimation. In summary, projection of water yield for future climatic and socio-economic scenarios was based on the water yield model as forced by the future predicted PDFs for precipitation and potential evapotranspiration, but employing historical soil, vegetation, and storm parameters.

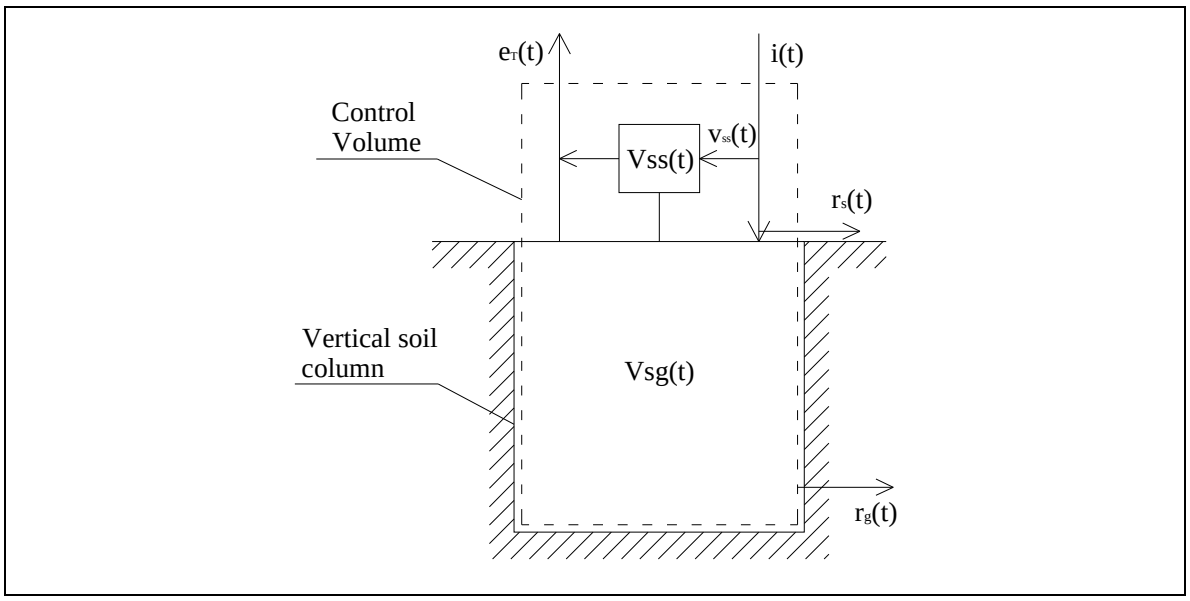


Fig. A.1 Control volume, input and output fluxes relative to Eagleson annual water balance model.

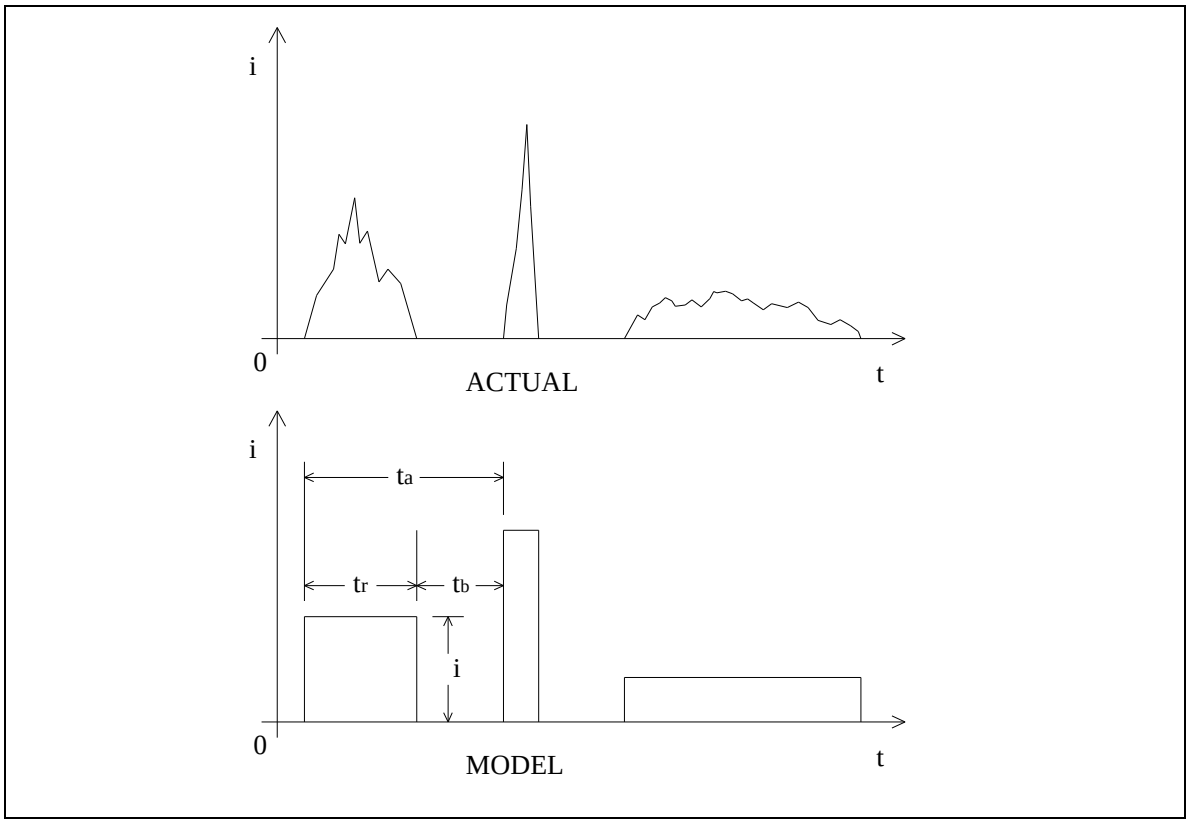


Fig. A.2 Comparison between the actual behaviour of rainfall and the one modelled for this analysis. In the horizontal axis is reported the time and in the vertical one the rain intensity. The parameters t_a, t_r, t_b represent respectively the inter-arrival time, the storm duration and the time between two storms.

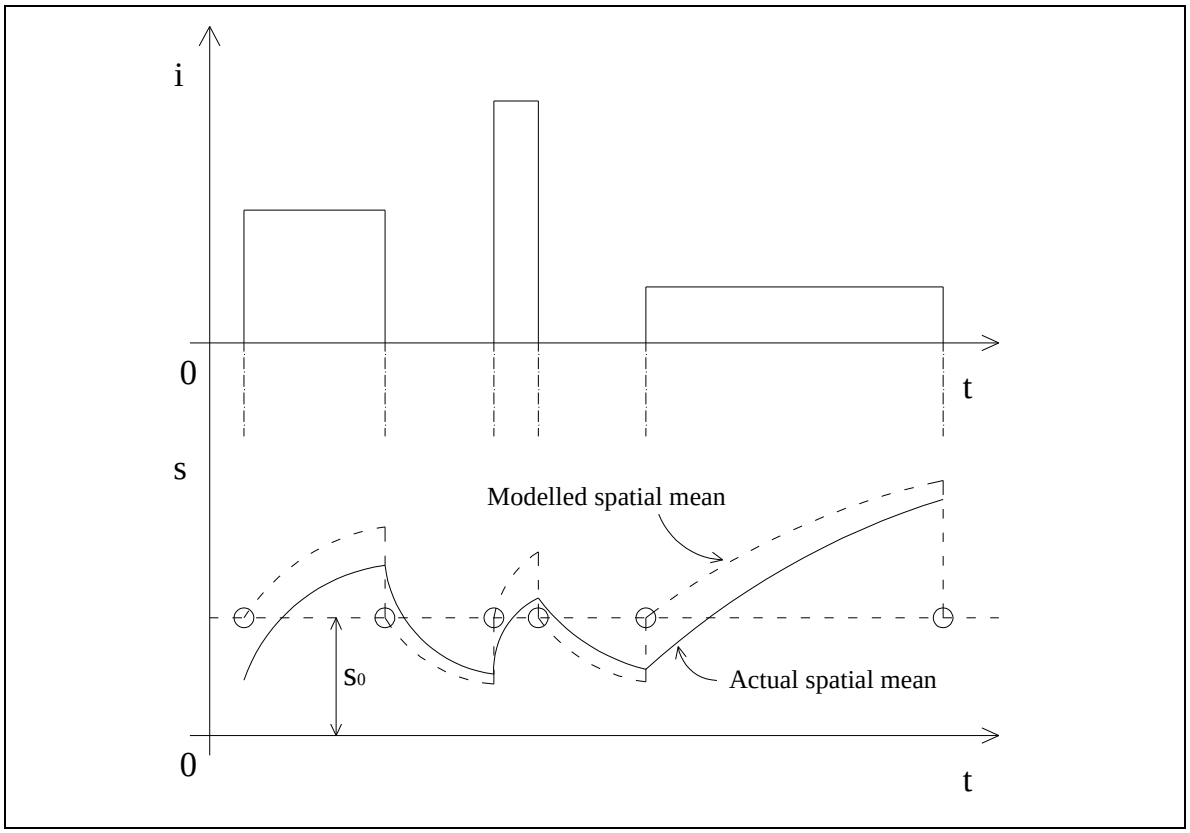


Fig. A.3 Spatial mean of soil moisture content. Comparison between the actual behaviour of the soil moisture concentration and the simplified model adopted in the analysis.

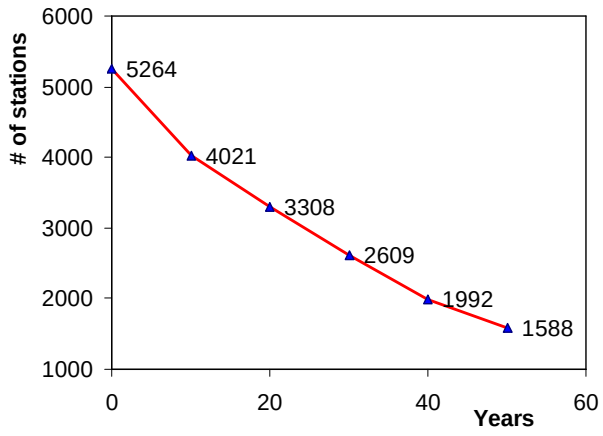


Fig. A.4 Distribution of the number of NCDC stations providing hourly data of precipitation as a function of record length.

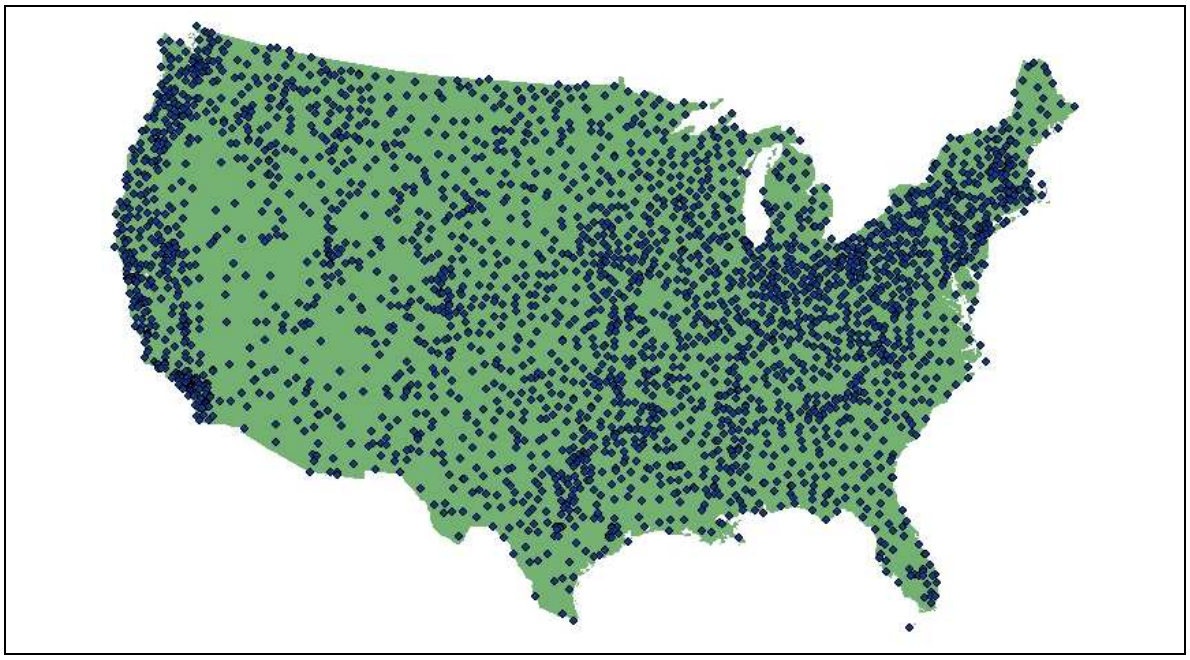


Fig. A.5 Spatial distribution of NCDC stations providing hourly datasets of precipitation.

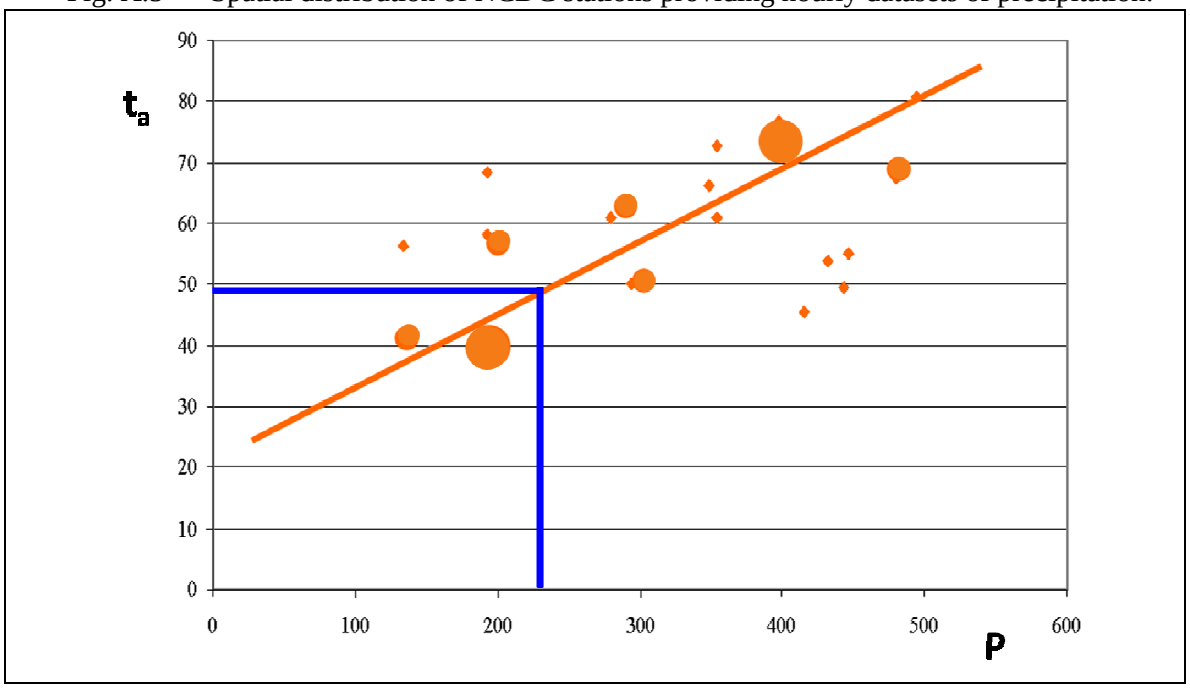


Fig. A.6 Sketch of weighted linear regression used to evaluate storm statistics (in this case the interarrival time) at 5x5 km resolution. Orange dots represent stations located near the cell under analysis. Dot size represents the weight given to the station in the regression, the weight being inversely proportional to the station's distance from the cell of interest. The orange line is the regression line as estimated by the weighted linear regression. Blue lines represent the procedure to estimate the value of the storm statistic of interest (in this case interarrival time) for the cell considered from its value of precipitation as extracted from PRIMS dataset at the 5x5 km spatial resolution.

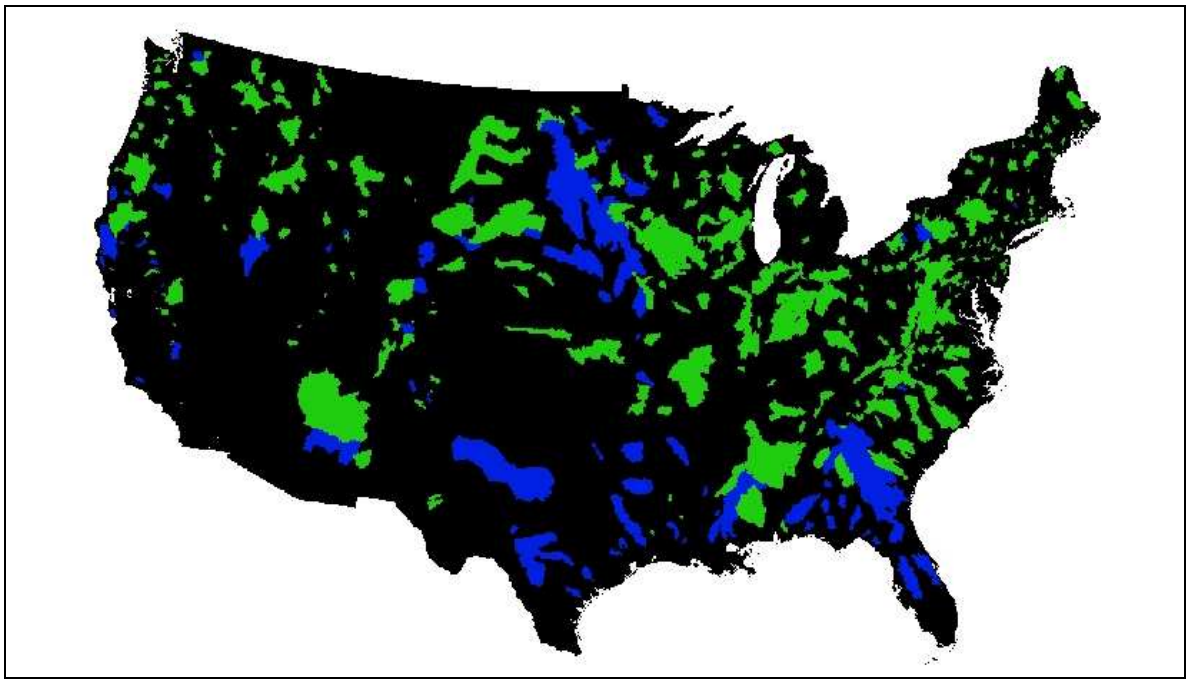


Fig. A.7 Test basins shown in green are those where a perfect match was possible by changing the transpiration efficiency only. Test basins in blue are those where such match was not possible only by changing k_y.

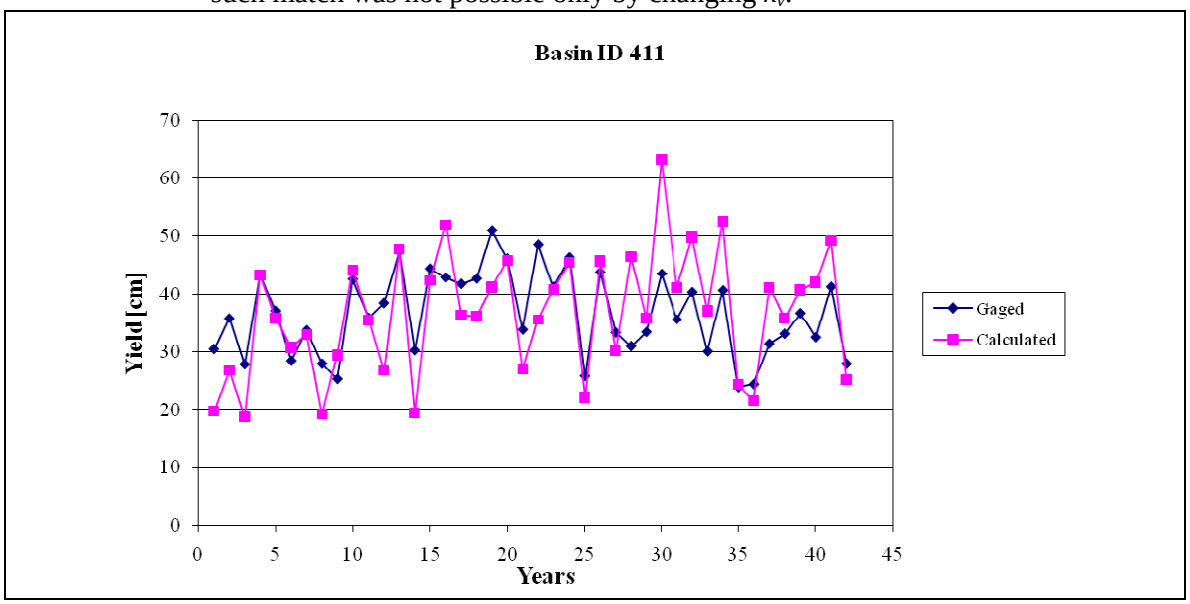


Fig. A.8 42-year sequence of calculated yield and measured stremflow for test basin 411 (northern Wyoming). The estimated sequence is obtained by using the plant transpiration efficiency which led to convergence between mean estimated yield and average measured streamflow.

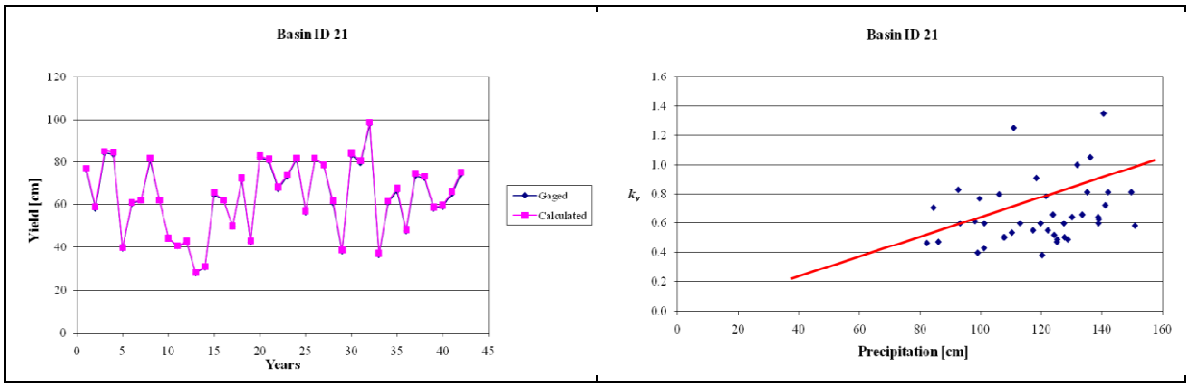


Fig. A.9 Calibration of test basin 21 (coastal Massachussets). Left: 42-year sequence of calculated yield and measured streamflow for test basin 21, with the calibration based on 42 annual estimates of k_v. Right: annual values of k_v that guarantee a perfect match between measured and calculated yield versus annual measured precipitation, also showing the linear regression line resulting from regressing k_v on precipitation.

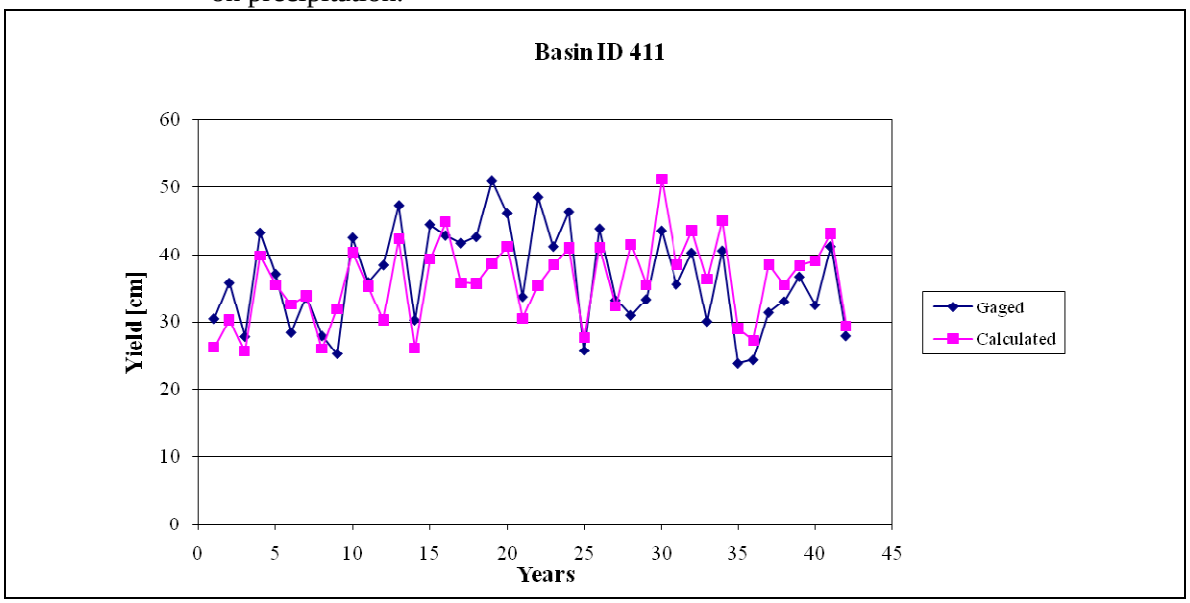


Fig. A.10 42-year sequence of calculated yield and measured streamflow for test basin 411 (northern Wyoming). The estimated sequence is obtained by using a variable plant transpiration efficiency factor as obtained from the regressive procedure.

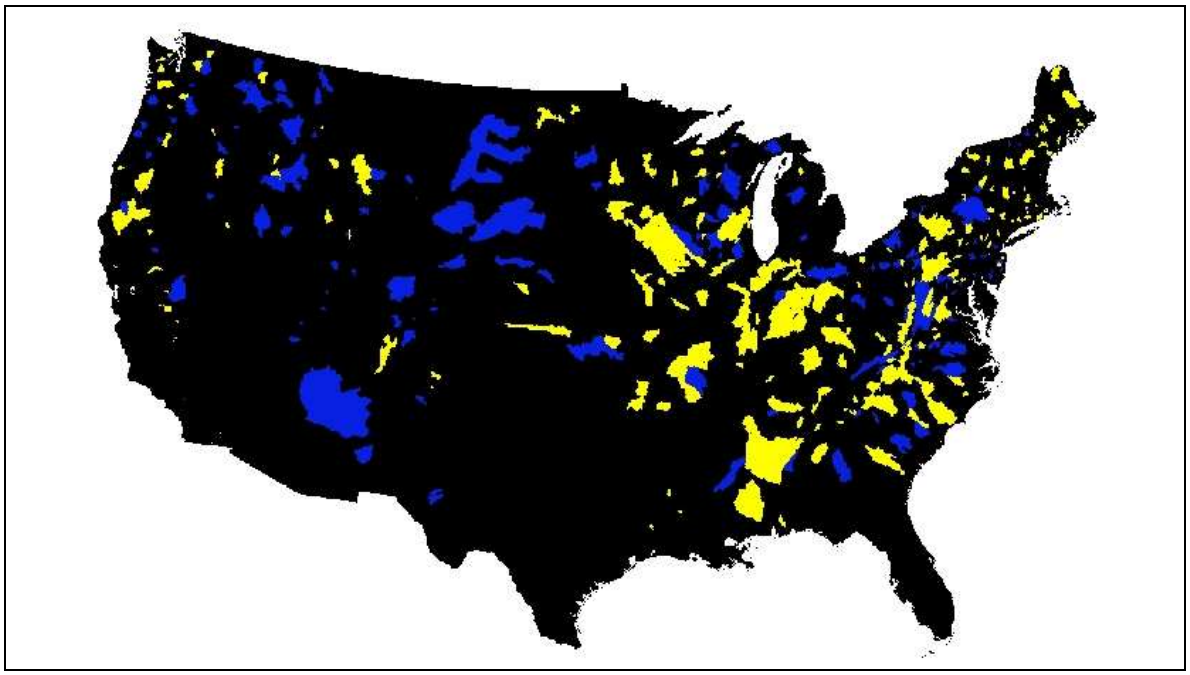


Fig. A.11 In blue are indicated those basins where a yearly variable plant transpiration efficiency leads to a better match between model output and observed streamflow (for basins where convergence was already achieved only by plant transpiration efficiency calibration, green in Fig. A.7). Yellow indicated basins where keeping the plant transpiration efficiency constant (and equal to the value obtained in step 1 of calibration process) lead to a better fit with the observed streamflow sequence (in terms of 42-year MSE).

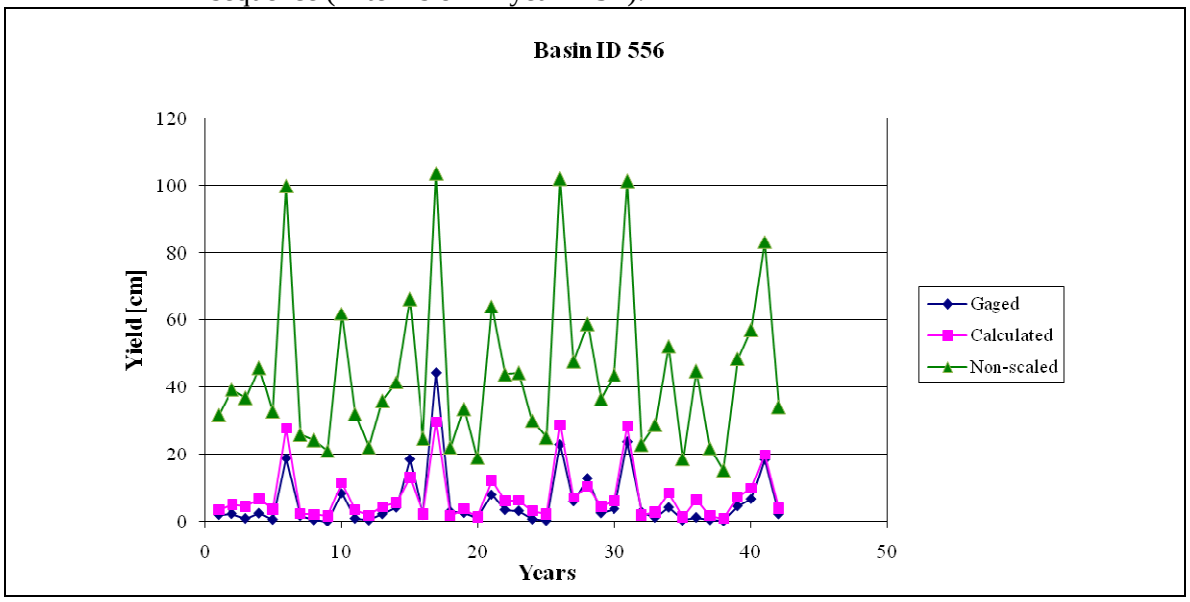


Fig. A.12 42-year sequence of calculated yield and measured streamflow for test basin 566 (coastal central California), showing measured streamflow (blue), calculated un-scaled yield (green), and effective (calculated and scaled) yield per equation 2.10 (pink).

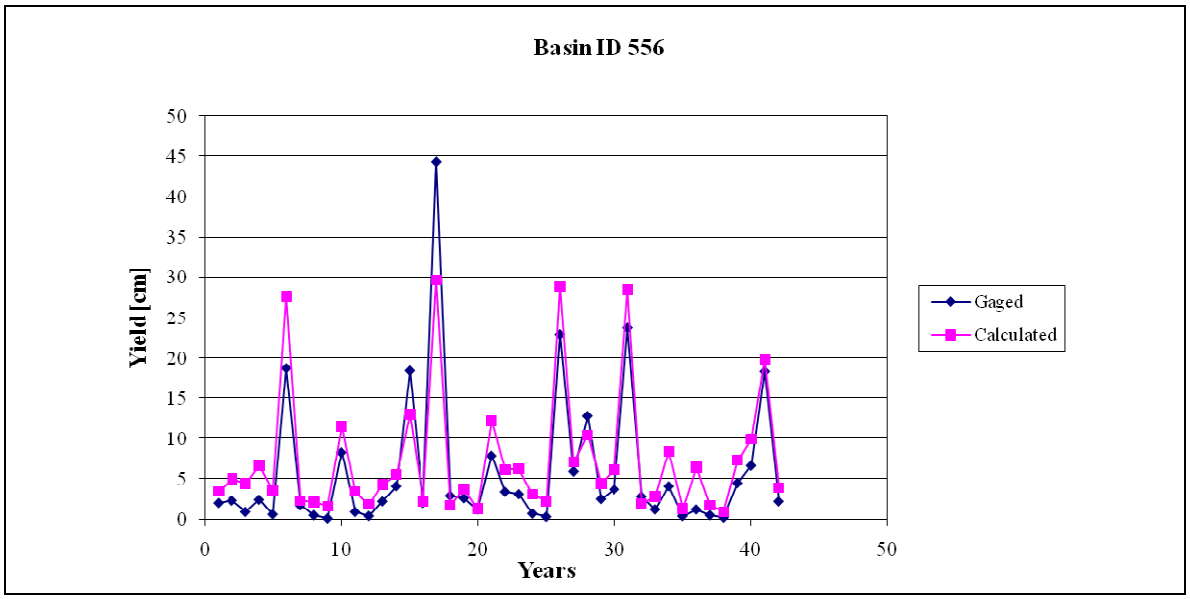


Fig. A.13 42-year sequence of calculated yield and measured streamflow for test basin 566 (coastal central California), showing measured streamflow (blue) and calculated yield (pink). Yield is calculated by scaling each year's prediction of annual yield by a factor equal to the ratio of the ratio mean estimated yield to average measured streamflow (YS) times the actual value of precipitation.

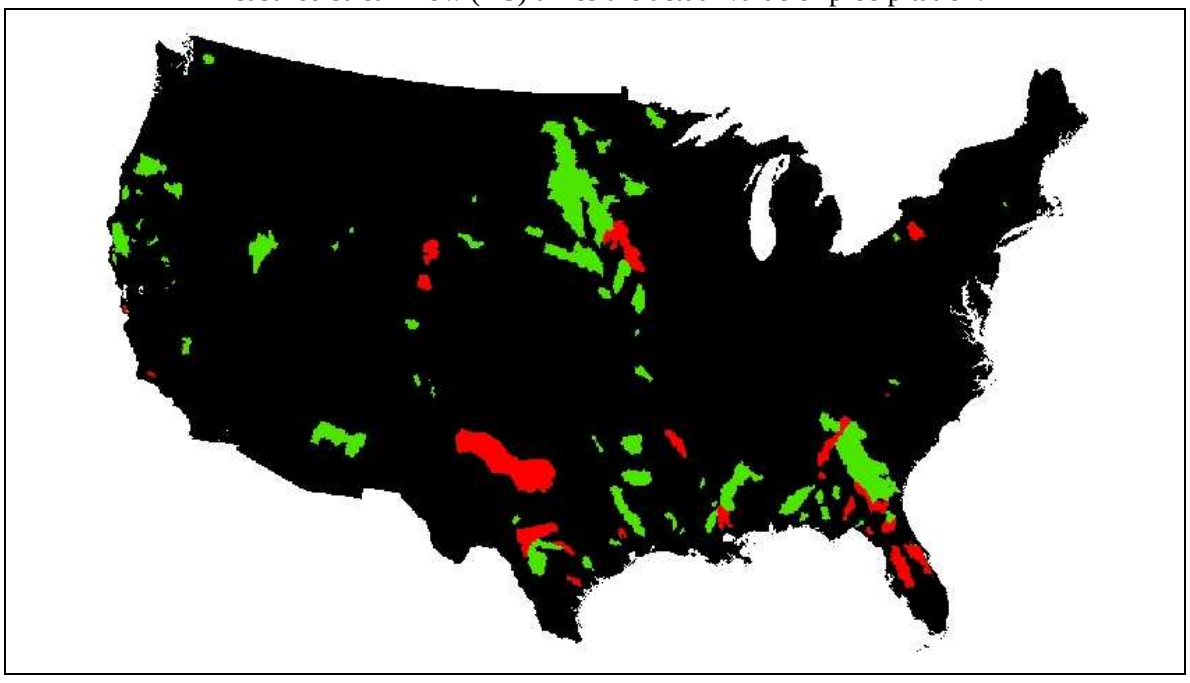


Fig. A.14 Basins in red are those where a yearly variable scaling factor leads to a better match between model output and observation (for basins where convergence was not achieved only by plants' transpiration efficiency calibration and only a fraction of yield is used to match the average streamflow). Green indicates basins where a better match between model and observations was found by keeping the scaling factor constant from year to year.

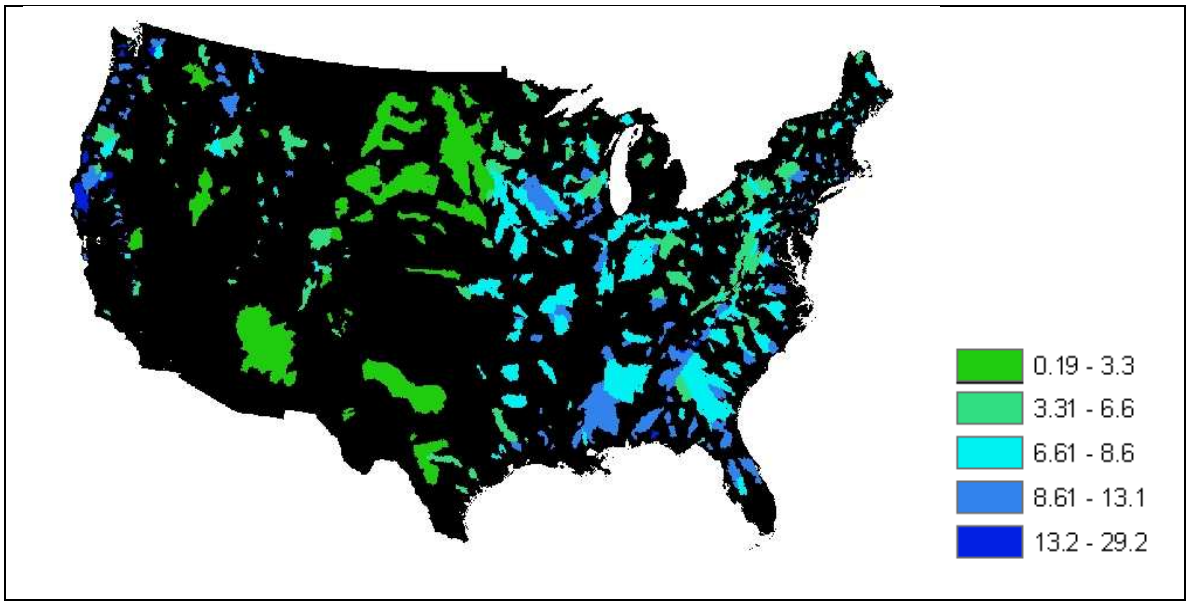


Fig. A.15 Square root of the mean squared error between modeled water yield and measured streamflow (MSE) for the test basins (cm).

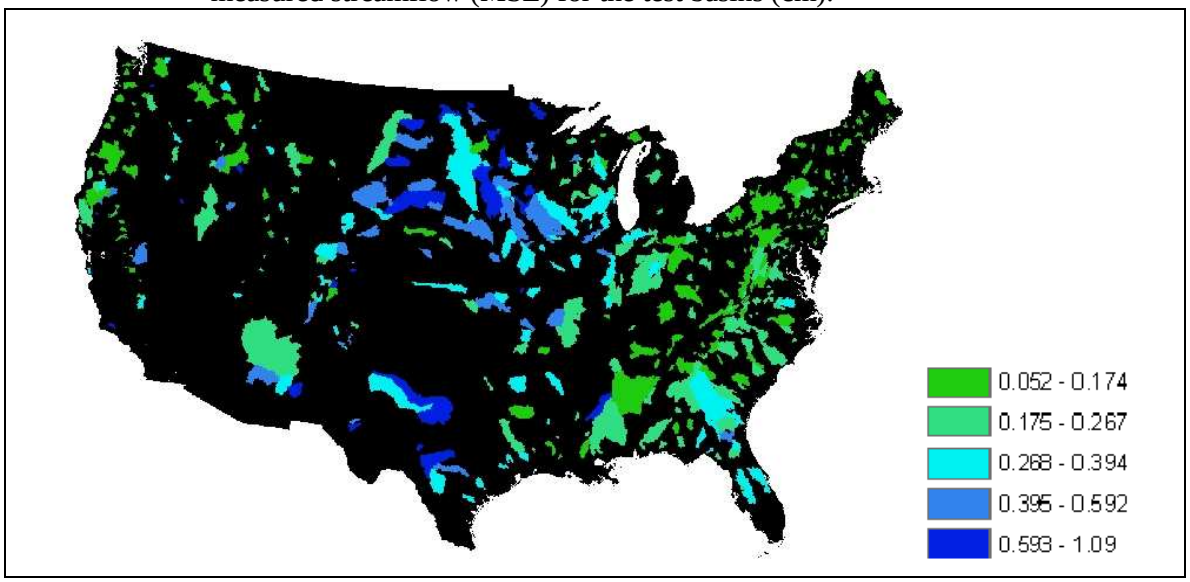


Fig. A.16 Ratio of the square root of MSE to average streamflow for the tests basins.

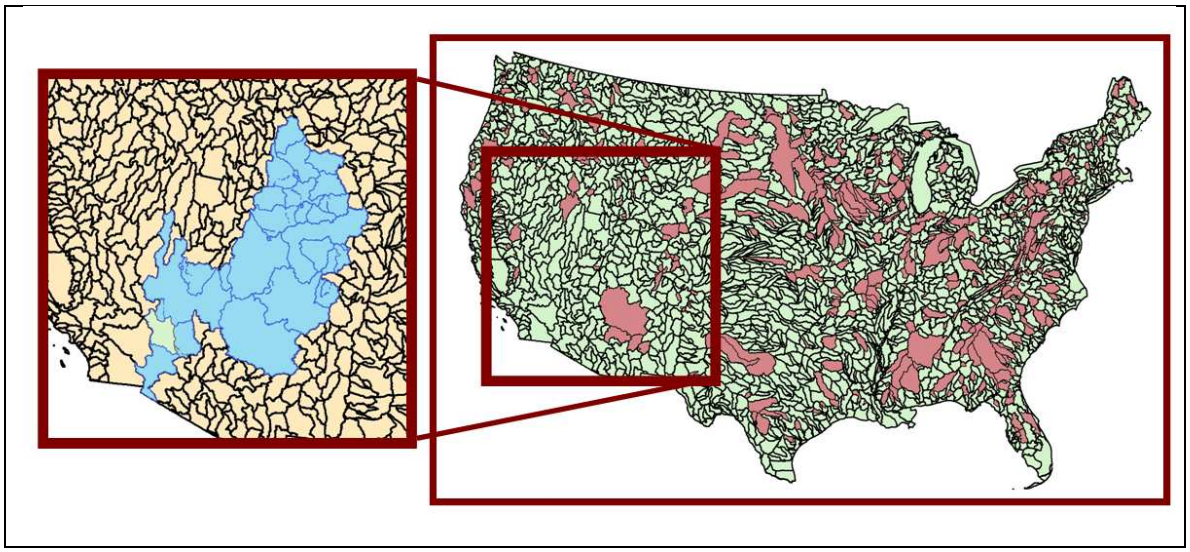


Fig. A.17 Map of the 655 test basins (red), 8-digit basins (green) with zoom of the and Colorado River Basin with the watersheds available in USBR records (light blue).

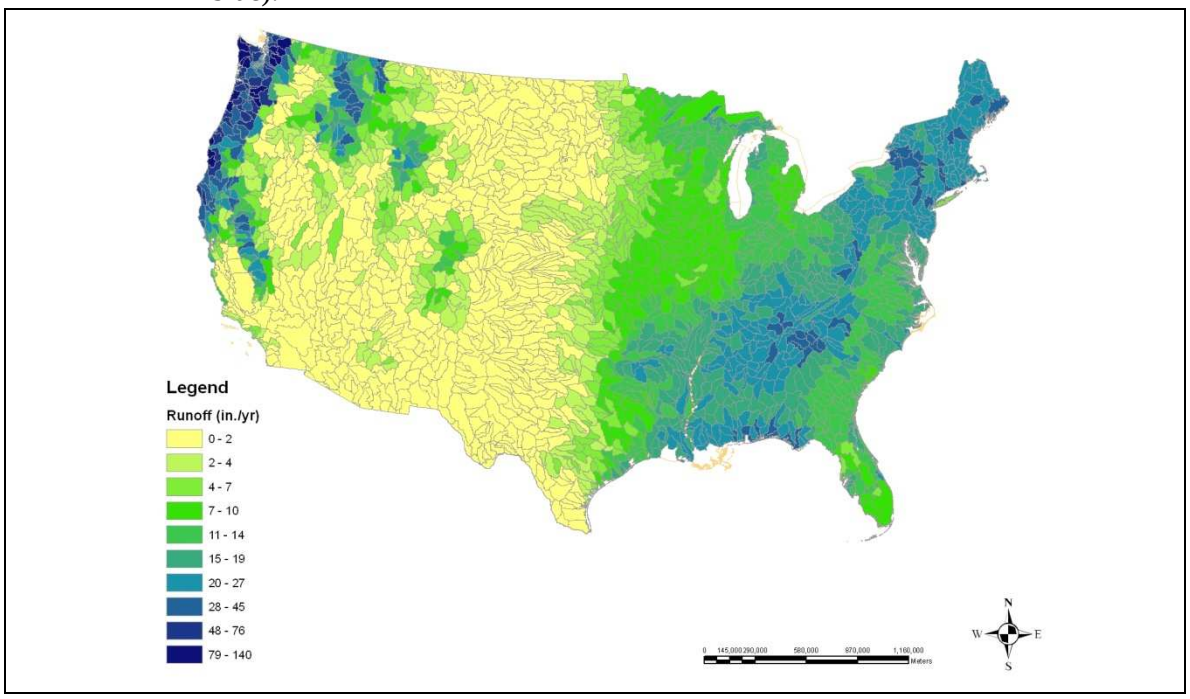


Fig. A.18 USGS average runoff at the 8-digit basins level.

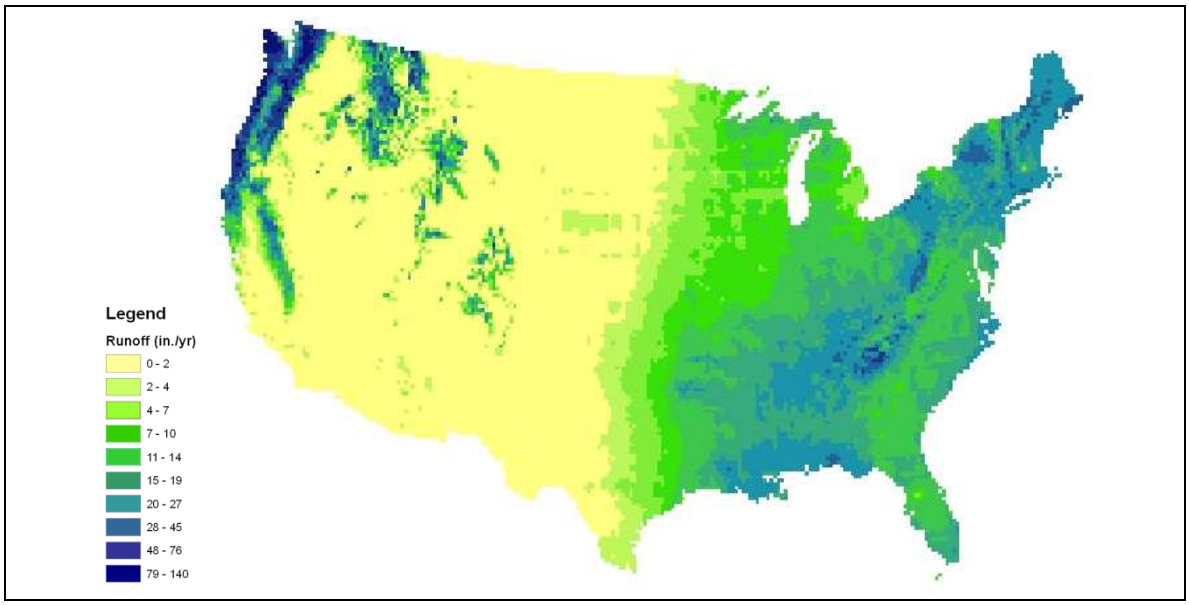


Fig. A.19 Modeled average runoff for the study grid.

Tab. A.I Results of multivariate linear regression of Brooks and Corey parameters (Abu Rizaiza 1991; Kochendorfer 2005).

Parameter	Intercept	Coefficients		R2
		%Sand	%Clay	
ξ	-0.0295	0.00076	0.00201	0.831
Ψ_1	14.6	-0.09340	0.45400	0.893
m	0.202	0.00329	-0.00318	0.868

Appendix B The US water supply network

B.1 Overview

Estimating the annual water yield as simply the difference between precipitation and evapotranspiration at the study grid level, or even at a basin level, does not provide, by itself, a valid estimation of the water effectively available in a certain point in space. Water, in fact, can be either diverted from a place where it is in excess to another where it is scarce, or stored in loco for future needs. Ultimately, if not needed, it simply flows from upstream to downstream, where it is made available for local demands or is lost to oceans. This implies that a thorough characterization of water supply for a large area such as the one we are dealing with in this study cannot omit from an accurate understanding of the complex water network existing within the area itself.

In this framework, we will subdivide the contiguous U.S. into 98 Assessment Subregions (ASR) and describe the complex system of water networks that takes place among them. ASRs were originally delineated by the Water Resources Council for its Second National Water Assessment (1978) to further subdivide Water Resource Regions (WRR). They are tracked with a four-digit code and their boundaries do not span the 18 boundaries or divide the standard four-digit basins. Therefore, ASRs can be aggregated to obtain estimates for the WRRs and are generally large enough to support the use of county-level data while still allowing analysis of some large regional differences within WRRs.

In the following sections we will describe the structure of the water networks existing within the conterminous U.S. at the ASR level (section B.2), the location and size of storages (section

B.3), the types of water demands considered (section B.4) and the set of rules adopted for the network simulation (section B.6).

B.2 Network Structure

Networks were characterized as systems of nodes connected by links. Each ASR, as well as each demand, storage capacity, and network sink (representing outflow to oceans), is represented by a node. Natural and artificial water routes connecting ASRs are represented by directional links. Links are also used to connect ASR nodes to storage and demand nodes (see Fig. B.1).

ASRs were considered linked, that is, part of the same network, when it was possible to individuate a water route between them. A water route is here defined as a sequence of water links, either natural (due to natural upstream to downstream flow) or artificial (via water diversions).

ASRs were considered linked, that is, part of the same network, if they were connected by either a natural flow path (upstream to downstream river flow) or an artificial diversion (via a canal, tunnel, or other constructed conveyance). An artificial connection between ASRs is called a trans-ASR diversion. Of the 98 ASRs, 83 are part of multi-ASR networks and the remaining 15 are unconnected. The unconnected ASRs drain to the sea or to Canada or are closed basins. Three multi-ASR networks were delineated, one with 69 ASRs that includes most of the central and western U.S., one with 10 ASRs in the Northeast, and one with 4 ASRs in the Southeast other (Fig. B.2).

B.3 Storages

To estimate reservoir storage we began with the National Inventory of Dams (NID) maintained by the U.S. Army Corps of Engineers, downloaded in June 2009.

The database was filtered in order to retain only reservoirs with a normal surface area of at least 5 km^2, resulting in 1509 dams and associated reservoirs. This initial list has then been

trimmed to retain only one record per reservoir, resulting in 1243 reservoirs. We further removed reservoirs that were for flood control only, were mining tailings ponds, or stored power plant cooling water. In addition, we independently checked reservoirs that had an unusual normal surface area to storage volume ratio (a common case was a storage volume of 0), and revised those values accordingly, resulting in the removal of several other reservoirs for which the research showed that the normal surface area was below 5 km^2. Finally, we added one reservoir that was not on the list. The final list includes 1196 reservoirs (of this list, we revised the surface area of 13 reservoirs and the storage volume of 44 reservoirs).

Storages were then aggregated at the ASR level, meaning that the storage volume of each given ASR is the sum all of those storages (among the final 1196 reservoirs list) that fall within the ASR borders.

Normal storage capacities were then aggregated at the ASR level, such that the storage volume of each ASR is the sum all of the storage volumes (among the final 1196 reservoirs list) that fall within the ASR borders. Normal storage volumes of the ASRs vary from 0 (ASR 1602) to over 40 million acre-feet (ASR 1005). Thirteen ASRs have at least 10 million acre-feet of storage (Tab. B.I).

The amount of stored water lost to evaporation in a given year was estimated by computing a surface area corresponding to the known storage volume and then multiplying that area by an evaporation rate. Because we aggregated reservoirs within an ASR, and therefore lost details specific to each reservoir, global relationships were needed.

The basic approach used to determine ASR reservoir surface area was to develop regional area-to-volume relations for four large groups of ASRs. The groups were formed as groups of WRRs whose reservoirs exhibited relatively distinct area-to-volume relations. The groups ranged from those tending to have relatively shallow reservoirs to those tending to have relatively deep reservoirs (Fig. B.5). Area-to-volume relations were determined for each group by regressing normal surface area on normal storage volume across all reservoirs in the group. For all

regressions, the constant was fixed at zero. All ASRs in a group were assigned the area-to-volume relation of its group. Individual evaporation rates for each ASR were set equal to the mean 1953-2005 potential evapotranspiration rate from a wet surface of all cells within the ASR.

A separate approach was used, however, for ASRs 1404 and 1503. For those two ASRs, Lakes Powell and Mead, respectively, make up nearly all of the storage capacity. Detailed surface area-to-storage relations were available for these reservoirs (Fig. B.4), as were average evaporation rates.

B.4 Water use classes

Three classes of water use were included in the network simulations: in-stream flow requirements, trans-ASR diversions, and consumptive uses. Each of these classes was individually examined for each ASR, as explained in the following sections.

B.4.1 In-stream Flow Requirements

In-stream flow requirements reflect the desire to leave some water in the stream for wildlife, fish, recreational activities, and aesthetic concerns. Ideally, the determination of a required minimum in-stream flow would be addressed locally, in order to consider properly the biological and environmental characteristics of each ecosystem. Careful consideration of local stream characteristics is unrealistic at the ASR scale, and beyond the scope of this study. In place of a more locally-specified minimum flow, we adopted the general guideline provided by Tennant (Tennant 1976), and specified the in-stream flow requirement as 10% of the average streamflow. Average streamflow was computed as the average total yield over the years 1953-1985. This constant amount was applied on an annual basis to both current and future conditions, without adjusting for shifts in the average water yield due to potential climate changes.

B.4.2 Trans-ASR Diversions

Trans-basin diversions, which move water across basin divides, are common throughout the U.S. and especially common in arid regions. Typically the amount of water diverted reflects long-standing legal agreements that specify the operating rules used to determine the diversion amount in a given year. Diversion amounts generally vary from year to year about a long-term average amount. Of course, the rules may also change over time. Because the operating rules differ from one diversion to the next and are not easily available, and because of the large number of such diversions, we adopted a simple procedure for including the trans-ASR diversions in the network simulations, setting each trans-ASR diversion equal to a constant amount computed from data on past diversions.

Most of the information regarding water diversions was taken from two publications of the USGS regarding respectively the western (Petsch 1985) and the eastern (Mooty and Jeffcoat 1986) U.S. These publications report on transfers between 4-digit basins for 1982 and any prior years back to 1973 if available. From these data we computed the average diversion across years 1980-1982. Information from these two sources was supplemented by more recent information for California (California Department of Water Resources 1998), Colorado (Colorado Water Conservation Board 1998, 2010; Litke and Appel 1989), the Lower Colorado River Basin (International Boundary and Water Commission 2004), and other locations from miscellaneous sources. Data for inter-basin transfers were aggregated at the ASR level in order to obtain an updated dataset of trans-ASR water diversions (Tab. B.IV). The trans-ASR diversion amount was held constant across all simulation years.

B.4.3 Consumptive Use

Historical records for water consumptive use are available at the ASR level for the years 1985-2005 from the USGS databases. These records were used to simulate historical and current conditions and as well as a useful source of information to predict future water demands.

Estimation of future demands relied mostly on the extrapolation of past trends in water use efficiency and projections of the main socio-economic and climatic drivers of water use listed in Tab. B.V. A considerable effort is also devoted to estimating the effects of climate change and future liquid fuel energy goals on future water use.

Trends in water use efficiency and drivers of water uses were analyzed for the six categories of water demands reported in Tab. B.VI.

In general consumptive uses are estimated by multiplying the estimated withdrawals for each category of water use by a consumptive use proportion and adjusted by adding a factor that takes into account potential future consumptive use attributable to climatic or other changes that are not reflected in data on past levels of water use. Further details are provided in Tom et al (Forthcoming).

B.5 Water use priorities

In the simulation of water allocation for the water networks, the three classes of water use were assigned distinct priorities. The priorities determine the order in which the classes are satisfied. In times of water abundance, the priorities have no practical impact, as all uses are met, but in times of water shortage the priorities determine which uses are met and which are not met or are only partially met. Note that the five categories of consumptive water use were treated as a block, and thus were assigned equal priority. Because of this simplification, we are unable to distinguish among the water use categories in time of shortage and cannot estimate how the shortage would affect each separate category. Although in reality the different water use categories may not suffer equally in times of shortage—the effect of shortage in a given ASR or sub-basin within an ASR would actually reflect the distribution of water rights or existing water allocation rules—fully accounting for local water allocation arrangements was beyond the scope of this national study. The three classes of water use were assigned priorities in the following order: in-stream flow requirements, trans-ASR diversions, and consumptive uses. Reservoir

storage was given the next lower priority. This order of priority guarantees a minimal amount of water for environment and ecosystem needs before any other needs are met, and satisfies major water diversion agreements before meeting local demands or storing water.

Note that water uses belonging to the same class were assigned the same level of priority irrespective of their position in the network. For example, the in-stream flow constraint was satisfied in all ASRs within a network before other use classes were met in any ASR of the network. Thus, spatial position within a network (e.g., upstream versus downstream) had no effect on the priority with which uses were satisfied.

B.6 Network simulation

Simulation of water allocation within each network was performed using the MODSIM (Labadie et al. 1984) network simulation package. Network simulation provides annual values of water flows in any network link, storage levels for each ASR, amount of water lost due to reservoir evaporation, and amount of water assigned to each water use class.

The various aspects, assumptions, specifications, etc. of the network simulations that have been described above, which allow simulation of water allocation within each network, are summarized here:

- Each ASR represents a node in the water network.
- Annual total water yield is accumulated over the entire ASR and is considered as a water input at the ASR node.
- Annual water supply for a given ASR is the sum of the annual water yield, the water inflow from ASRs located upstream, net water received from other ASR via trans-ASR diversions, and the water previously stored in the ASR itself.
- Water uses were grouped into three classes: in-stream-flow requirements, trans-ASR diversions, and consumptive uses.

- Each water use class was assigned a different priority, in the following order: in-stream flow requirements, trans-ASR diversions, and consumptive use.
- If ASR water yield plus contributions from upstream are insufficient to meet the requests of the three water classes in the ASR, water stored in the ASR is used if available, irrespective of the actual location of the reservoirs and demands within the ASR.
- Any water in an ASR not needed to satisfy the within-SR or downstream requests of the three water classes is stored up to the total available storage capacity.
- Water that cannot be stored in the most downstream ASR is released to a network sink (an ocean or Canada).
- Water loss due to reservoir evaporation is estimated by assigning to each reservoir an area-volume relationship and multiplying aggregate reservoir surface area by an annual potential evaporation rate for the ASR.

Modeling demand and supply at the ASR level, it should be noted, will fail to realistically represent conditions in some localized areas within an ASR. Perhaps the most likely instance of this scale-dependent failure is where a major demand area is located in the upper reaches of an ASR. Such a location would in fact—in the absence of pumping water uphill—place the area upstream of the bulk of the water supply of the ASR, although in the simulation the full supply of the ASR would be available to meet demands within the ASR.

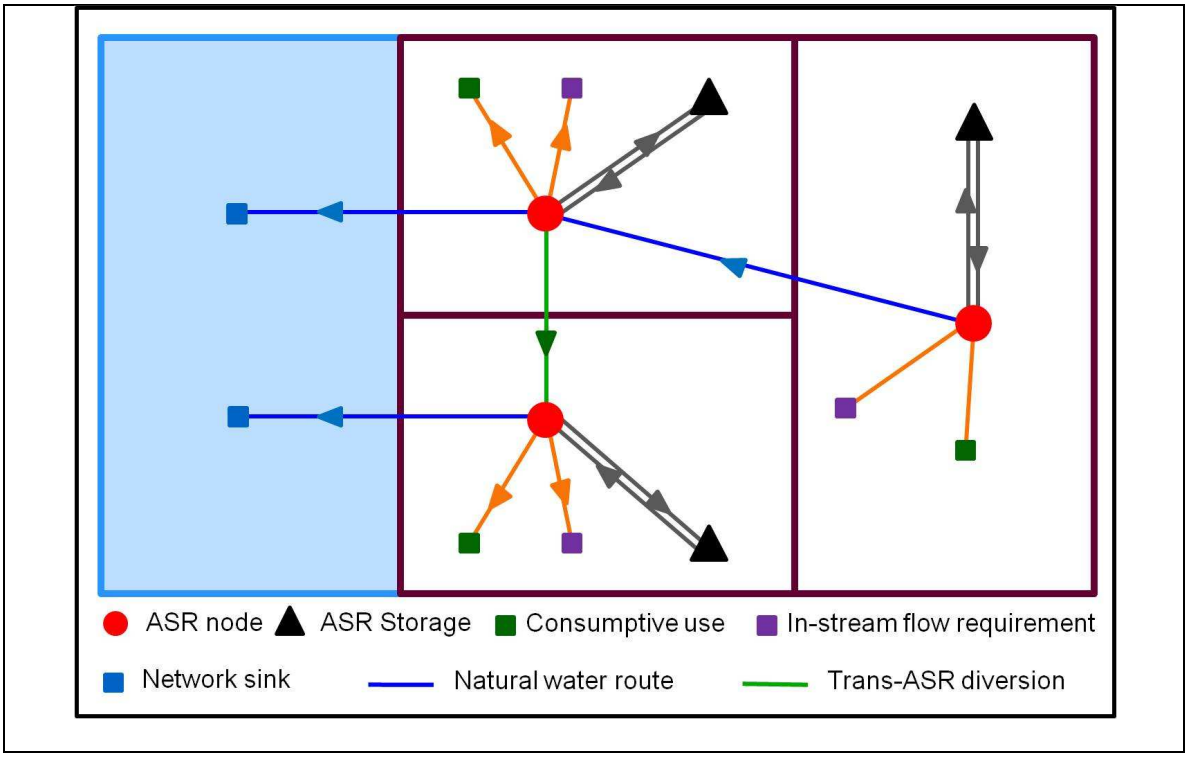


Fig. B.1 Schematization of a network containing three ASRs and including both natural and artificial water routes.

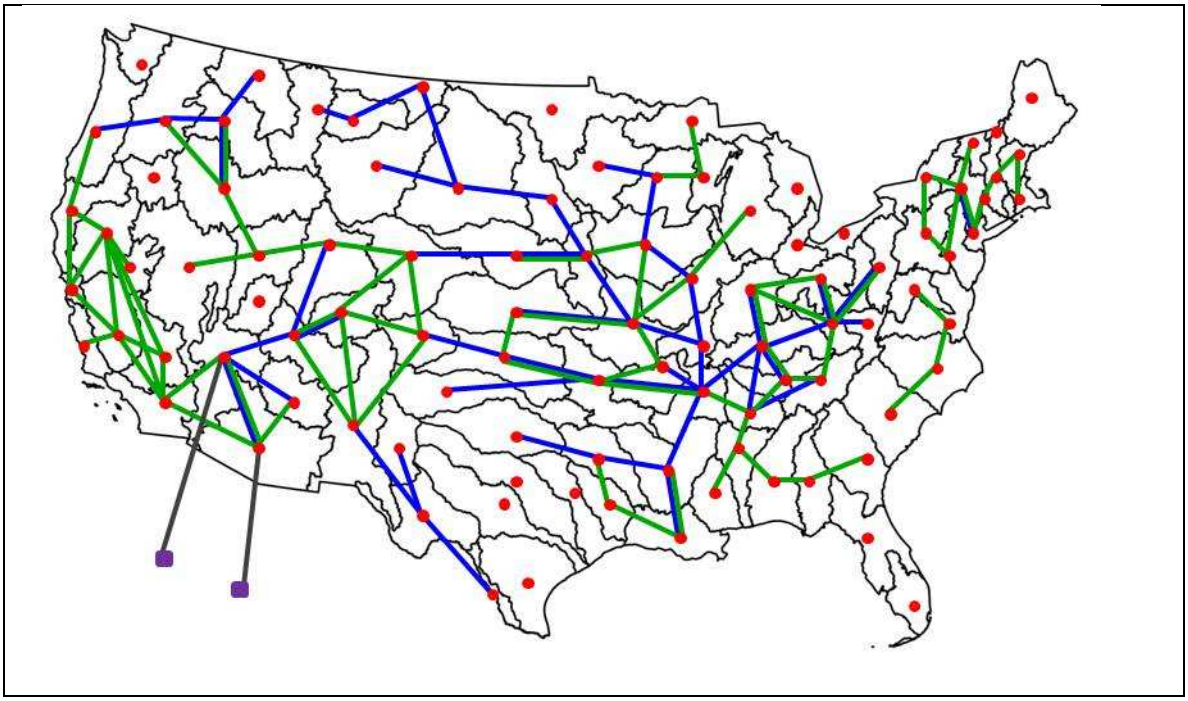


Fig. B.2 Water networks across the conterninous U.S. at the ASR level. Natural links are indicated with blue lines, artificial links (water diversions) with green lines. Grey lines indicate diversion to Mexico.

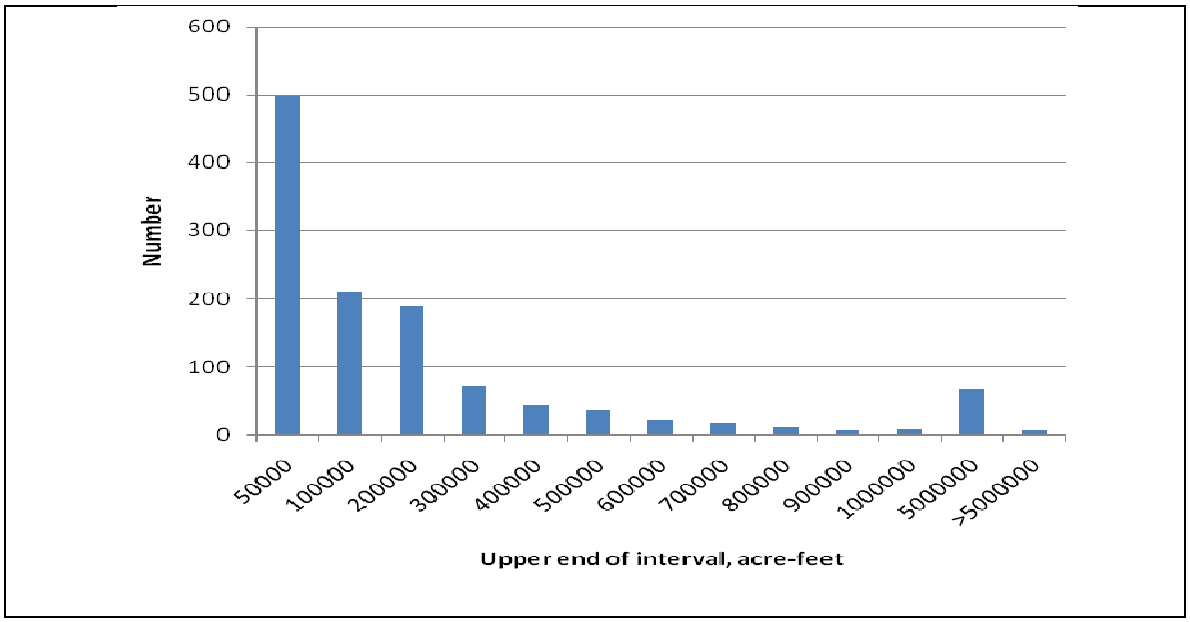


Fig. B.3 Storage volume size distribution of 1196 reservoirs.

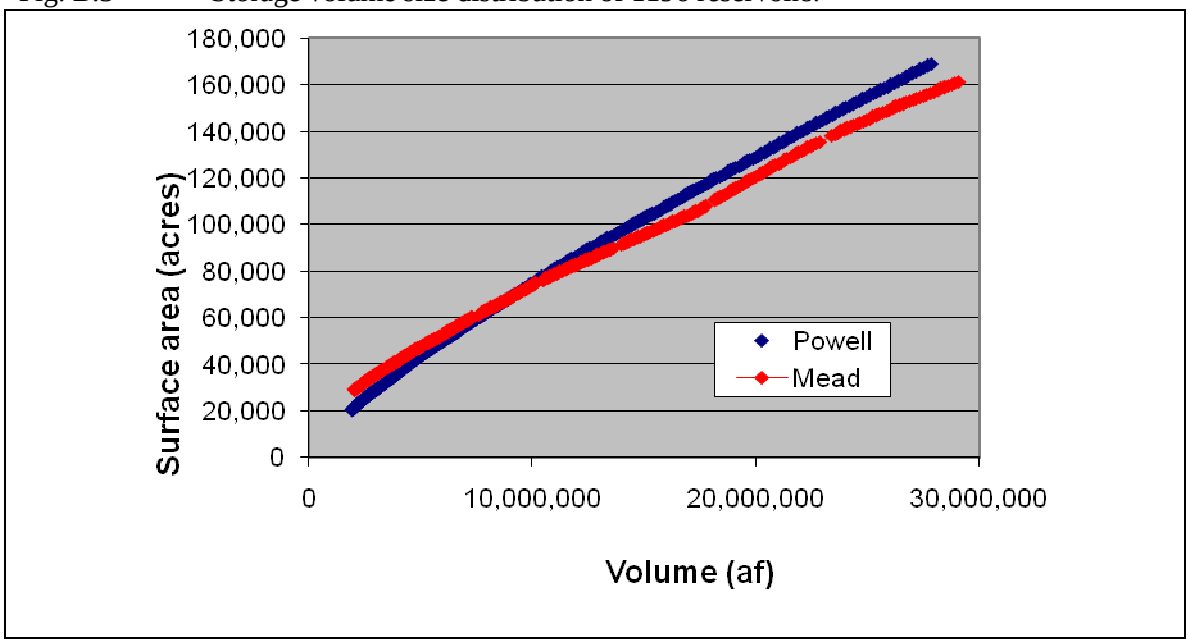


Fig. B.4 Surface area-to-storage relationships for Lakes Powell and Mead.

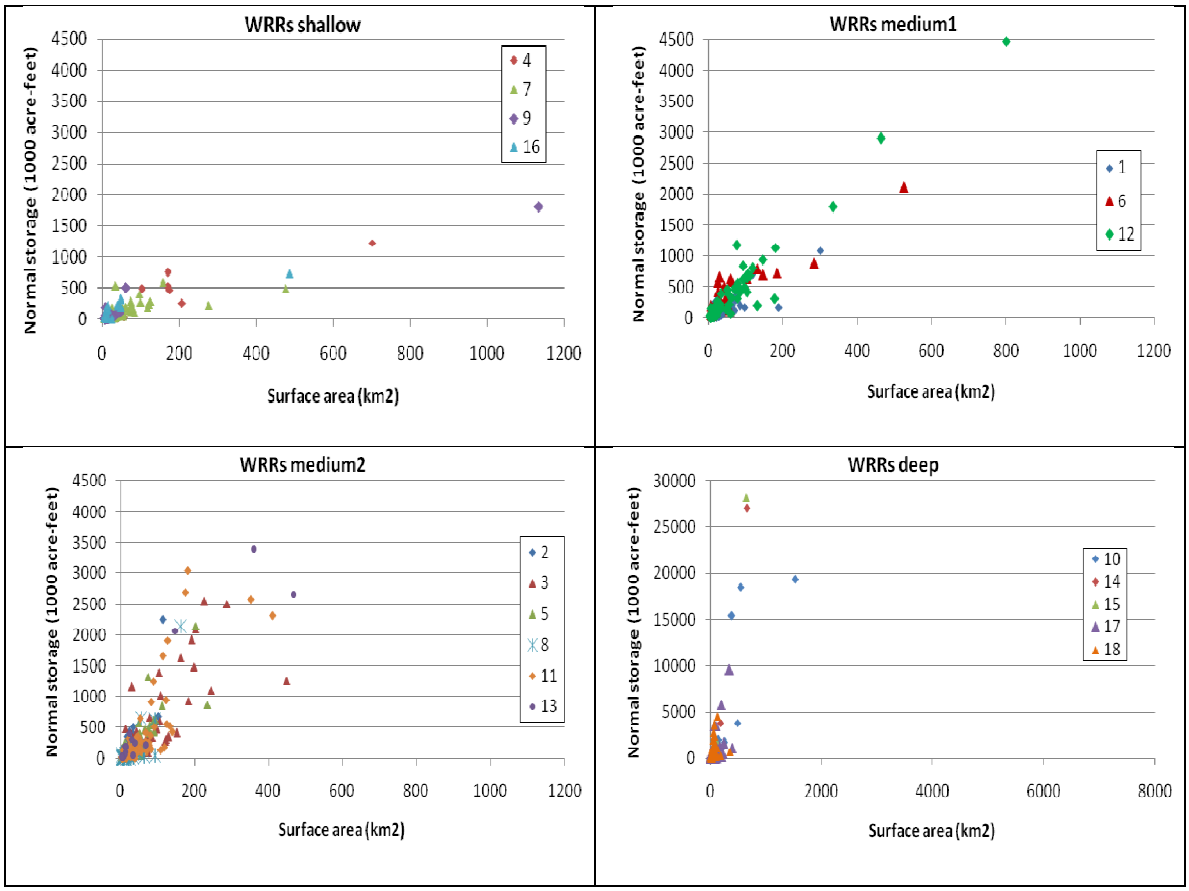


Fig. B.5 Groups of surface area-to-storage relationships for the conterninous U.S.

Tab. B.I Normal surface areas and storage volumes of reservoirs by ASR.

ASR	Number of reservoirs	Surface area (km^2)	Storage volume (acre-feet)	Storage volume (km^3)
101	53	2983	9,161,932	11301
102	11	303	339,694	419
103	1	14	114,000	141
104	4	280	1,450,200	1789
105	14	270	1,069,773	1320
106	1	7	3,500	4
201	14	271	1,914,887	2362
202	9	100	880,684	1086
203	6	70	950,208	1172
204	11	184	1,533,703	1892
205	8	108	731,648	902
206	1	7	33,700	42
207	8	337	4,763,164	5875
301	9	441	2,635,134	3250
302	12	2295	10,745,641	13255
303	14	2024	13,326,110	16437
304	12	228	498,205	615
305	1	15	33,324	41
306	10	781	4,450,780	5490
307	15	847	5,292,915	6529
308	14	400	2,521,472	3110
309	2	145	387,538	478
401	7	185	374,031	461
402	7	1049	2,361,175	2912
404	11	185	255,805	316
405	8	108	105,470	130
406	1	5	17,780	22
407	1	18	69,500	86
408	17	1711	4,910,008	6056
501	14	199	1,198,789	1479
502	22	610	3,481,610	4294
503	13	123	540,130	666
504	5	44	399,690	493
505	18	615	3,601,434	4442
506	16	185	781,314	964

507	9	875	6,347,200	7829
601	20	758	6,560,611	8092
602	8	1250	5,422,900	6689
701	19	1161	1,272,053	1569
702	30	1036	2,597,039	3203
703	18	950	1,867,580	2304
704	16	490	1,576,975	1945
705	9	378	1,293,851	1596
801	4	214	803,180	991
802	27	970	4,875,232	6014
803	6	60	85,600	106
901	27	1567	3,115,572	3843
1001	5	194	617,306	761
1002	11	445	4,658,098	5746
1003	3	397	15,502,400	19122
1004	7	269	3,059,354	3774
1005	13	2663	40,713,453	50219
1006	4	646	4,232,535	5221
1007	27	993	4,895,080	6038
1008	6	63	365,996	451
1009	1	29	144,600	178
1010	18	722	3,191,245	3936
1011	10	500	2,926,538	3610
1101	6	694	10,585,000	13056
1102	12	179	1,311,262	1617
1103	8	239	1,098,580	1355
1104	34	1256	4,982,058	6145
1105	17	729	4,309,899	5316
1106	16	667	2,849,744	3515
1107	41	1688	6,626,673	8174
1201	17	5143	23,303,728	28745
1202	28	1524	8,222,684	10143
1203	31	651	3,819,774	4712
1204	19	643	4,206,219	5188
1205	10	582	2,286,848	2821
1302	9	516	5,729,860	7068
1303	2	391	3,692,180	4554
1304	3	114	310,384	383

1305	3	479	2,698,340	3328
1401	13	375	5,631,191	6946
1402	10	191	3,095,031	3818
1403	4	730	28,874,500	35616
1501	1	8	2,610	3
1502	5	1029	30,508,313	37631
1503	12	205	3,251,572	4011
1601	6	93	774,780	956
1602	0	0	0	0
1603	4	88	260,800	322
1604	7	607	1,587,300	1958
1701	15	1220	11,989,814	14789
1702	26	1694	19,047,780	23495
1703	36	1426	12,754,945	15733
1704	7	223	5,520,957	6810
1705	20	325	6,208,326	5578
1706	9	340	3,320,425	4096
1707	4	61	85,190	105
1801	9	634	4,295,705	5299
1802	31	1454	23,556,700	29057
1803	31	1358	23,171,777	28582
1804	4	46	663,983	819
1805	4	74	1,051,550	1297
1806	12	108	2,273,358	7403
1807	2	30	199,870	247
Total	1196	61590	471,219,056	11301

Tab. B.II Surface area-to-storage regression for Lakes Powell and Mead and potential evaporation rates.

	Intercept	Slope	Evaporation rate
	km2	km2 /Mm3	m/yr
Powell	59	0.01881	1.21
Mead	91	0.01613	1.98

Tab. B.III Reservoir categories characterization across the 18 WRRs.

Group	Description	WRRs	Slope coefficient	Adjusted R^2
1	Shallow	4, 7, 9, 16	0.39660	0.83
2	Medium 1	1, 6, 12	0.14481	0.90
3	Medium2	2, 3, 5, 8, 11, 13	0.09032	0.76
4	Deep (without Powell and Mead)	10, 14, 15, 17, 18	0.03868	0.65
4	Deep (with Powell and Mead)	10, 14, 15, 17, 18	0.02777	0.68

Tab. B.IV Trans-ASR diversions.

ASR		Acre-feet	$m^3 / 10^3$
From	To		
102	103	899	1,109
104	105	1,084	1,337
105	102	580	716
201	202	1,474,267	1,818,427
201	207	56,900	70,183
202	104	617	761
203	201	898,170	1,107,844
204	203	44,311	54,655
204	408	635,644	784,033
206	205	213	263
301	205	6,702	8,267
301	302	12	15
306	303	67,345	83,066
307	306	8,350	10,299
307	308	11,550	14,246
309	308	100,000	123,345
402	401	1,617	1,994
404	704	2,231,913	2,752,943
408	201	7,860	9,695
501	502	23,314	28,757
502	506	63	78
502	601	1,077	1,328
503	502	4,142	5,109
505	506	101	125
506	503	1,477	1,822
507	505	662	817

507	601	508	627
602	308	3,300	4,070
602	507	492	607
702	402	3,620	4,465
801	602	43	53
801	1104	23,195	28,610
803	802	12,589	15,528
1008	1009	41,400	51,065
1009	703	83	102
1010	1011	697	859
1011	703	617	761
1011	704	309	381
1011	1101	11,607	14,317
1101	1104	7,720	9,523
1102	1007	20,000	24,669
1103	1010	169	208
1103	1104	23	28
1201	803	65,000	80,174
1201	1107	9,397	11,591
1302	1102	834	1,029
1401	1007	7,727	9,530
1401	1601	80,153	98,865
1402	1007	357,453	440,899
1402	1102	143,454	176,943
1402	1302	1,156	1,426
1402	1403	130,509	160,975
1403	1302	93,479	115,301
1501	1503	8,860	10,928
1502	1503	1,906,977	2,352,152
1502	1806	4,441,333	5,478,142
1503	1806	0	0
1502	8599[a]	1,500,000	1,850,168
1503	8599	124,927	154,091
1603	1601	724	893
1604	1802	4,233	5,222
1702	1703	1,700	2,097
1703	1601	1,233	1,521
1704	1703	6,323	7,799

1801	1705	21,833	26,930
1801	1802	879,000	1,084,199
1801	1804	33,000	40,704
1802	1803	3,896,000	4,805,503
1802	1804	317,000	391,002
1802	1806	1,571,000	1,937,743
1802	1807	80,000	98,676
1803	1804	616,000	759,802
1803	1805	28,000	34,536
1803	1806	0	0
1807	1806	360,000	444,040

[a] Mexico.

Tab. B.V Water use drivers.

Population
Income
Electric energy
Irrigated acres
Temperature
Precipitation
Potential evapotranspiration

Tab. B.VI Categories of water uses.

Domestic and Public
Industrial and commercial
Thermoelectric
Irrigation
Livestock
Aquaculture

Appendix C Future climatic and socio-economic projections

C.1 Overview

Estimating future vulnerability of U.S. water supply to shortage requires projections of water supply and demand. Climate directly affects both supply and demand. In addition, population and economic conditions directly affect demand. Because future climate, population, and economic conditions are uncertain and may take a variety of paths, we project supply and demand for alternative scenarios of future conditions. Further, climate under each scenario is projected using three different global climate models. The resulting set of nine different possible futures provides a range of estimates of demand and supply, and therefore of vulnerability. The mixture of results offers a rough indication of the uncertainty about future conditions.

This study is one of several assessments performed every ten years pursuant to the Renewable Resources Planning Act of 1974 (Langner in preparation-a). The assessments are commonly known as the RPA assessments. The act specifies that the assessments will project demand and supply conditions 50 years into the future, which in this case would be to 2060. However, because the potential effects of climate change on water supply and demand become more significant in the latter half of the century, we extend this assessment to beyond 2060.

In this chapter we give a brief description of the projected socio-economic scenarios (section C.2) and GCM models (section C.3) analyzed in this study. Projections of future precipitation, potential evapotranspiration and water yield are shown respectively in sections C.3.2, C.3.3 and C.3.4.

C.2 Future Socio-Economic Scenarios

The increasing globalization of the world economy and the possibility of substantial climatic change have created considerable uncertainty about future U.S. water supply and demand. One way to capture this uncertainty, adopted by the Intergovernmental Panel on Climate Change (IPCC), is to examine various possible future scenarios. The emissions scenarios examined by the IPCC differ in population growth, economic growth, and energy use, among other things, and thus in GHG emissions, CO_2 levels, and climatic changes. As a starting point for the RPA assessments, three scenarios—A1B, A2, or B2—each based on a different storyline, were chosen from the IPCC set (Nakicenovic, Alcamo et al. 2000).

The IPCC scenarios are internally consistent possible global futures that differ in many ways having to do with fertility rate, technological change, international trade, income growth, and energy development. Most importantly for this and the other RPA assessments, the scenarios specify alternative future population, income, and CO_2 levels, with implications for climatic variables that can be modeled using GCMs and spatial downscaling methods. The scenarios thus capture a range of potential futures that may substantially affect future water supply and use in the U.S.

Of the three scenarios, the A2 scenario is the most extreme and B2 the least extreme in terms of atmospheric CO_2 concentration. For example, year 2100 CO_2 concentrations are 856 ppm with the A2 scenario and 621 ppm for the B2 scenario, with the A1B scenario falling roughly midway between these extremes at 717 ppm (Tab. C.I). However, it is important to note that the CO_2 concentrations of these scenarios do not differ greatly until later in the 21st century. The CO_2 concentrations of the A2 and A1B scenarios are very similar in 2060 (at 572 and 580 ppm, respectively), although the B2 concentration begins diverging from the other two in about 2020 (and is 504 ppm in 2060). Clearly, one must extend the purview of the study beyond 2060 to observer the greatest differences in the scenarios and their impacts.

As with CO_2 concentration, global temperature differences among the scenarios are relatively small until the latter half of the 21st century. The multi-model projected global average surface warming projected by 2060 (relative to 1980–1999) are about 2º C for the A1B scenario and 1.9º C for the A2 scenario (a 2060 estimate for the B2 scenario was not available) (Tab. C.I). However, by 2100 the surface warmings of the A1B and A2 scenarios are projected to be 2.8º C and 3.6º C, respectively, with an increase of 2.4º C for the B2 scenario.

The population and economic projections of the IPCC scenarios do not use the most recent United States Census or recent economic data, and thus are somewhat dated. The IPCC projections were updated for the 2010 RPA assessments based on more recent information for the U.S. (Langner in preparation-b) (Tab. C.II). The population projections for the RPA assessment A1B scenario incorporate the 2000 census and presume a continuation of past changes in U.S. population growth level, and the A2 and B2 scenario populations were determined in relation to the revised A1B scenario by maintaining the relative differences among the original IPCC scenarios. Scenario A2 expects a higher population growth rate than the A1B scenario and the B2 scenario a lower growth rate (Fig. C.1). Scenario A1B expects much higher economic growth in the U.S. than do the other two scenarios (Fig. C.2).

C.3 Future Climate: General Circulation Models

The scenarios were used in combination with the following General Circulation Models (GCMs): the Canadian Centre for Climate Modelling and Analysis Third Generation Coupled Global Climate Model Version 3.1, Medium Resolution (CGCM31MR), for the A1B, A2 and B2 scenarios; the Australian Commonwealth Scientific and Industrial Research Organisation Mark 3.5 Climate System Model (CSIROMK35) for the A1B, A2 and B2 scenarios; the Japanese Centre for Climate System Research Model for Interdisciplinary Research on Climate Version 3.2, Medium Resolution (MIROC32MR), for scenarios A1B and A2; and the United Kingdom Met Office Hadley climate model (HadCM) for the B2 scenario (Tab. C.III). See Joyce et al. (in

preparation 2) for details. Monthly estimates of precipitation and temperature were available from all GCMs for the period 2001 to 2100 at the 5 arc minute grid level for the US. The specific variables were precipitation in millimeters (mm) and mean daily minimum and maximum air temperatures in degrees Celsius (°C). The use of these distinct and well-established GCMs ensured that the downscaled scenarios met the IPCC criteria for selecting scenarios for climate change impacts studies. The criteria include (1) consistency of regional scenarios with global projections; (2) physical congruence across climate variables; and (3) applicability to impact assessment, which is facilitated by the downscaled data being reported as change factors that can be referenced to locally observed climate data.

C.3.1 Downscaling and bias removal methodology

The spatial resolution of GCM output is too large to support most river basin studies and thus must be downscaled to a finer scale if it is to be used with water basins like ASRs or smaller watersheds. Further, GCM output commonly contains a bias, which is recognized by comparing the GCM estimates for a past period with field-based measurements for the same period. Downscaling and bias correction of the GCM data occurred in two steps. The first step, performed by Joyce et al. (in preparation 1), required downscaling the raw GCM simulations from their original spatial resolutions to the 5 arc minute scale (roughly a 10-km grid). The second step, performed by the authors of this document, consisted of further downscaling the data to match the 5-km grid resolution of this study for initial water yield estimation and to remove residual bias.

The first step began with converting the monthly values from the GCM datasets to monthly change factors, using the means of the simulated monthly values for the 30-year period 1961-1990, taken from the PRISM data set (Daly et al. 1994), as the baseline. In the case of temperature variables (monthly mean daily minimum and maximum air temperature), the change factor was computed as the arithmetic difference between the monthly value and the

corresponding 30-year mean (1961-1990) of the same temperature variable for that month. For monthly precipitation, the change factor was the ratio of the GCM-based monthly value to the 1961-1990 mean for that month. The change factors were then interpolated using the ANUSPLIN software (McKenney et al. 2006; Price et al. 2006) to create time series for the period over which the GCM simulations were carried out, extending to 2100. ANUSPLIN produced a fitted spline "surface" equation for each monthly variable, which was then used to create gridded data for that monthly variable at the 10-km grid scale. The bias is eliminated by matching the GCM output with historical observations for a period over which the two datasets overlap. For precipitation (*P*) as an example, the downscaling and bias correction procedure was as follows, where *Y* indicates a year from 2001 to 2100, *j* indicates month, *G* indicates GCM data and *H* indicates historical data:

- Compute deltas at GCM scale: $deltaP_{G,Y,j} = P_{G,Y,j} - P_{G,1961-90,j}$
- Downscale the deltas to 10-km grid
- Compute final values at 10-km scale: $P_{Y,j} = P_{H,1961-90,j} - deltaP_{G,Y,j}$
- Compute ETp from final values.

In the second step, the 10-km data were further downscaled, using simple spatial interpolation, to match our 5x5 km water yield grid. Then the data were corrected to remove a residual bias, using a recent eight-year period as the baseline. Specifically, observed precipitation and temperatures from the PRISM dataset were averaged over the period 2001-2008 and compared with averaged GCM values for the same period. Biases were estimated by computing the difference between the 2001-2008 averages predicted by the GCMs and the observed averages. Potential evapotranspiration in mm per day was computed using a modification of Penman's equation by Linacre (1977).

C.3.2 Future precipitation

Fig. C.3 shows mean annual values of precipitation by scenario-GCM combination for the U.S. for five time periods centered at years 2005, 2020, 2040, 2060, and 2080. The 2005 point represents a mean for the period 2001-2010, and the other points are 20-year means centered at those years. Average annual precipitation changes from 77 cm for 2005 to from 63 to 80 cm in 2060, with most scenario-GCM combinations showing little change over time in aggregate precipitation (Fig. C.3).

Looking at aggregate U.S. precipitation may mask regional differences. For example, with the A1B-CGCM combination mean precipitation is projected to consistently increase in most of the Northeast and in Texas and decrease in the West (Fig. C.7). Besides this large scale behavior, however, coherent patterns of changes in precipitation are not easily identifiable.

The disagreement about future precipitation found among scenario-GCM combinations at the aggregate scale (Fig. C.3) is also apparent at the ASR scale, as is seen by comparing the maps shown in Fig. C.7-Fig. C.15. Although the scenario-GCM combinations agree that precipitation will increase in the Northeast, there is little agreement elsewhere. Some consistent trends, however, are identifiable for specific scenario-GCM combinations, as seen in the following examples: for A2-CGCM, an increase in the North and decrease in the South); for B2-CGCM a decrease in the Central U.S. and in the Southwest; for CSIRO a decrease in the Southeast for the A1B scenario and in the Southeast and Northeast for the A2 and B2 scenarios; and for MIROC a large decrease in the Southeast, and decreases everywhere else with the exception of the Northeast. In a few cases precipitation is not expected to monotonically increase (or decrease) throughout the century, but rather alternate from periods of increase to periods of decrease. This is expected to happen in the southern East Coast for the A1B-CGCM combination, in the central Great Plains for A2-GCMC, in the southern Great Plains for A1B-CSIRO, and in Texas and eastern California for B2-HADN.

C.3.3 Future potential evapotranspiration

Future potential evapotranspiration is deeply influenced by the temperature projections, which are expected to rise in all of the scenarios-GCM combinations analyzed. Average (the midpoint between minimum and maximum temperature) annual temperature rises from 11.8° C for 2005 to from 13.5 to 15° C for 2060 depending on the scenario-GCM combination (Fig. C.4). Annual average potential evapotranspiration, therefore, is projected to rise as well, from 3.5 mm per day for 2005 to from 4.0 to 4.6 mm/d for 2060 (Fig. C.5). Spatial distributions of potential evapotranspiration increases for all the scenario-GCM combinations are presented in Fig. C.16-Fig. C.24. Noticeably, the only instances of expected decrease are limited to scattered areas and isolated periods, the most evident of which is the East Coast and eastern Great Plains for B2-CGCM in 2020. The projections from the MIROC model are the most extreme, with increases in average U.S. potential evapotranspiration that exceed 30% by 2060. The MIROC and CSIRO models predict large increases in the Southeast, especially for the A2 scenario. Changes predicted by CGC model are less dramatic and suggest increases of 15.8%, 25.9% and 14.1% respectively for the A1B, A2 and B2 scenarios by 2060. However, for the Colorado River Basin the CGC model projects increases in potential evapotranspiration of 21.1%, 25.1% and 18.3%, respectively.

C.3.4 Future water yield

Projections for precipitation and temperatures (and therefore potential evapotranspiration) given by the nine GCM-scenario combinations were used in the water balance model to estimate future traces of water wield for the period 2006-2090. Because the changes in projected water yield are the direct result of the changes in precipitation and potential evapotranspiration rates, some of the trends of future yield can easily be anticipated. The overall picture, in fact, shows water yield decreases throughout the 21st century (Fig. C.6). Using the CGCM model, for

example, decreases in mean annual yield across the U.S. of 15.8%, 21.7% and 17.2% are expected by 2060 with the A1B, A2 and B2 scenarios, respectively.

Maps of future water yield for the all GCM-scenario combinations are provided in Fig. C.25 through Fig. C.33 and confirm the overall decrease seen in Fig. C.6 (note that in the maps, the scale is truncated at +10 cm and -10 cm). In general, the magnitude of the decrease is larger in the eastern U.S., although areas of the central and western U.S. are expected to experience the largest percentage decreases. Besides the case of the northern East Coast and noticeably in the East for the B2-CSIRO case, increases in yield are often localized in small areas and only occur for a few scenario-GCM combinations. Knowing only the changes in the mean values of precipitation and potential evapotranspiration may not, by itself, suffice to indicate the direction of the change in yield. In some circumstances, in fact, the higher moments of the distribution functions of precipitation and potential evapotranspiration strongly affect the distribution function of yield, leading to some apparently counterintuitive results. This is the case in the lower Colorado River Basin, where an increase in yield (of over 20% by 2060) is projected for the A1B-CGCM future despite a projected decrease in precipitation and increase in potential evapotranspiration. This large percentage increase in yield (which, by the way, is only a fraction of a cm in absolute terms) is caused by the increases in the variance of both predicted precipitation and potential evapotranspiration. As a result of a larger variance of the climatic forcing, the distribution of extreme events can be affected so as to lead to an increase rather than a decrease in the average water yield. This is especially true in arid climates, as a consequence of the highly skewed distributions of precipitation and water yield (note that precipitation can increase, above past levels, during wet times, but cannot drop below 0 during dry times).

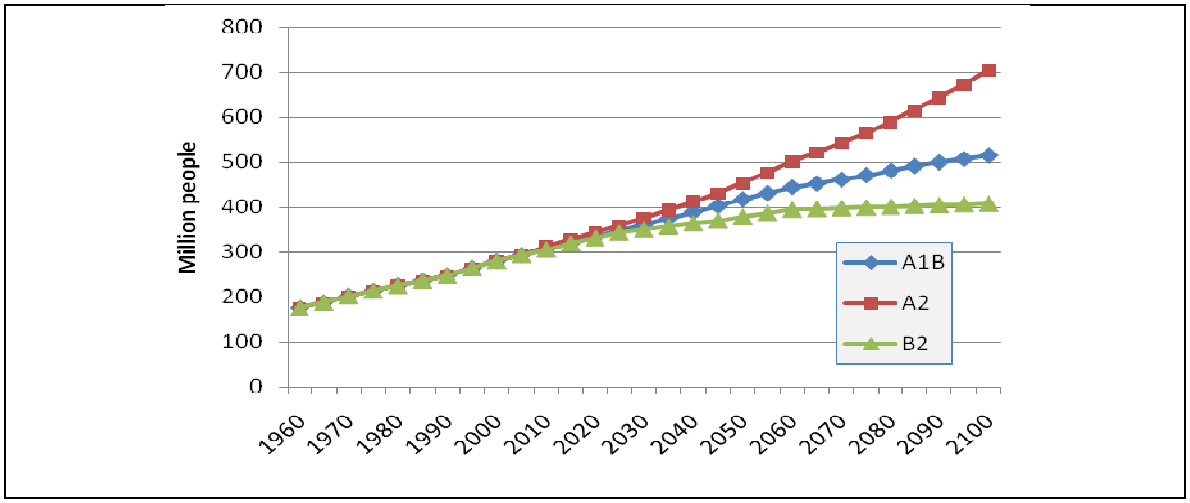


Fig. C.1 Past and projected population of the U.S.

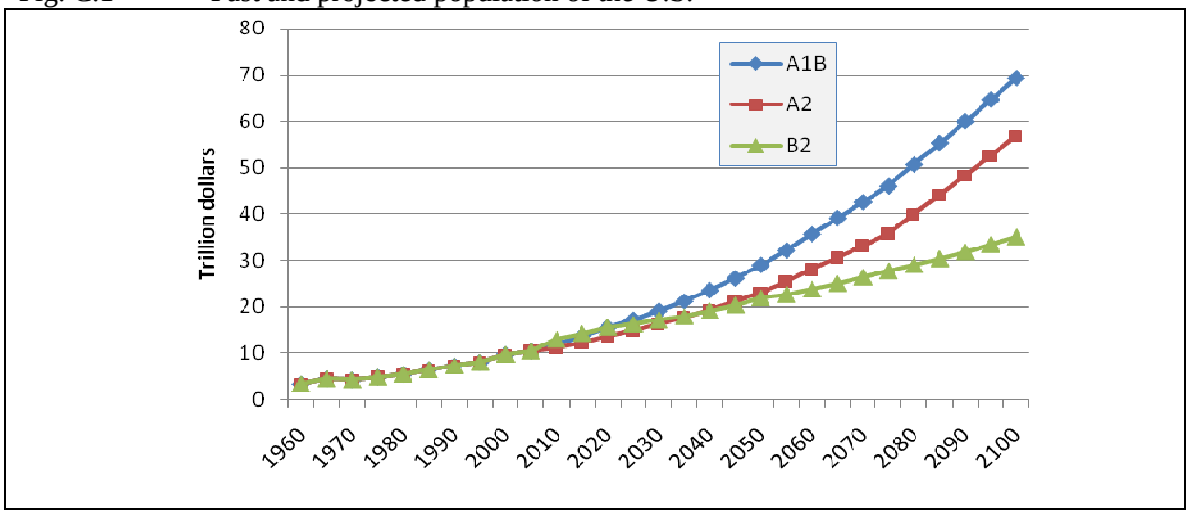


Fig. C.2 Past and projected total annual personal income of the U.S., 2006 dollars.

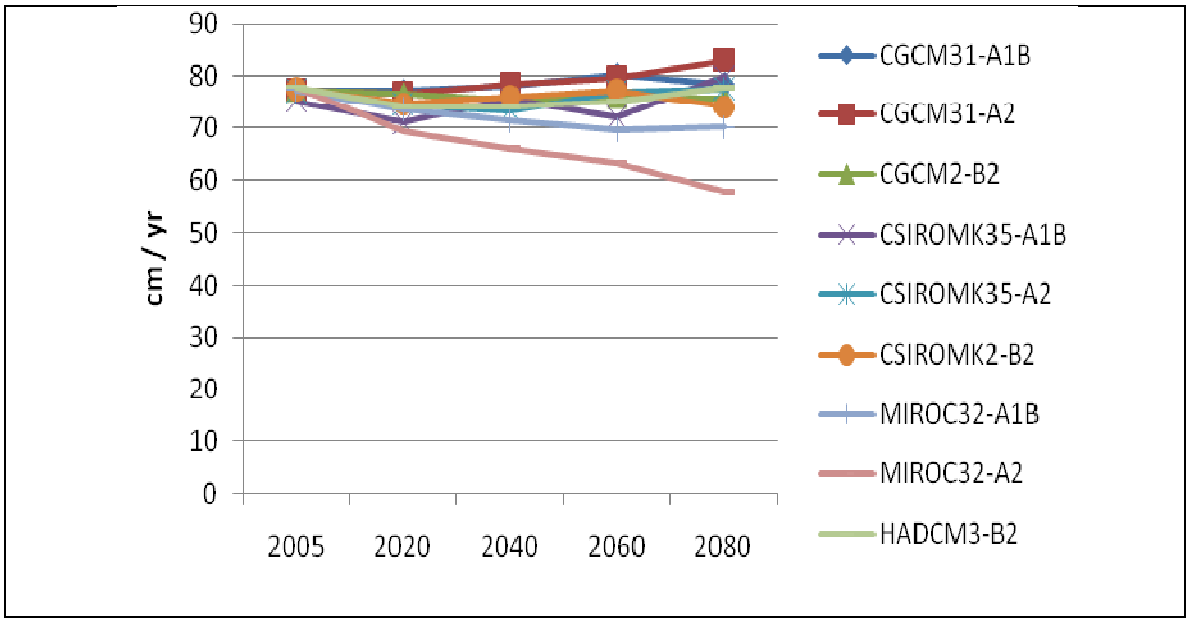


Fig. C.3 Mean annual precipitation in the coterminous US by scenario-GC combination.

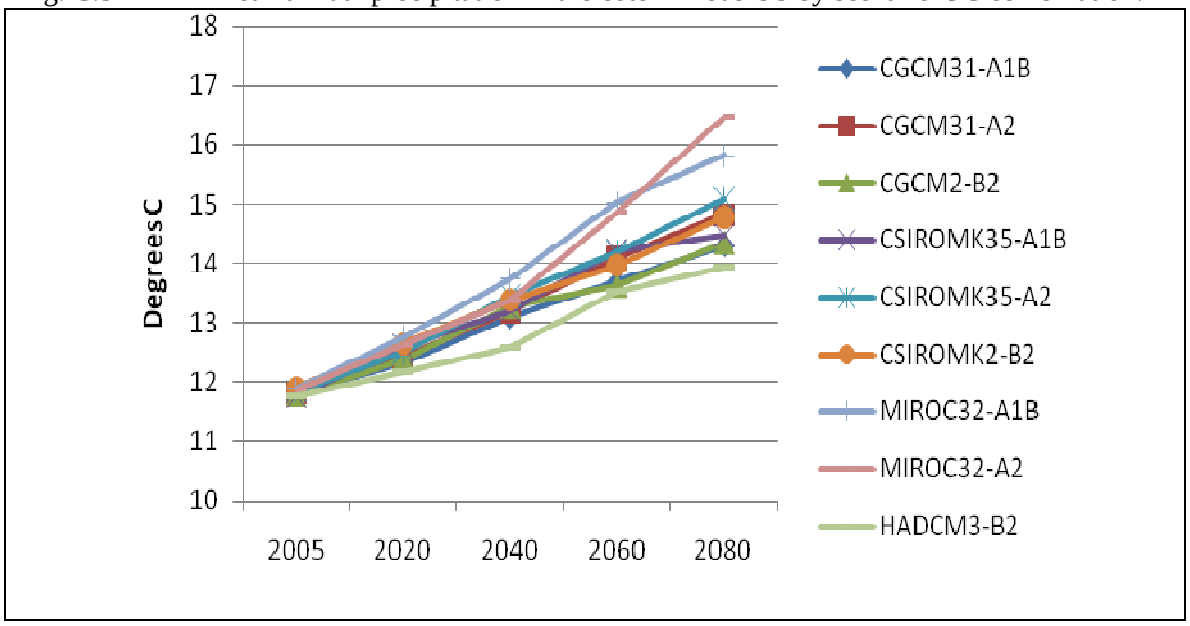


Fig. C.4 Mean annual temperature in the coterminous US by scenario-GC combination.

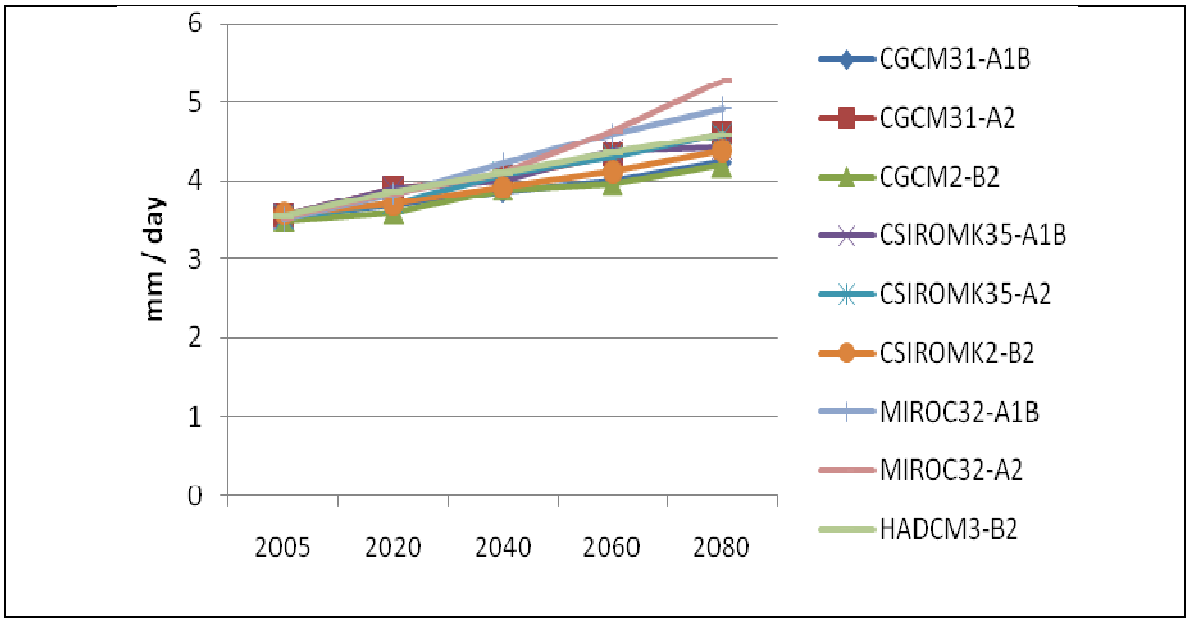


Fig. C.5 Mean annual *ETp* in the coterminous US by scenario-GC combination.

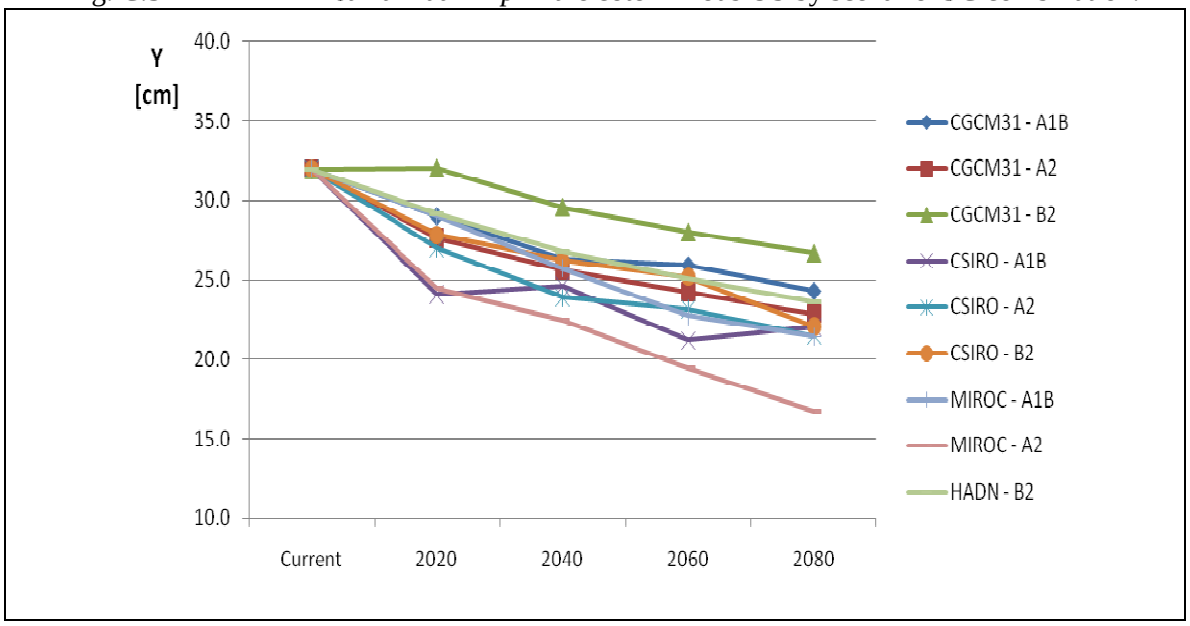


Fig. C.6 Mean annual Yield in the coterminous US by scenario-GC combination.

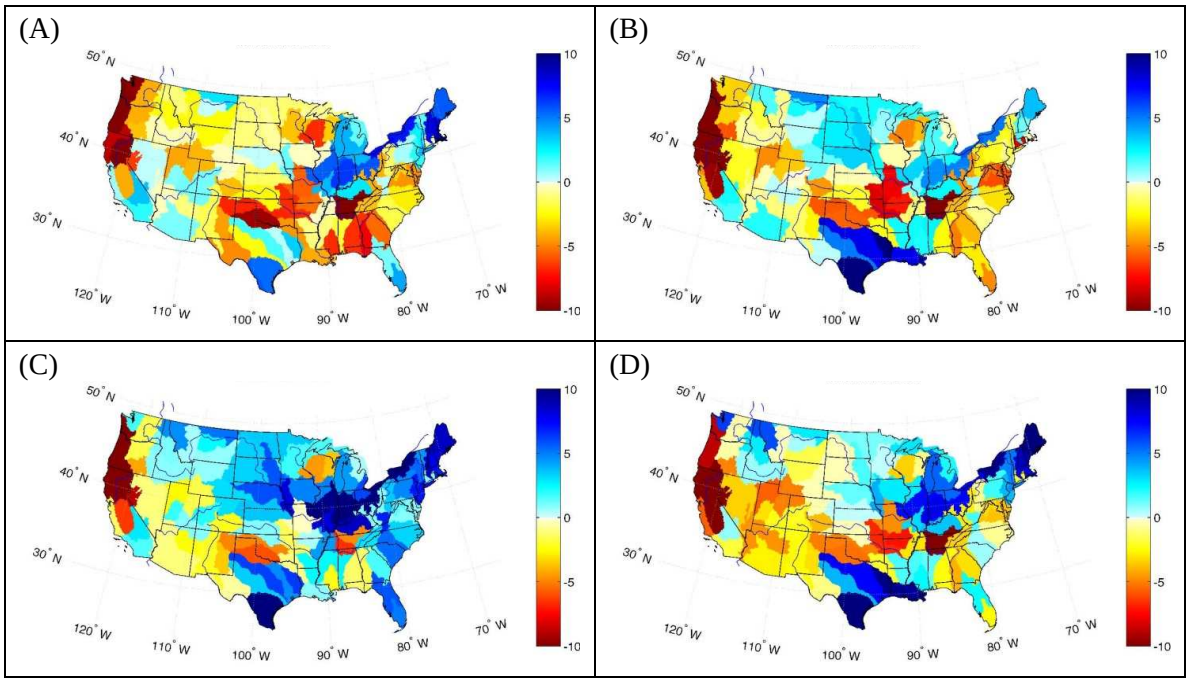


Fig. C.7 Changes in mean precipitation projected by the CGCM/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

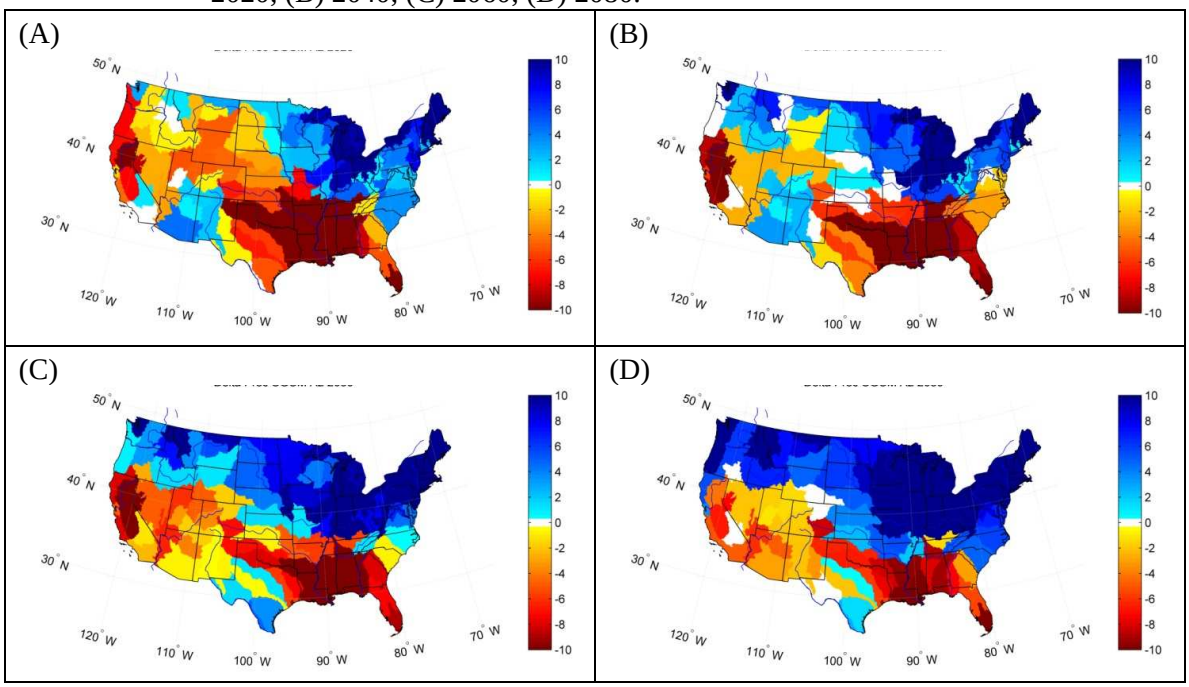


Fig. C.8 Changes in mean precipitation projected by the CGCM/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

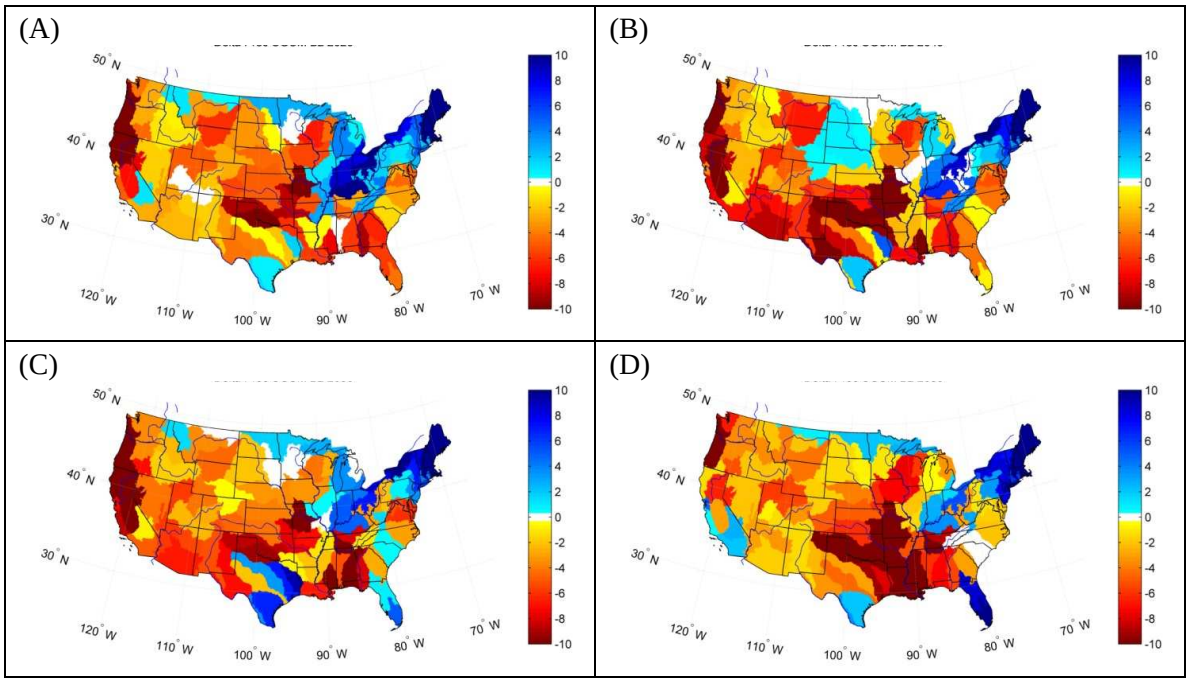


Fig. C.9 Changes in mean precipitation projected by the CGCM/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

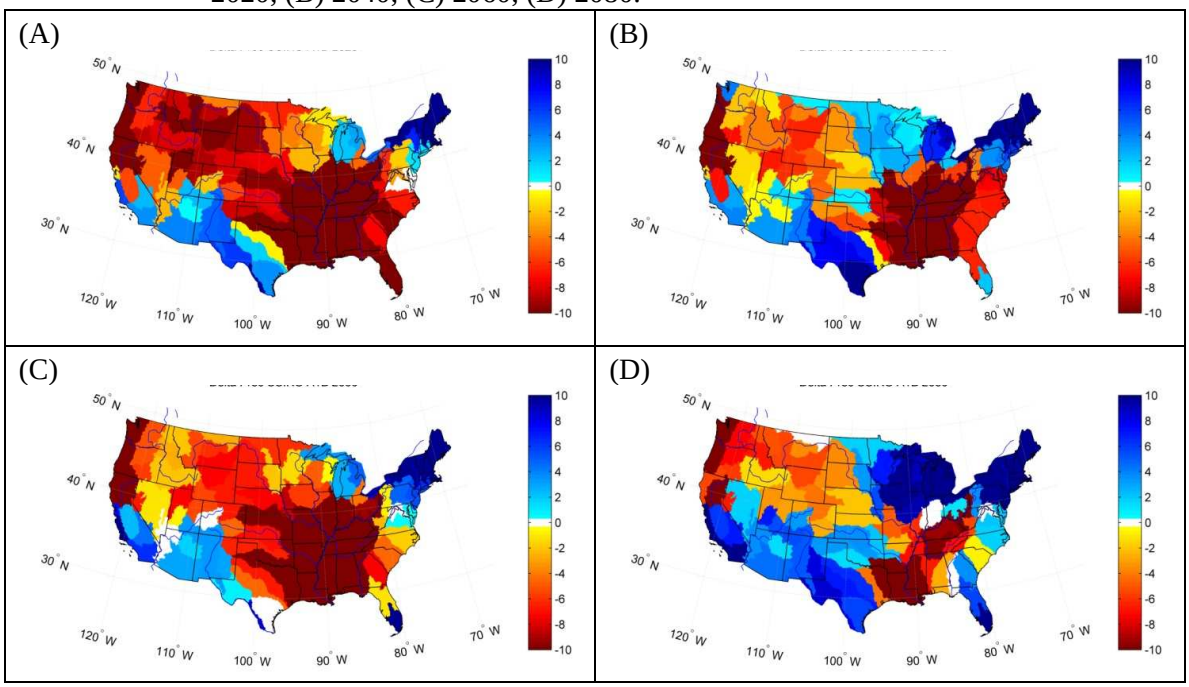


Fig. C.10 Changes in mean precipitation projected by the CSIRO/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

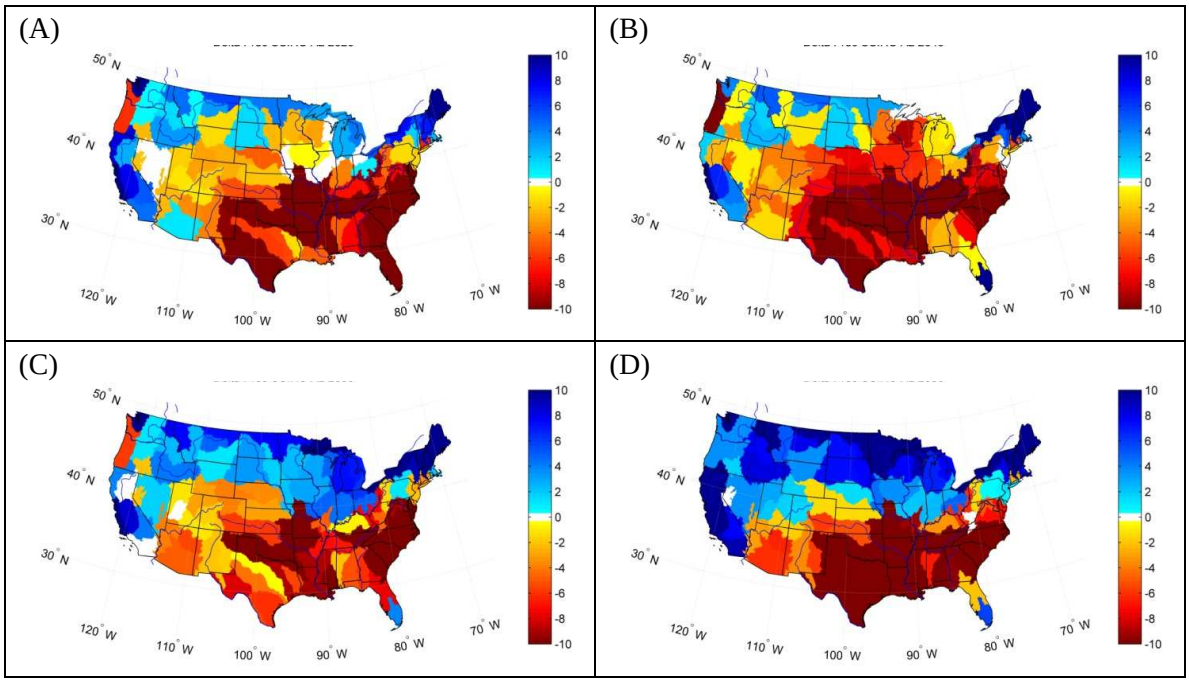


Fig. C.11 Changes in mean precipitation projected by the CSIRO/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

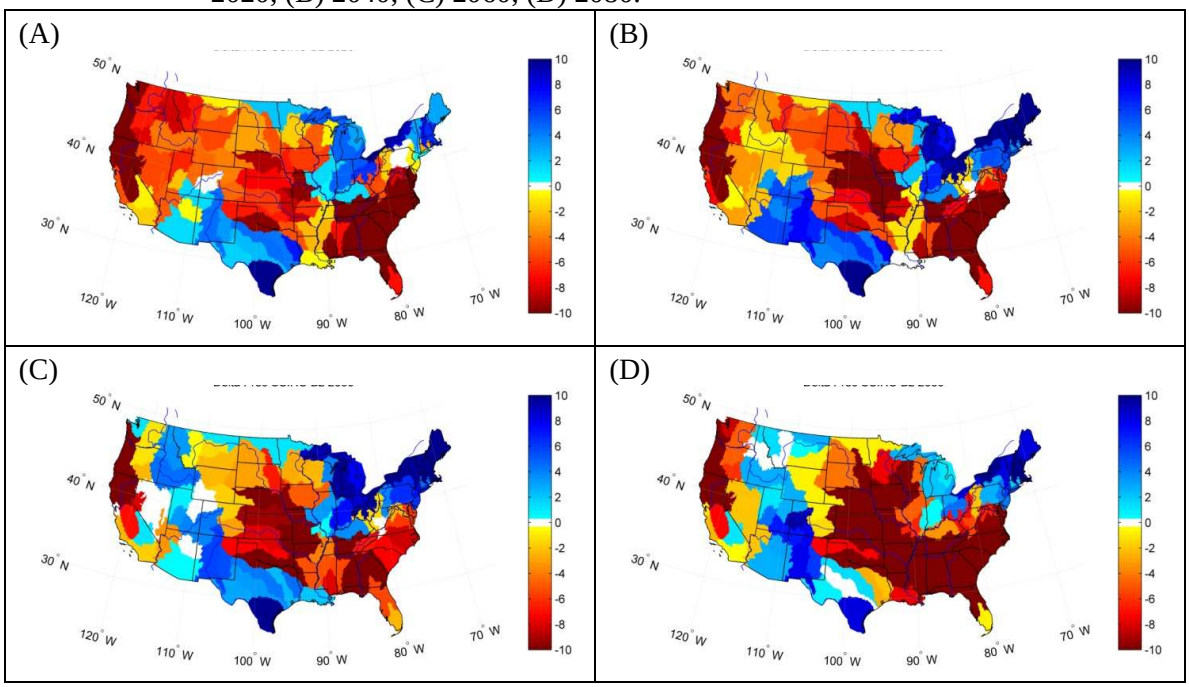


Fig. C.12 Changes in mean precipitation projected by the CSIRO/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

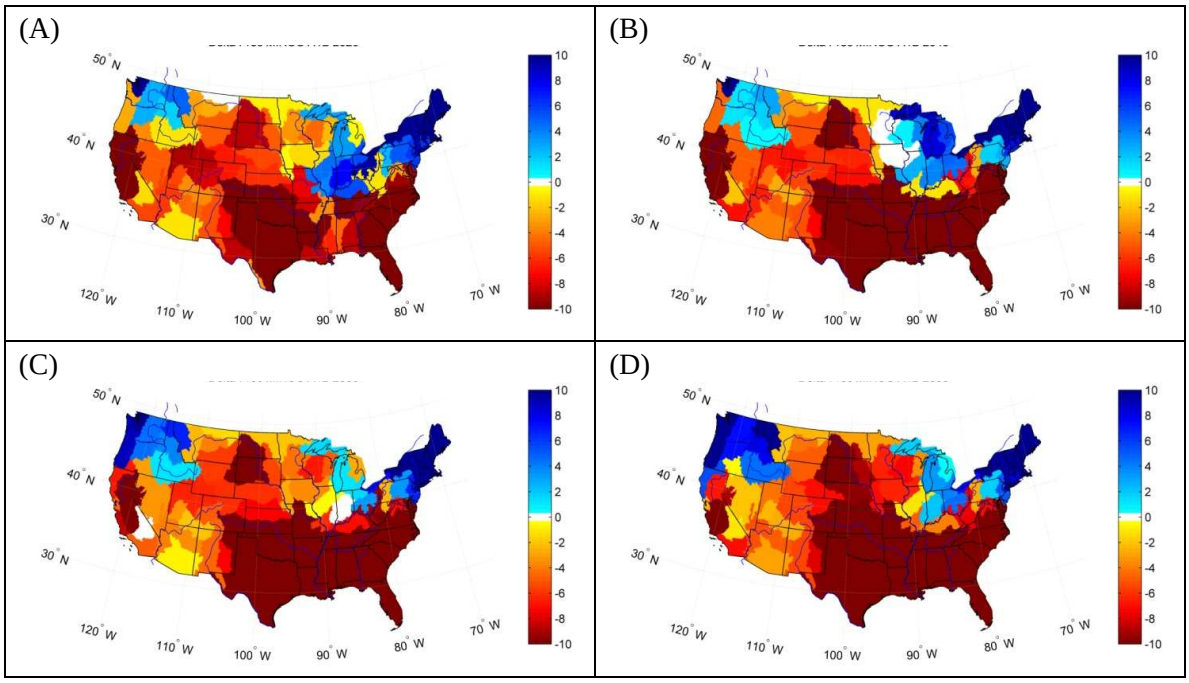


Fig. C.13 Changes in mean precipitation projected by the MIROC/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

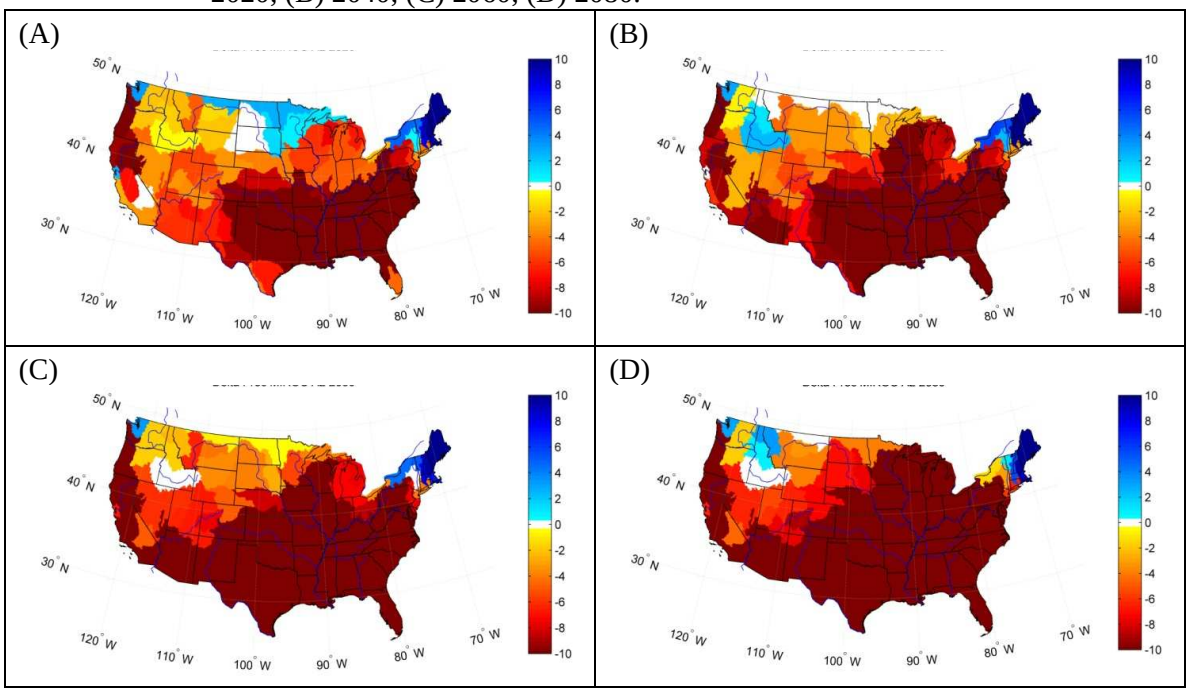


Fig. C.14 Changes in mean precipitation projected by the MIROC/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

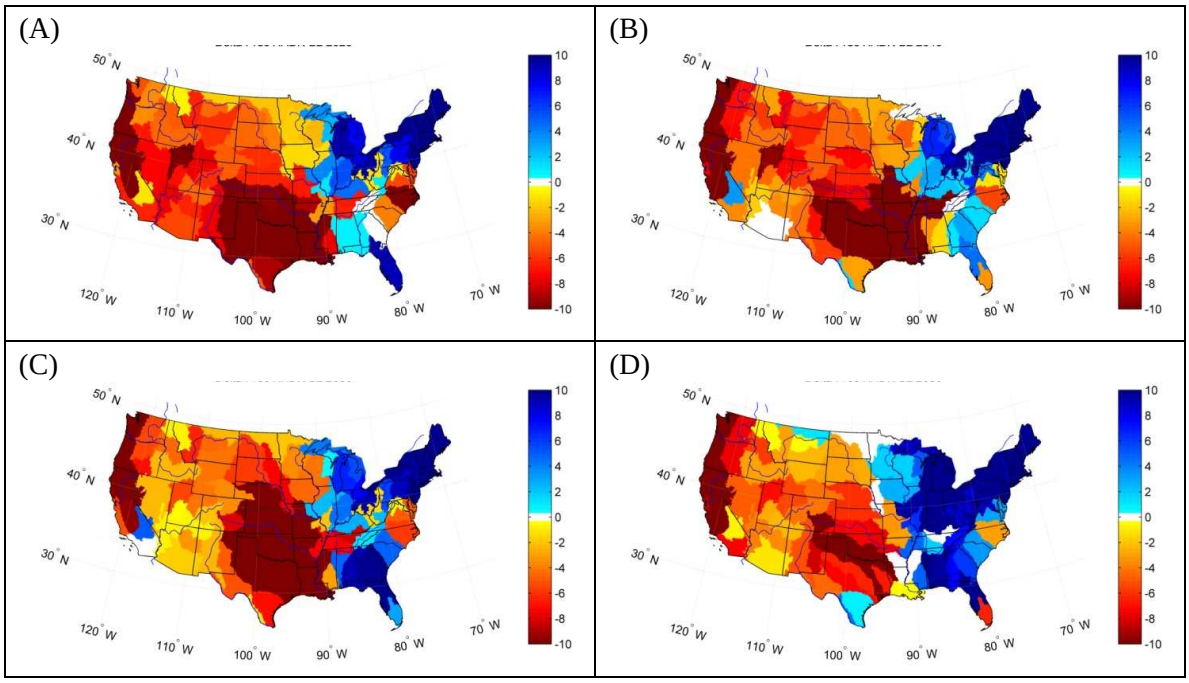


Fig. C.15 Changes in mean precipitation projected by the HADN/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

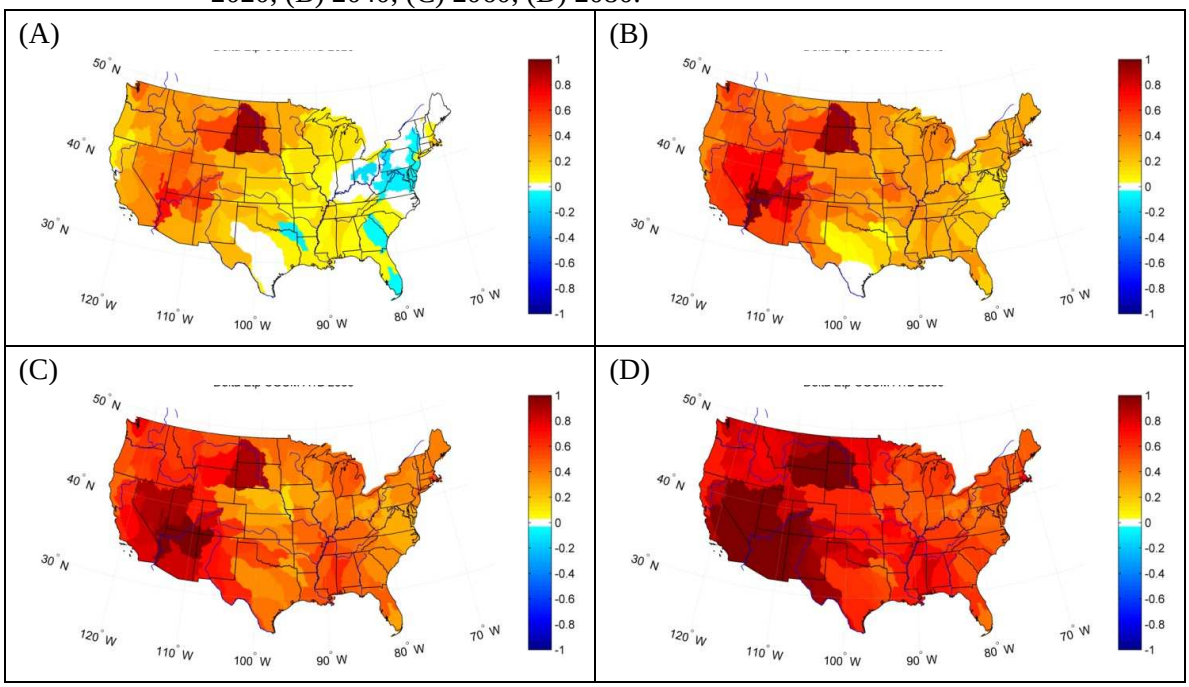


Fig. C.16 Changes in mean potential evapotranspiration projected (in mm/day) by the CGCM/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

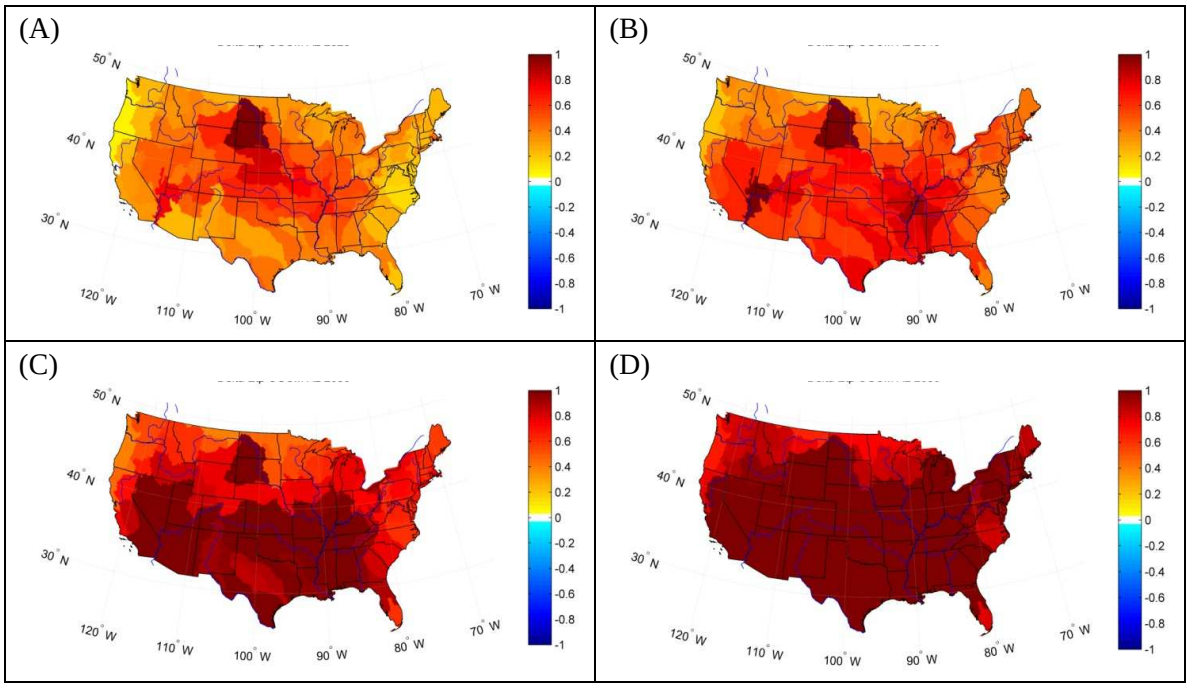


Fig. C.17 Changes in mean potential evapotranspiration (in mm/day) projected by the CGCM/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

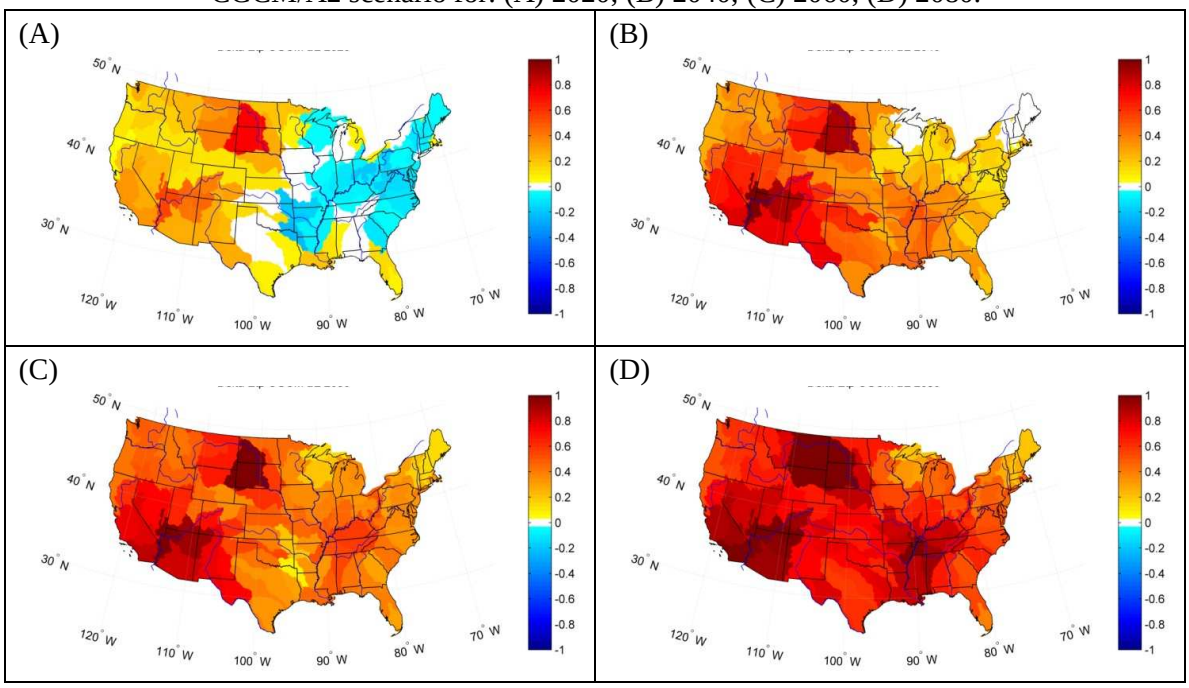


Fig. C.18 Changes in mean potential evapotranspiration (in mm/day) projected by the CGCM/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

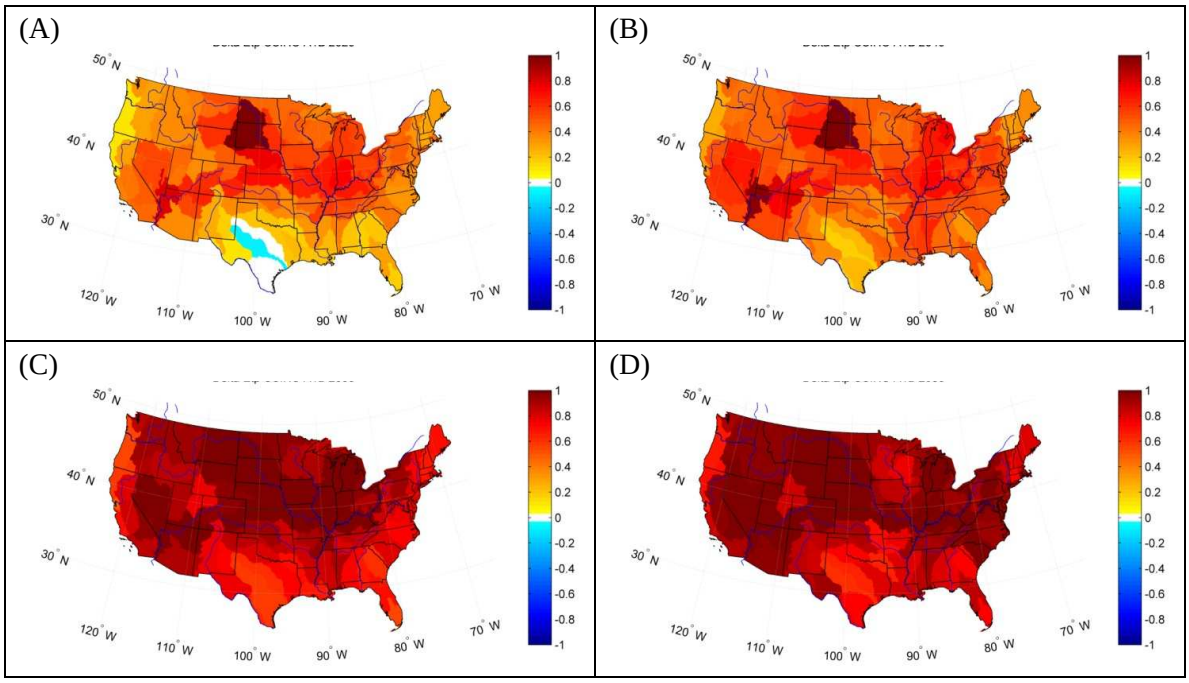


Fig. C.19 Changes in mean potential evapotranspiration (in mm/day) projected by the CSIRO/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

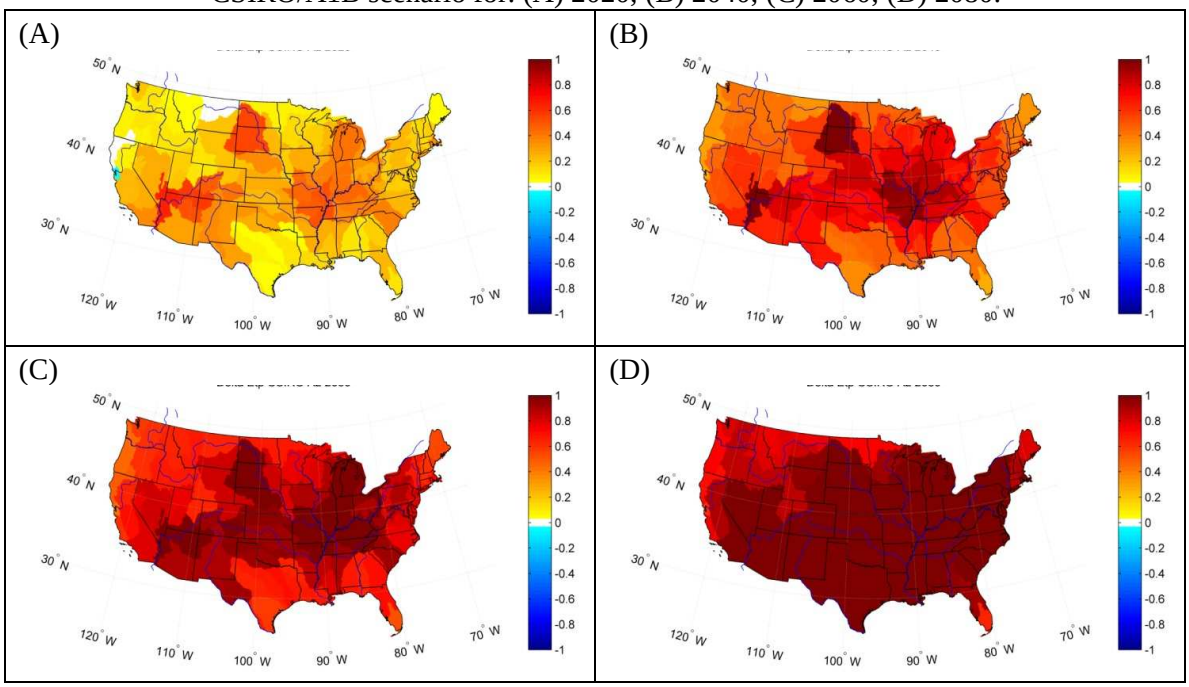


Fig. C.20 Changes in mean potential evapotranspiration (in mm/day) projected by the CSIRO/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

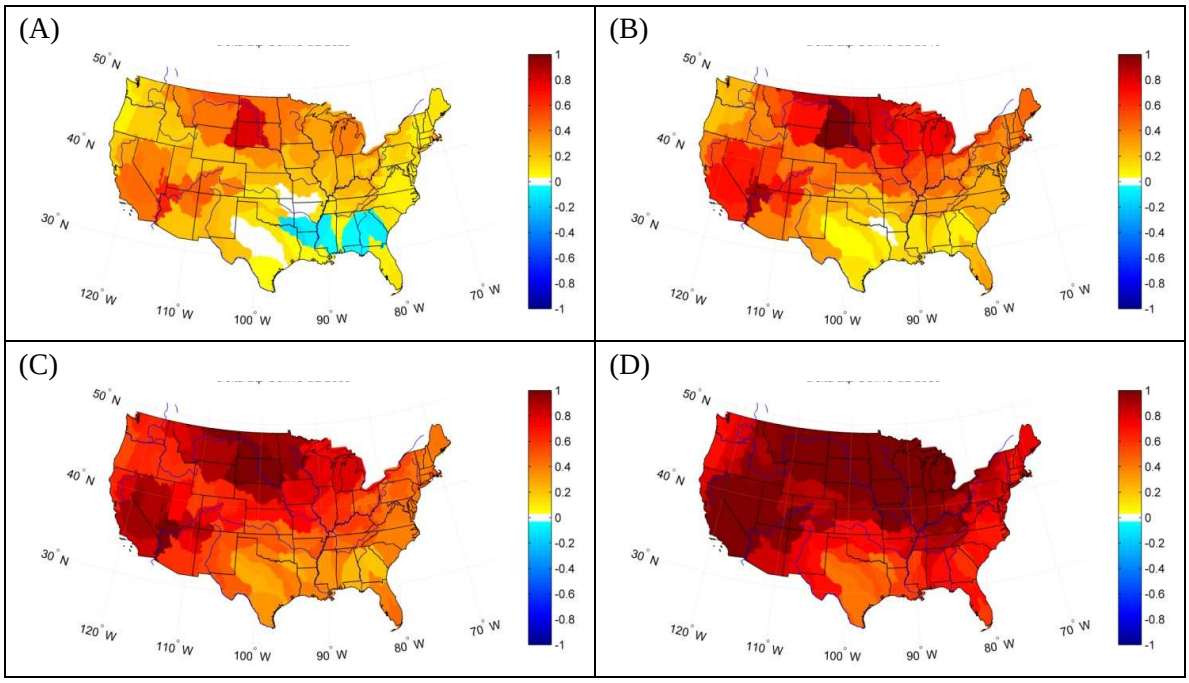


Fig. C.21 Changes in mean potential evapotranspiration (in mm/day) projected by the CSIRO/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

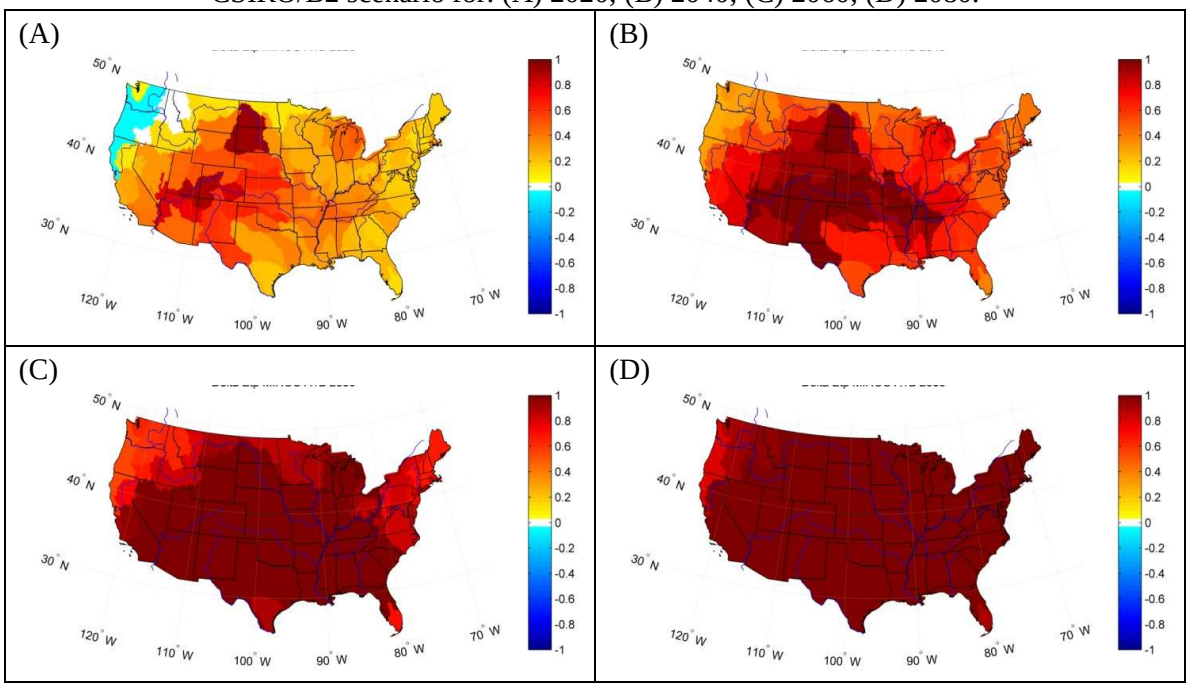


Fig. C.22 Changes in mean potential evapotranspiration (in mm/day) projected by the MIROC/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

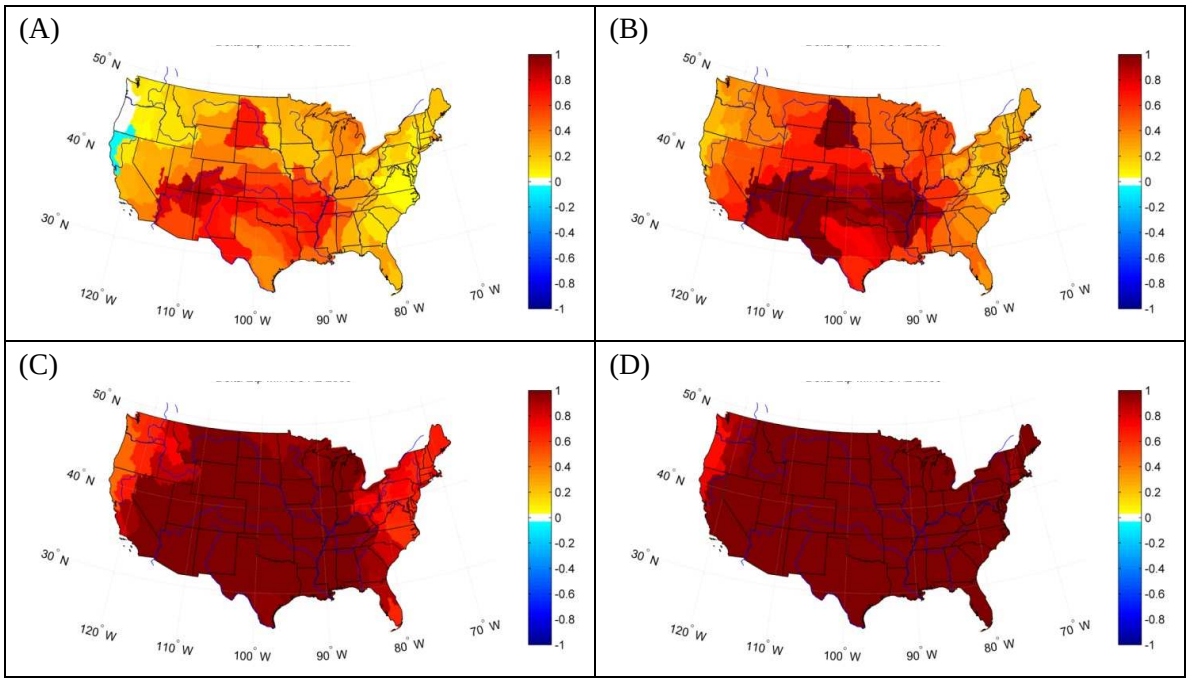


Fig. C.23 Changes in mean potential evapotranspiration (in mm/day) projected by the MIROC/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

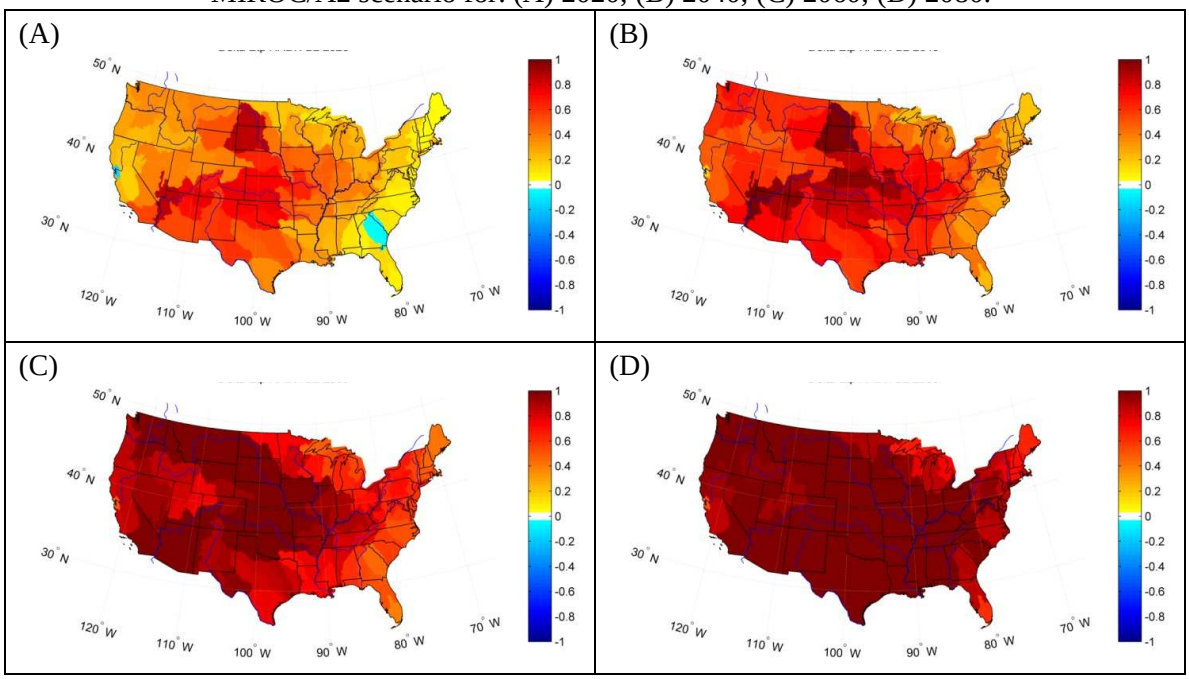


Fig. C.24 Changes in mean potential evapotranspiration (in mm/day) projected by the HADN/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

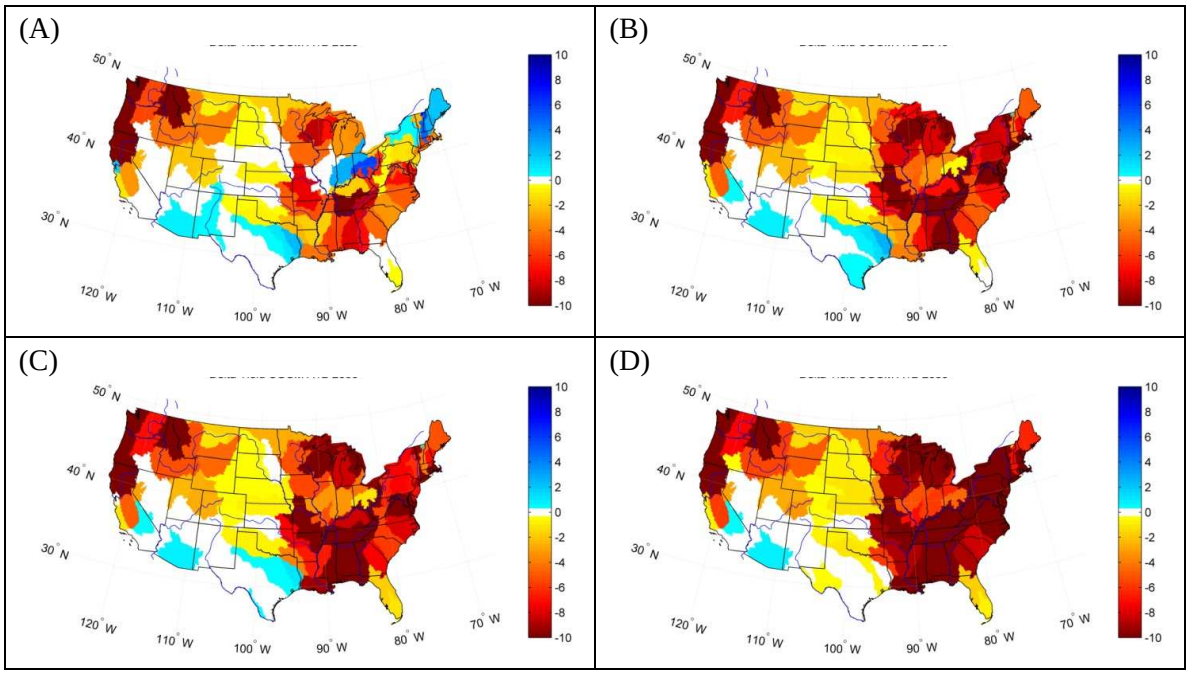


Fig. C.25 Changes in mean water yield for the CGCM/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

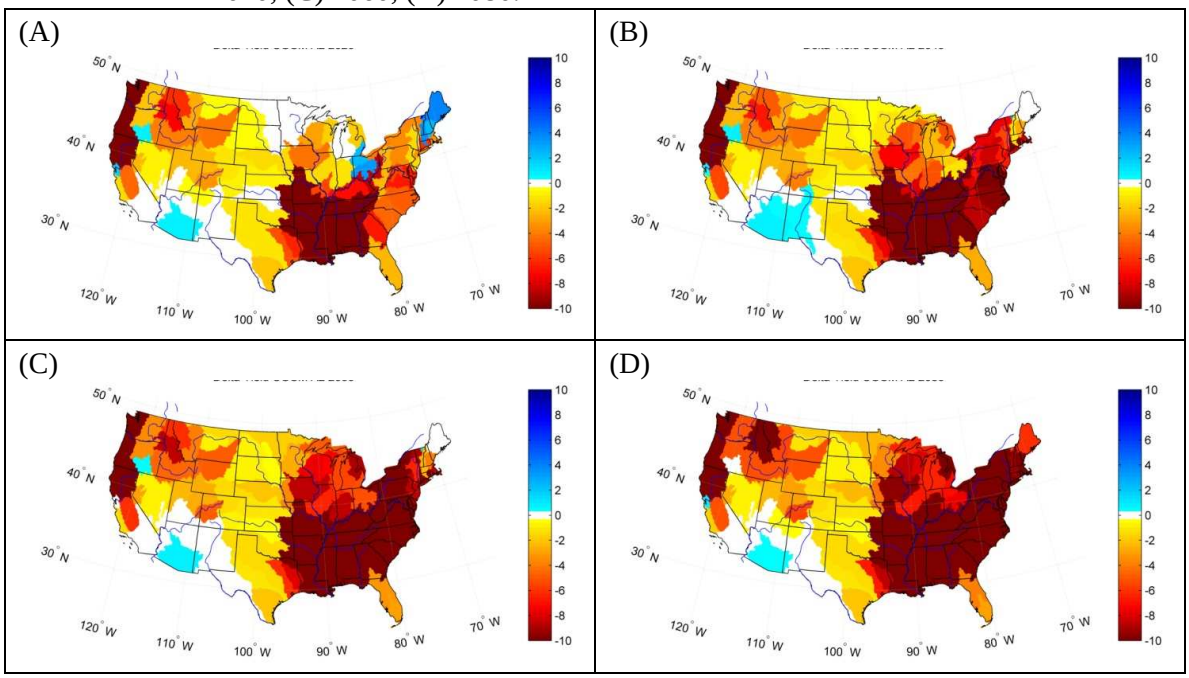


Fig. C.26 Changes in mean water yield for the CGCM/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

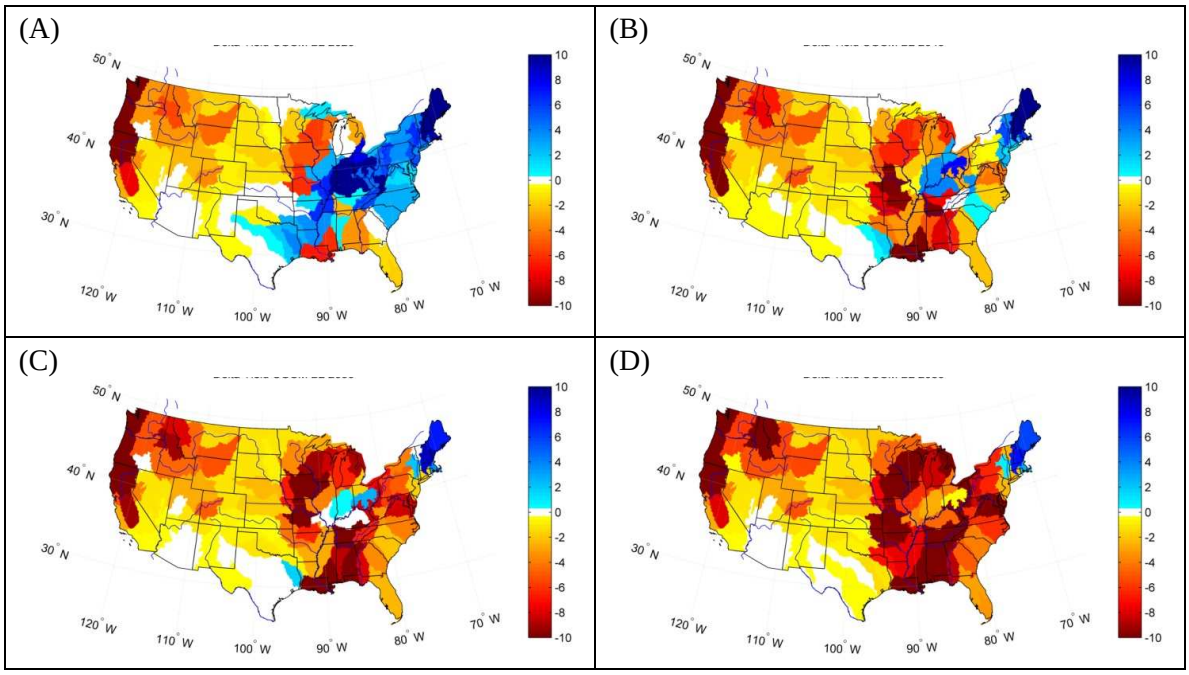


Fig. C.27 Changes in mean water yield for the CGCM/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

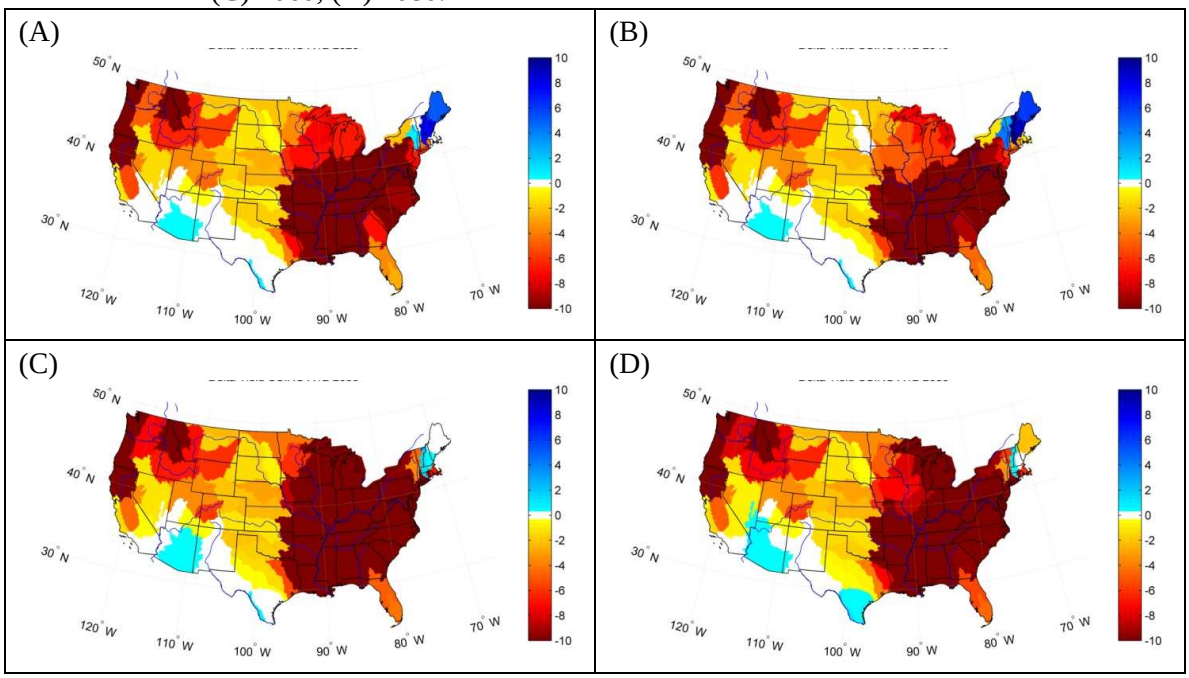


Fig. C.28 Changes in mean water yield for the CSIRO/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

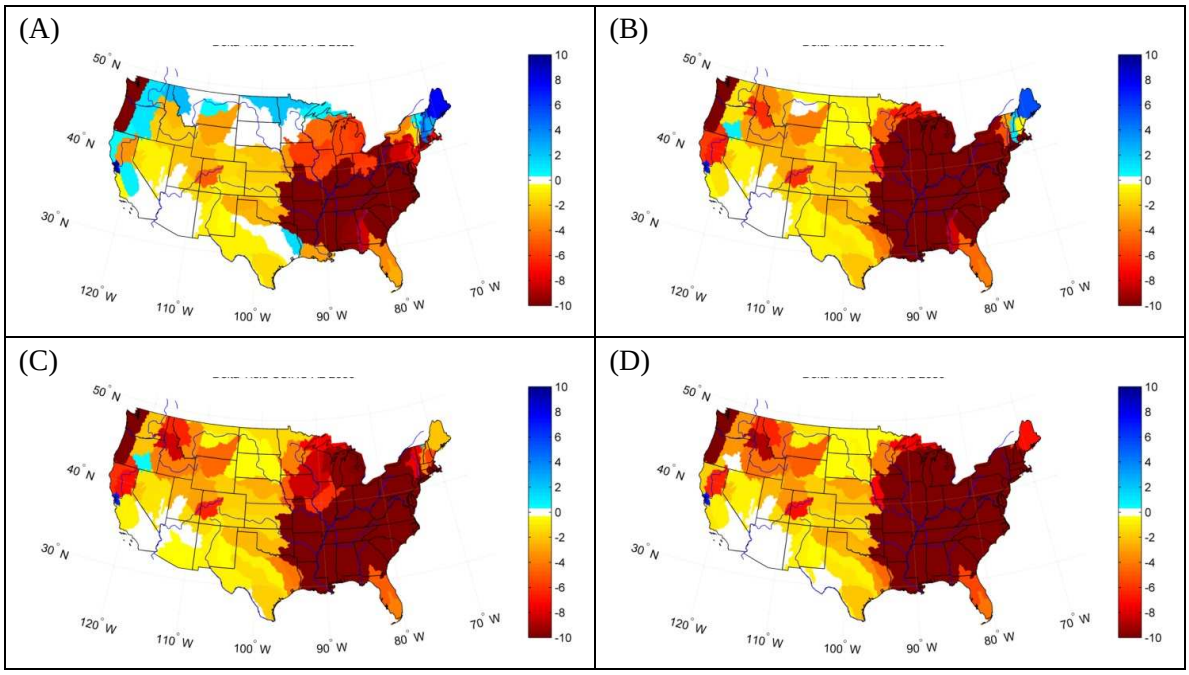


Fig. C.29 Changes in mean water yield for the CSIRO/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

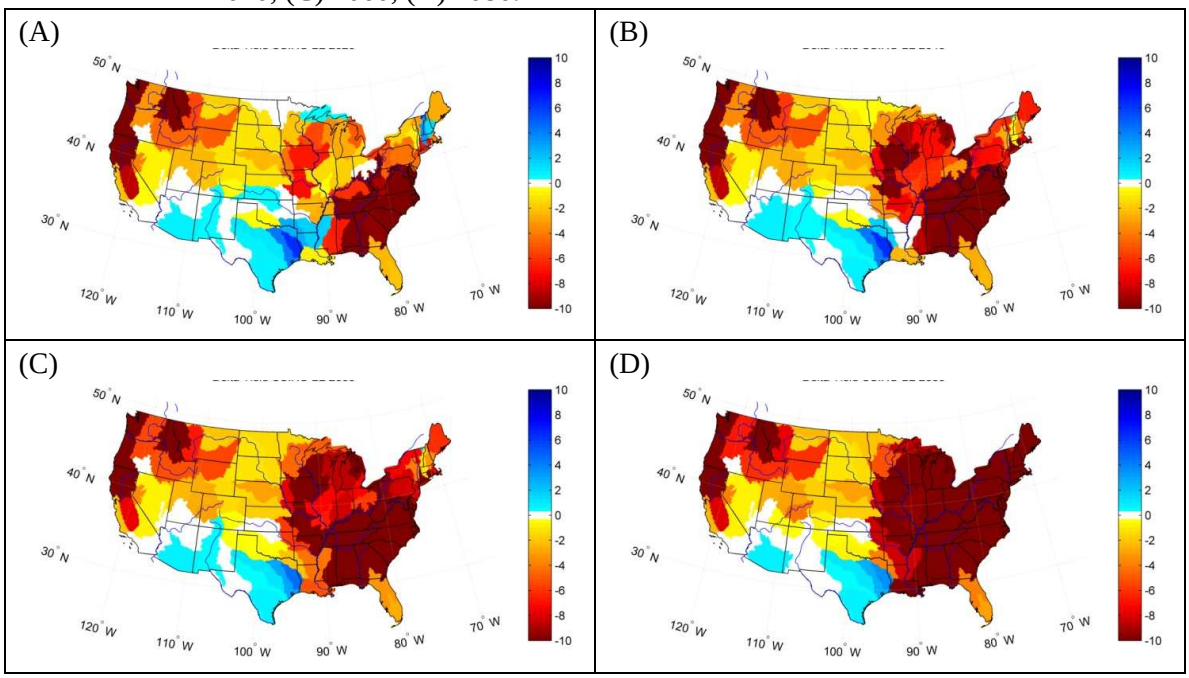


Fig. C.30 Changes in mean water yield for the CSIRO/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

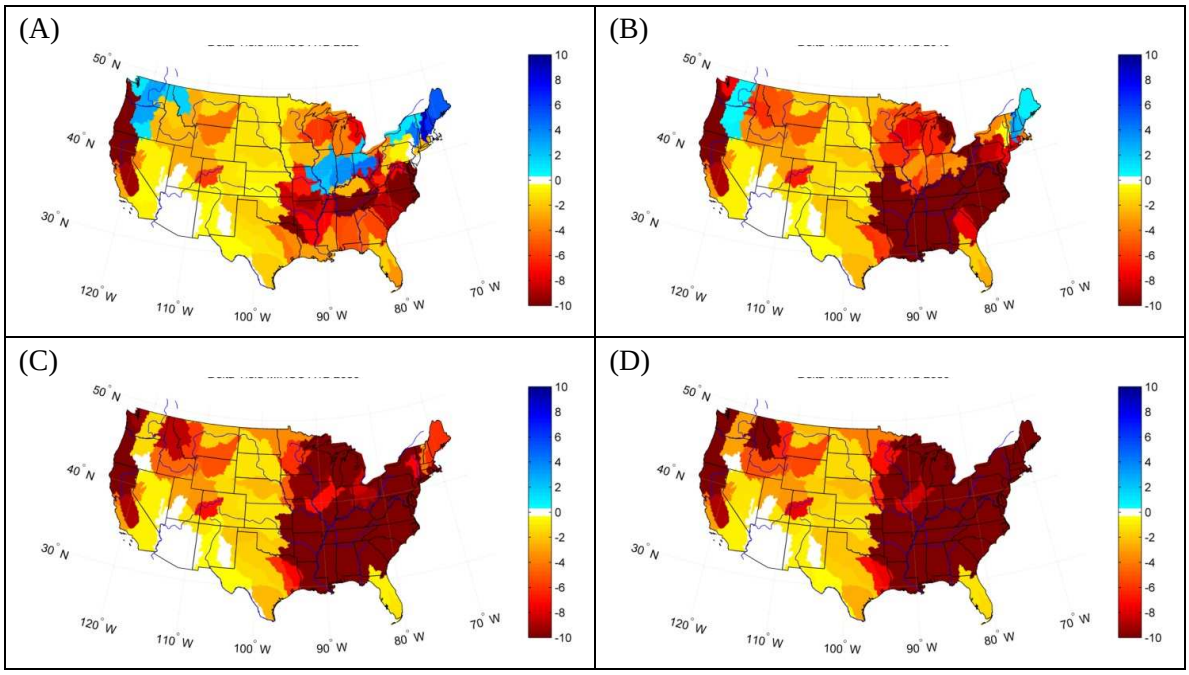


Fig. C.31 Changes in mean water yield for the MIROC/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

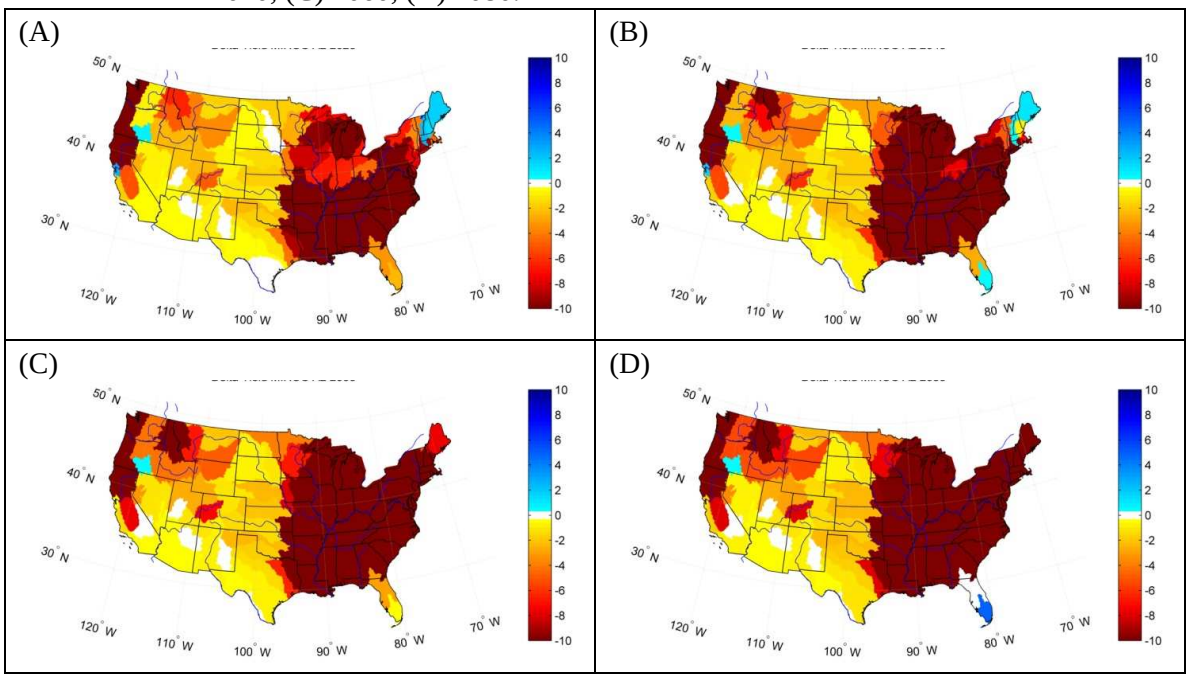


Fig. C.32 Changes in mean water yield for the MIROC/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

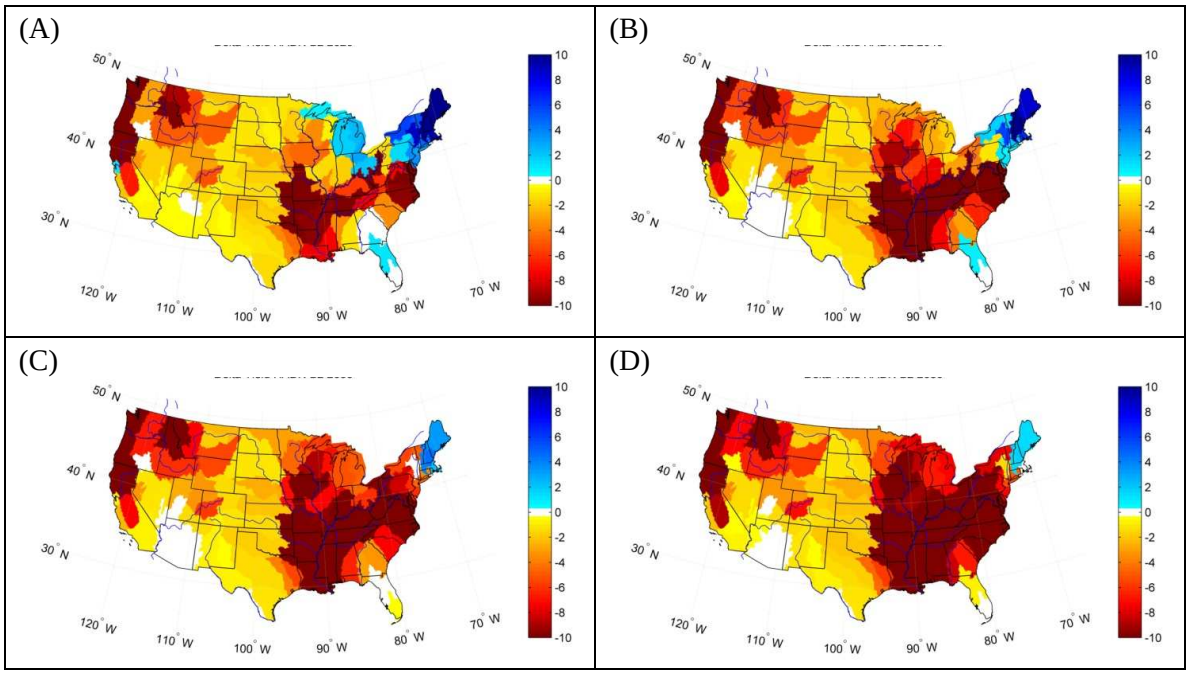


Fig. C.33 Changes in mean water yield for the HADN/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

Tab. C.I Atmospheric CO_2 concentrations and global mean temperature changes of the IPCC scenarios.

Year	A1B-AIM		A2-ASF		B2-MESSAGE	
	CO_2[a]	ΔT[b]	CO_2	ΔT	CO_2	ΔT
1970	325		325		325	
1980	337		337		337	
1990	353		353		353	
2000	369	0.2	369	0.2	369	0.2
2010	391	0.5	390	0.4	388	na
2020	420	0.7	417	0.7	408	na
2030	454	1.0	451	0.9	429	na
2040	491	1.4	490	1.2	453	na
2050	532	1.7	532	1.5	478	na
2060	572	2.0	580	1.9	504	na
2070	611	2.2	635	2.3	531	na
2080	649	2.4	698	2.8	559	na
2090	685	2.6	771	3.2	589	na
2100	717	2.8	856	3.6	621	na

[a] In ppm. Source: http://www.ipcc-data.org/ddc_co2.html, reference model runs.
[b] Multi-model °C change from 1980-1999 mean. Source: http://www.ipcc.ch/pdf/assesment-report/ar4/syr/ar4_syr_spm.pdf. Decadal changes were not listed for the B2 scenario.

Tab. C.II Scenarios of future conditions in the U.S.

	A1B	A2	B2
Population growth	Medium	High	Low
Economic growth	High	Low-Medium	Low
Temperature increase	Medium	High	Low

Tab. C.III Scenario-GCM combinations.

A1B	A2	B2
CGCM31 MR	CGCM31 MR	CGCM31 MR
CSIROMK 35	CSIROMK 35	CSIROMK 35
MIROC32 MR	MIROC32 MR	HADCM3

Appendix D Vulnerability Assessment

D.1 Overview

Vulnerability, defined as the probability of water shortage, was assessed at the ASR level for current and future climatic and socio-economic conditions. An instance of shortage occurs whenever the water supply in a given ASR is insufficient to meet the demand.

Vulnerability for both current and future conditions was evaluated by simulating water allocation within the water networks of the U.S. Individual water network simulations were performed for the nine alternative futures. Each simulation started in 1953 and proceeded to year 2090. Vulnerability was measured for five 20-year time periods within the 1953-2090 time span. Current vulnerability was estimated as the probability of shortage for the period 1986-2005. Future vulnerability was estimated for four 20-year periods centered at 2020, 2040, 2060 and 2080.

This chapter is organized as follows: section D.2 gives a definition of vulnerability as we intend it in this framework; vulnerability assessments for current climatic and socio-economic conditions and projected future climatic and socio-economic conditions are discussed respectively in sections D.3 and D.4.

D.2 Vulnerability: definition and approach

In general, vulnerability of a system is a function of the extent at which it can be damaged by the impact of an external hazard. The definition of vulnerability and the implication that it has on the approach used to estimate it has been the focus of several papers (Blaikie, Cannon et al. 1994),(Kelly and Adger 2000). In its forth assessment, the Intergovernmental Panel on Climate

Change (IPCC) defined vulnerability as “the degree to which these systems are susceptible to, and unable to cope with, adverse impacts” (Schneider, Semenov et al. 2007). While this last definition seems to be widely accepted, it still leaves room for debate concerning the actual quantification of the “degree of susceptibility to adverse impacts”. In the context of water resources, many studies have estimated the potential impact of future climate and socio-economic scenarios on water resources by estimating future projections of a set of water stresses indicators (Postel 2000),(Vörösmarty, Douglas et al. 2005), (Weib and Alcamo 2011).

In agreement with the IPCC definition, we observe that, conceptually, the vulnerability of a system is a function of its ability to respond to (*i.e.*, cope with; adapt to) inherently variable stressors. However, modeling the ability to respond to stresses—via, for example, construction of new reservoirs or alteration of allocation priorities—is a step beyond our goals in this assessment. Rather, we seek to measure the likelihood that adaptation will be needed, and to objectively address the uncertainty about the stressors affecting the system. In particular, we estimate vulnerability as the probability that a critical system threshold, itself a function of both the capacity and the stressors of the system, will be crossed (Korchendorfer and Ramirez 1996). In the context of the US water supply system, this definition translates into evaluating the probability that, at given time and place, water demand exceeds water supply. In other words, we define vulnerability as the probability of shortage:

$$V = \Pr[S < D] = \Pr[S - D < 0] \tag{D.1}$$

where S is water supply, and D is water demand. In general, supply is defined as:

$$S = P - E + I + Q_{div} \tag{D.2}$$

where P is precipitation, E is actual evapotranspiration, I is the input from upstream (which may include reservoir releases), and Q_{div} is the net trans-basin diversion (the difference between diversions into and diversions out of the ASR)[3].

Setting Z equal to S - D, equation (D.1) can be rewritten as:

$$V = \Pr[S < D] = \Pr[S - D < 0] = \Pr[Z < 0] \tag{D.3}$$

or more explicitly:

$$V = \Pr[Z < 0] = \frac{1}{\sqrt{2\pi\sigma_Z^2}} \cdot \int_{-\infty}^{0} e^{-\frac{(z-\mu_Z)^2}{2\sigma_Z^2}} dz \tag{D.4}$$

where:

$$\mu_Z = \mu_S - \mu_D \tag{D.5}$$

$$\sigma_Z^2 = \sigma_S^2 + \sigma_D^2 - 2\,\mathrm{cov}(S, D) \tag{D.6}$$

and μ_S, μ_D, σ_S, and σ_D, *cov(S, D),* are the mean, standard deviation and covariance of water supply and water demand.

Equation (D.4) is the exact expression for vulnerability in the case of *correlated* normally distributed S and D. Or, in the case of non-Gaussian variables, it corresponds to a First Order Second Moment approximation.

Carrying out the integral of equation (D.4) yields,

---

[3] Unlike P, which is an exogenous input each year, S is an endogenous quantity, because it depends on storage and delivery decisions made in response to the priorities determining water allocation within a network. S of an ASR potentially is affected by water stored in that ASR the previous year. And if the ASR is part of a multi-ASR network, S of the ASR potentially also is affected by water previously stored in reservoirs of upstream ASRs. Thus, S can only be obtained as an output of the water routing model. The estimates of S used here follow the specification in equation 6.2, with I to a given downstream ASR in a given year including not only all inflows from upstream (which may include releases from upstream reservoirs) but also releases from reservoirs within the given ASR. Note also that S includes the water that must be used to satisfy the required in-stream flow release from the ASR.

$$V(\mu_S, \mu_D, \sigma_Z) = \frac{1}{2} + \frac{1}{2} erf\left[\frac{-\mu_S + \mu_D}{\sqrt{2\sigma_Z^2}}\right] \tag{D.7}$$

where *erf()* is the Gauss error function[4].

Therefore, as is clear from equation (D.7), the vulnerability of water supply to shortage as defined in equation (D.1) is a function of the mean, standard deviation and covariance of supply and water demand, that is, μ_S, μ_D, σ_S, σ_D, *cov(S, D)*. We may then express the total change in vulnerability, *dV*, as a function of the individual contributions of changes in each of those variables as follows:

$$dV = \frac{\partial V}{\partial \mu_S} d\mu_S + \frac{\partial V}{\partial \mu_D} d\mu_D + \frac{\partial V}{\partial \sigma_S} d\sigma_S + \frac{\partial V}{\partial \sigma_D} d\sigma_D + \frac{\partial V}{\partial \operatorname{cov}(S,D)} d\operatorname{cov}(S,D) \tag{D.8}$$

where each of the partial derivatives represent the sensitivity of the vulnerability to unit changes in each of the independent variables μ_S, μ_D, σ_S, σ_D, and *cov(S, D)*. In addition, each of the five terms of equation (D.8) represents the contribution to the total change in vulnerability resulting from the changes in μ_S, μ_D, σ_S, σ_D, and *cov(S, D)*. The partial derivatives appearing in equation (D.8) are obtained differentiating equation (D.7) with respect to μ_S, μ_D, σ_S, σ_D, and *cov(S, D)*. The resulting derivatives are shown below in equations (D.9)-(D.13).

$$\frac{\partial V}{\partial \mu_S} = -\frac{1}{\sqrt{2\pi\sigma_Z^2}} \cdot e^{-\frac{(-\mu_S + \mu_D)^2}{2\sigma_Z^2}} \tag{D.9}$$

$$\frac{\partial V}{\partial \mu_D} = \frac{1}{\sqrt{2\pi\sigma_Z^2}} \cdot e^{-\frac{(-\mu_S + \mu_D)^2}{2\sigma_Z^2}} \tag{D.10}$$

$$\frac{\partial V}{\partial \sigma_S} = -\frac{\sigma_S \cdot (-\mu_S + \mu_D)}{\sqrt{2\pi(\sigma_Z^2)^3}} \cdot e^{-\frac{(-\mu_S + \mu_D)^2}{2\sigma_Z^2}} \tag{D.11}$$

[4] The Gauss error function is also known as the probability integral.

$$\frac{\partial V}{\partial \sigma_D} = -\frac{\sigma_D \cdot (-\mu_S + \mu_D)}{\sqrt{2\pi(\sigma_Z^2)^3}} \cdot e^{-\frac{(-\mu_S+\mu_D)^2}{2\sigma_Z^2}} \tag{D.12}$$

$$\frac{\partial V}{\partial \operatorname{cov}(S,D)} = \frac{(-\mu_S + \mu_D)}{\sqrt{2\pi(\sigma_Z^2)^3}} \cdot e^{-\frac{(-\mu_S+\mu_D)^2}{2\sigma_Z^2}} \tag{D.13}$$

Although not included here for brevity, the above analysis can be easily extended to define changes in vulnerability as a function of changes in the probabilistic characteristics of *P*, *E*, *I*, and Q_{div}, explicitly[5].

D.3 Current Vulnerability of the U.S. water supply to shortage

The climate of the period 1986-2005 is taken to represent the *current climate*. Distribution functions of precipitation, temperature and potential evapotranspiration for the period 1986-2005 thus were used as input to the water balance model to determine the PDF of water yield for the current climate.

Water demands, as mentioned above, are also characterized by a stochastic component. The annual consumptive use of a given ASR is dependent in part on the amount of precipitation as well as on temperature and potential evapotranspiration. PDFs of water supply and water demand, therefore, are strongly correlated, because they both derive from the climatic input (precipitation, temperature and potential evapotranspiration).

[5] Initially we implemented an alternative approach to measuring vulnerability, which involved creating alternative synthetic traces water yield and demand. We used a multivariate AR(1) model to generate the synthetic traces based on the statistical properties of the original estimates of precipitation, temperature, potential evapotranspiration, and related water yield. Simulations using those traces provided alternative, statistically-identical versions of past and future supply and demand. Combining the supply and demand results of four synthetic traces with those of the original trace provided a total of 100 years of results for each 20-year period of interest. This relatively cumbersome approach provided estimates of vulnerability very similar to those obtained using the approach described here.

The historical values for climate variables (precipitation, temperatures and potential evapotranspiration) and for water use allow us to compute vulnerability estimates for what we call *current conditions*.

The procedure of water network simulation to estimate current vulnerability was as follows:

- The ASR network was simulated for a total of 53 years, from 1953 to 2005.
- As initial condition, storages in 1953 were set as half full.
- A flat value of water demand, corresponding to the demand in 1985, was used for those 33 years.
- The first 33 years of simulation were considered as transient and discarded for any vulnerability estimation purposes;

The analysis of the period 1986-2005 is here carried out not only to provide an estimate of the current probability of shortage, but mostly to set the benchmark to which future projections can be compared.

Estimates of μ_S, μ_D, σ_S, σ_D, and *cov(S, D)* for the current period for each ASR were computed from the annual values of *S* and *D* produced by the network simulations. For example, μ_S for a given ASR is the mean of the 20 years (1986-2005) of S_i from the multi-year simulation.

Recalling that water surplus, Z, was defined as the difference between water supply and water demand, vulnerability is the probability that the water surplus is zero or negative. By simply looking at the first moments of the water surplus PDF, one can notice that vulnerability increases as the mean of the surplus μ_z decreases and as its variance σ^2_z increases. Taking into account both effects simultaneously, one may also quantify vulnerability as a function of the ratio of the mean surplus to the corresponding standard deviation, $\beta = \mu_z/\sigma_z$, referred to hereafter as the reliability index. The reliability index quantifies in units of the standard deviation how far from shortage a given ASR is.

Maps of current vulnerability and reliability ratios are presented in Fig. D.1 and show that the water supply system for much of the US west of the Mississippi river is vulnerable under current

hydro-climatic and socio-economic conditions. However, only a few areas show vulnerability values exceeding 0.05 at the ASR scale, and they tend to be those that rely heavily on mining of groundwater.

D.4 Future vulnerability of the U.S. water supply to shortage

Future vulnerability is evaluated for each of the nine GCM-scenario combinations for the target 20-year periods centered around 2020, 2040, 2060 and 2080.

D.4.1 Sensitivity of vulnerability to changes in the drivers

Understanding how a given location responds to potential changes in climatic and socio-economic conditions is essential for future water management planning. As explained above, future changes in vulnerability of water supply to shortage are a function not only of the magnitude of the changes in future supply and in future demands, but also of the sensitivity of vulnerability to unit changes in supply and demand. In turn, those sensitivities are functions of the means, variances and covariances of P, E, and D.

The sensitivity of vulnerability to changes in P, that is, the change in vulnerability per unit change in mean precipitation, μ_P, for the CGCM/A1B scenario (Fig. D.2a) is larger in the western US in a region coinciding with the region exhibiting reliability the lower reliability index (Fig. D.2b). A similar behavior occurs with respect to changes in mean demand, μ_D, but with opposite sign (Fig. D.2). Therefore, these areas, in addition to being quite vulnerable under the current conditions, are more prone to large increases in vulnerability for the same change in S, E and D. In other words, these areas are vulnerable because their mean surplus, μ_Z, is close to zero, and because they are more sensitive to unit changes in mean surplus. As discussed above, it is not only changes in the mean of the drivers that affect vulnerability but also changes in their variances and co-variances. For the CGCM/A1B scenario, as the variance of S increases so does the vulnerability of water supply to shortage over most of the US (Fig. D.2c). Similar behavior is

observed with respect to changes in the variance of D (Fig. D.2d). However, as shown above, for the arid southwestern US, where S is close to zero, increases in the variance actually lead to decreases in the vulnerability. This is because a larger variance implies that, although large surpluses are still rare, their likelihood increases, therefore leading to reductions in vulnerability.

D.4.2 Vulnerability under the CGCM/A1B scenario

Changes in future vulnerability reflect changes in the probability distribution functions of S and D. Fig. D.5 shows the changes in vulnerability predicted for the CGCM31/A1B scenario for the periods 2020, 2040, 2060, and 2080. It is noticeable how the southwestern US and the Great Plains areas are expected to face the greatest increases in vulnerability, in addition to being already the areas where shortages are more likely. Dramatic increases in vulnerability are expected throughout the entire 21st century for the lower CRB, the central Great Plains, and the central California. Large increases are also expected in the Rio Grande basin, Texas and Utah. Noticeably, all those areas interested by the larger increase in vulnerability are characterized by having a current reliability index less than two (Fig. D.1). Interestingly, vulnerability seems to monotonically increase overtime in most of the country, especially in those areas where the likelihood of shortage is already large. Decreases in vulnerability, on the other hand, are quite small and are only expected in few areas of the eastern US, Midwest and northern US.

It is interesting to observe the separate effects of changes in S and changes in D on resulting changes in vulnerability. The effects of changes in S or D are each a combination of changes in the mean and the standard deviation. The combined effects of changes in the mean and the standard deviation of S for the A1B-CGCM future are shown in Fig. D.3, for changes from the current period to the four future periods centered at years 2020, 2040, 2060 and 2080. The effects of changes in the mean and standard deviation are represented by the first and third term of equation 6.6, respectively. Similarly for demand, Fig. D.4 shows the combined effects (i.e., the sum of the second and fourth terms of equation (D.8)) of changes in the mean and in the standard

deviation of D for A1B-CGCM. Changes in mean and variance of S lead by 2060 to increases in vulnerability over the central and southern Great Plains and in central California, and to decreases in the Sacramento basin, the Sierras to the east of the San Joaquin Valley, Nevada, parts of Arizona, parts of Washington and Oregon, the northern Great Plains, the Midwest, and the Mid-Atlantic region (Fig. D.3). Changes in mean and variance of D are projected by 2040 to induce increases in vulnerability in the Southwest, central California, and the southern Great Plains, and to have little effect on vulnerability over the rest of the country (Fig. D.4). The increases in D in many areas of the East, although substantial, especially in the Southeast (see Tom et al., in preparation), are insufficient to result in much shortage at the ASR scale.

The cumulative effect of changes in mean S and in the standard deviation of S is expected to increase vulnerability for the central Great Plains throughout the entire 21st century. The Rio Grande Basin and the Colorado River Basin, on the other hand, are expected to undergo alternate periods of increasing and decreasing vulnerability due to changes the PDF of S. A similar behavior is detected in the central California, where changes in the PDF of S are expected to determine an increase in vulnerability for the periods 2020 and 2060 and a decrease in 2040 and 2080. As for the rest of the country, the contribution of changes in the moments of S is projected to have negligible effect or to lead to decreases in vulnerability. Although not shown, the latter situation is primarily due to smaller variance in water supply projected through the course of the 21st century for the northwest, the northern Great Plains, the northern California and Texas.

Unlike the case of the water supply, the effect of changes in both mean and standard deviation of water demand is projected to always increase future vulnerability, with the noticeable exception of the Sevier Lake and the Rio Grande basin in 2020. The effect of water demand change is larger in the Southwest, in central California and in the southern Great Plains, while being essentially negligible in the rest of the country.

Finally, changes in vulnerability resulting from projected changes in the S and D are of the same magnitude or larger than those from changes in the corresponding means over most of the

US, except in central coastal California, in the San Joaquin river basin and in southern Florida where the opposite is true.

D.4.3 Vulnerability by GCM and Scenario

The A1B-CGCM future represents only one possible future set of hydro-climatic and socio-economic conditions. Analyzing alternative scenarios and utilizing alternative CGMs is one way to characterize the uncertainty that exists about the vulnerability projections.

Fig. D.6 through Fig. D.13 present estimates for future vulnerability projected for the other eight alternative futures. Those maps show pictures of the future broadly similar to that of the A1B-CGCM future. In particular, consistent increases in probability of shortage are projected for all nine futures predominantly in the Great Plains area and in the Southwest. Nevertheless, the magnitude of those increases can vary considerably among the alternative futures, as does the areal extent of the most vulnerable areas.

The CGCM model generally projects the less dramatic increases in vulnerability. The CSIRO model projects the largest increases in vulnerability in the Eastern U.S., as shown by comparing its projections with those of the other GCMs for corresponding scenarios. The MIROC model (as well as the HADN for the B2 scenario), on the other hand, projects the largest increases in vulnerability in the Great Plains, in the southern central U.S., and in the Colorado River Basin.

Among scenarios, the A2 is the one for which all the GCMs project the largest increases in vulnerability. Notable exceptions are the Colorado River Basin for the CGCM and the central California for the CSIRO, where the highest vulnerability are projected by the B2 scenario.

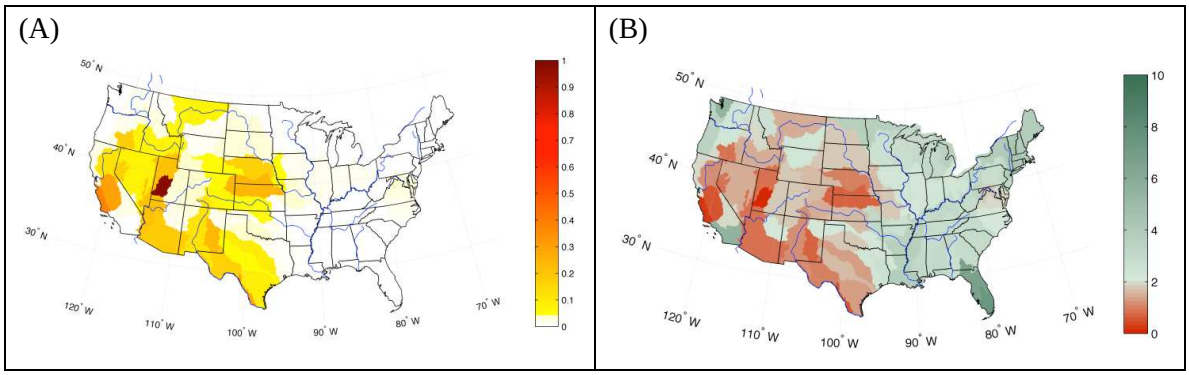


Fig. D.1 Current vulnerability (A) and reliability index (B).

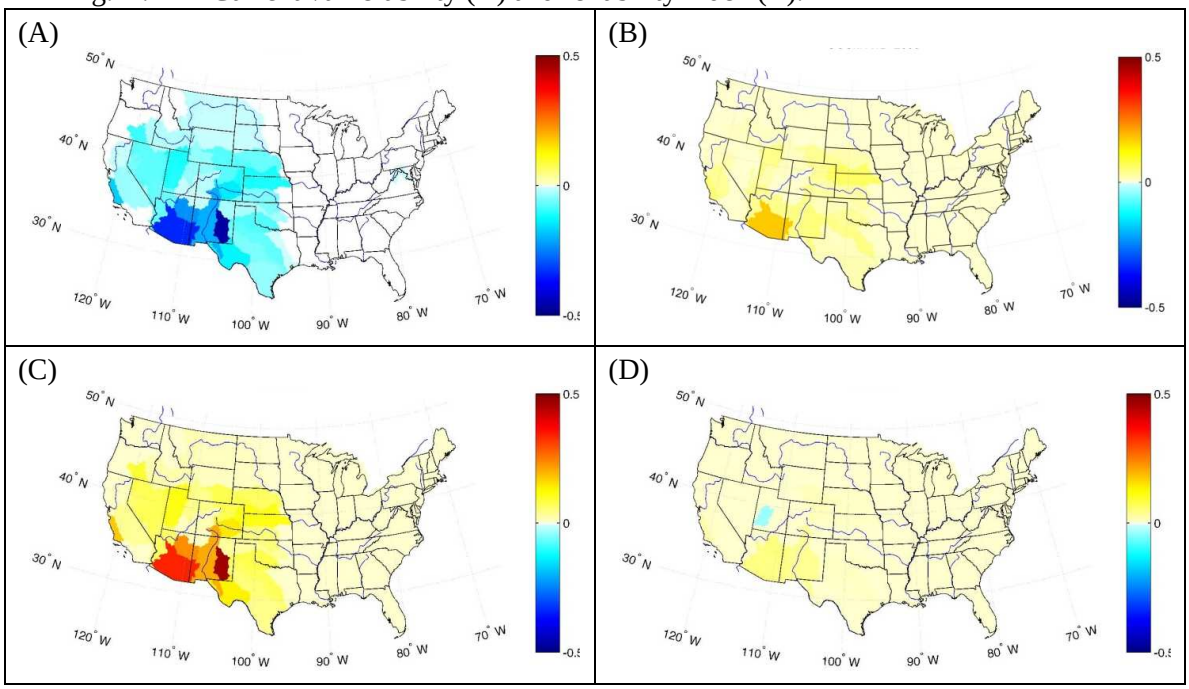


Fig. D.2 Current sensitivity of vulnerability to unit changes in: (A) mean water supply; (B) mean water demand; (C) standard deviation of water supply; (D) standard deviation of water demand; for the CGCM/A1B scenario.

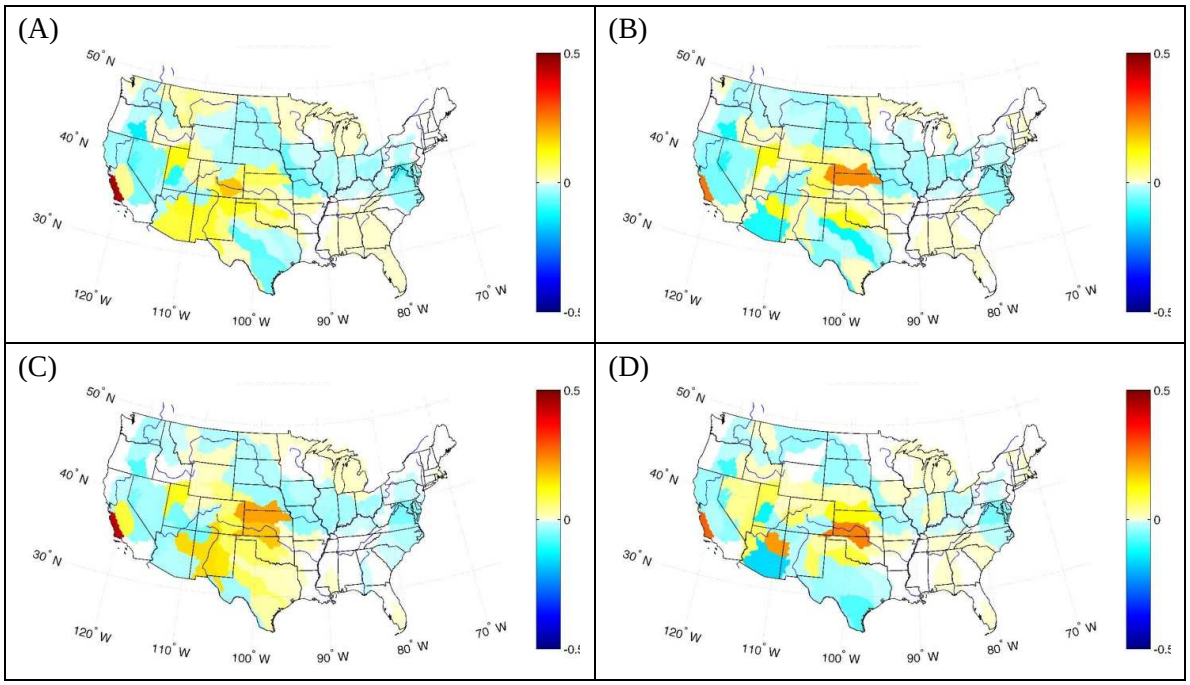


Fig. D.3 Changes in vulnerability for the CGCM/A1B scenario due to changes in average water supply and standard deviation of water supply for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

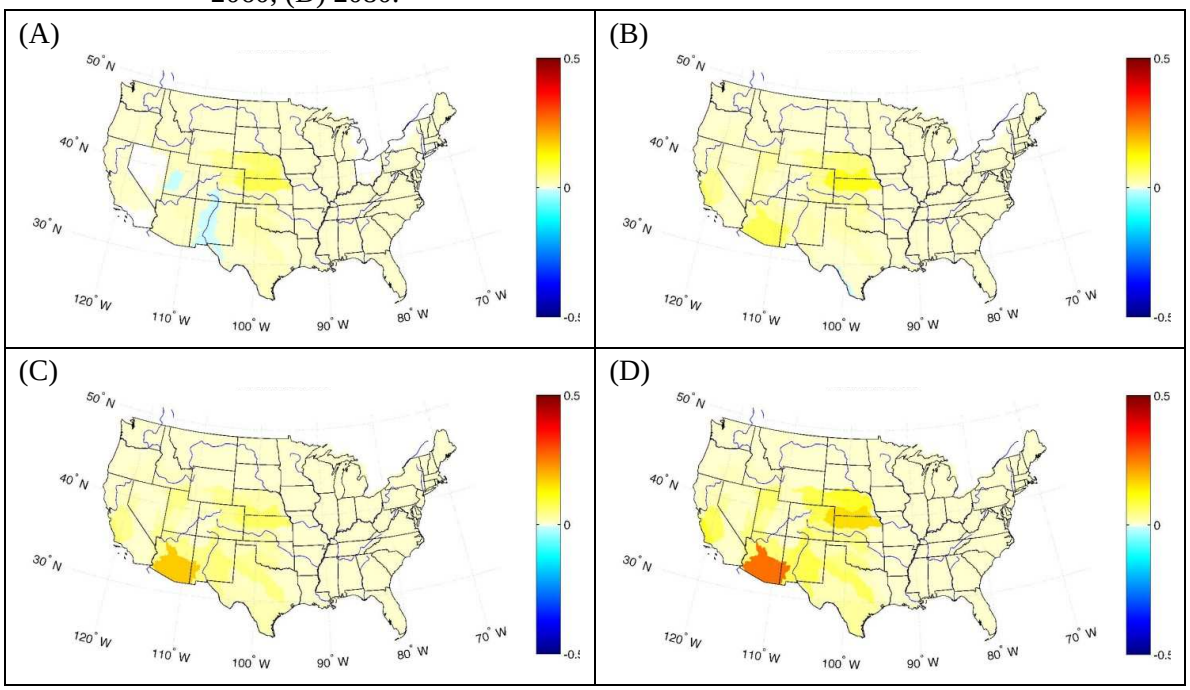


Fig. D.4 Changes in vulnerability for the CGCM/A1B scenario due to changes in average water demand and standard deviation of water demand for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

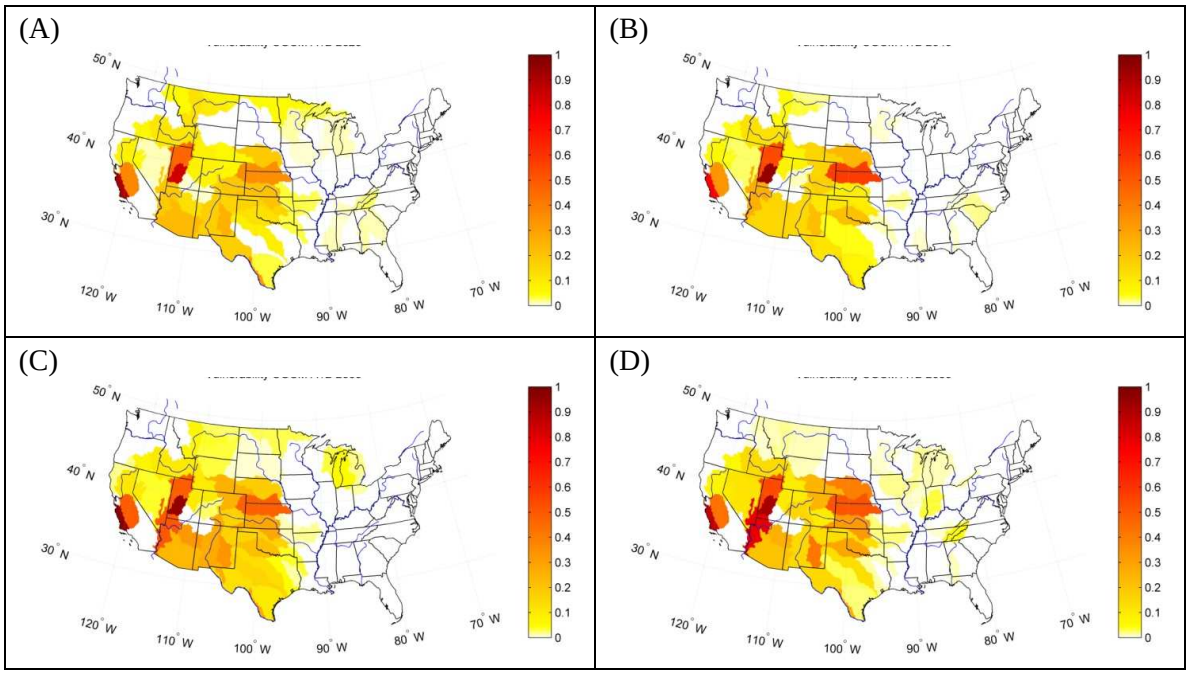


Fig. D.5 Future vulnerability for the CGCM/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

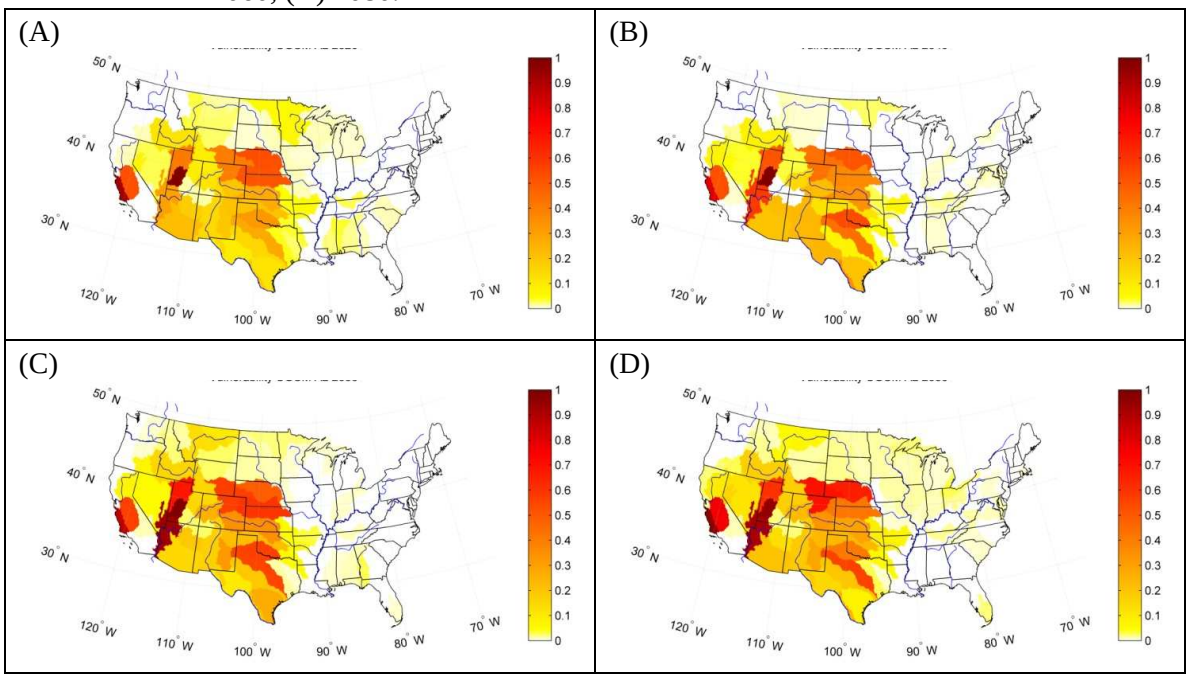


Fig. D.6 Future vulnerability for the CGCM/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

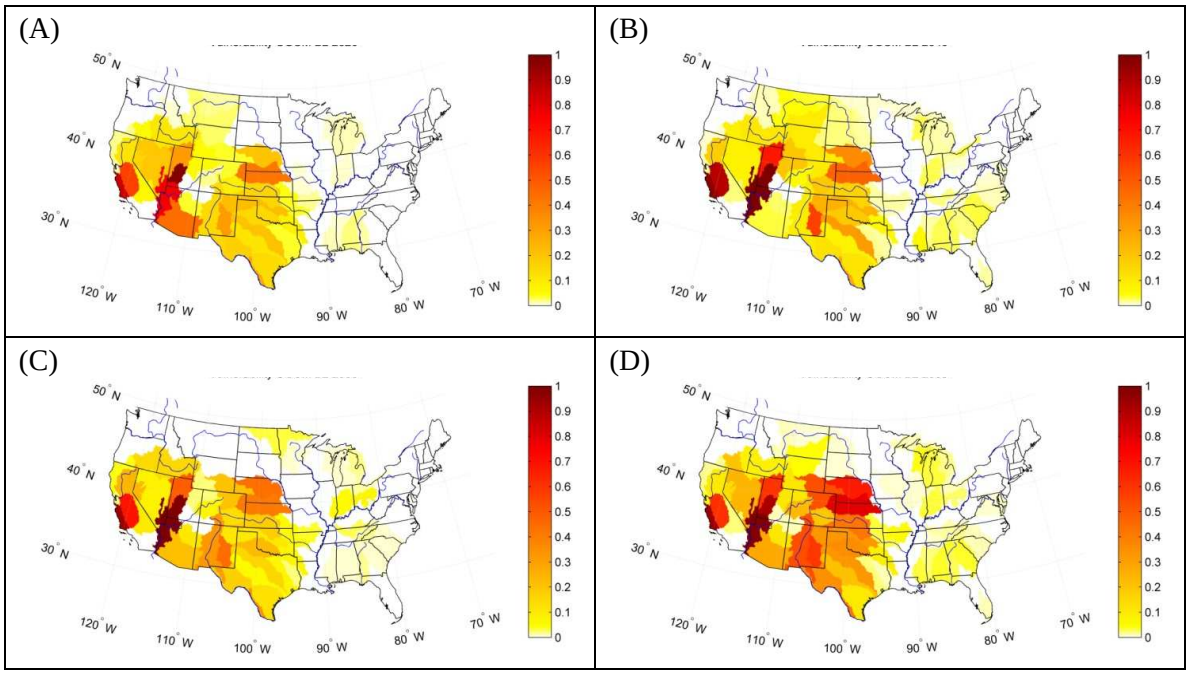


Fig. D.7 Future vulnerability for the CGCM/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

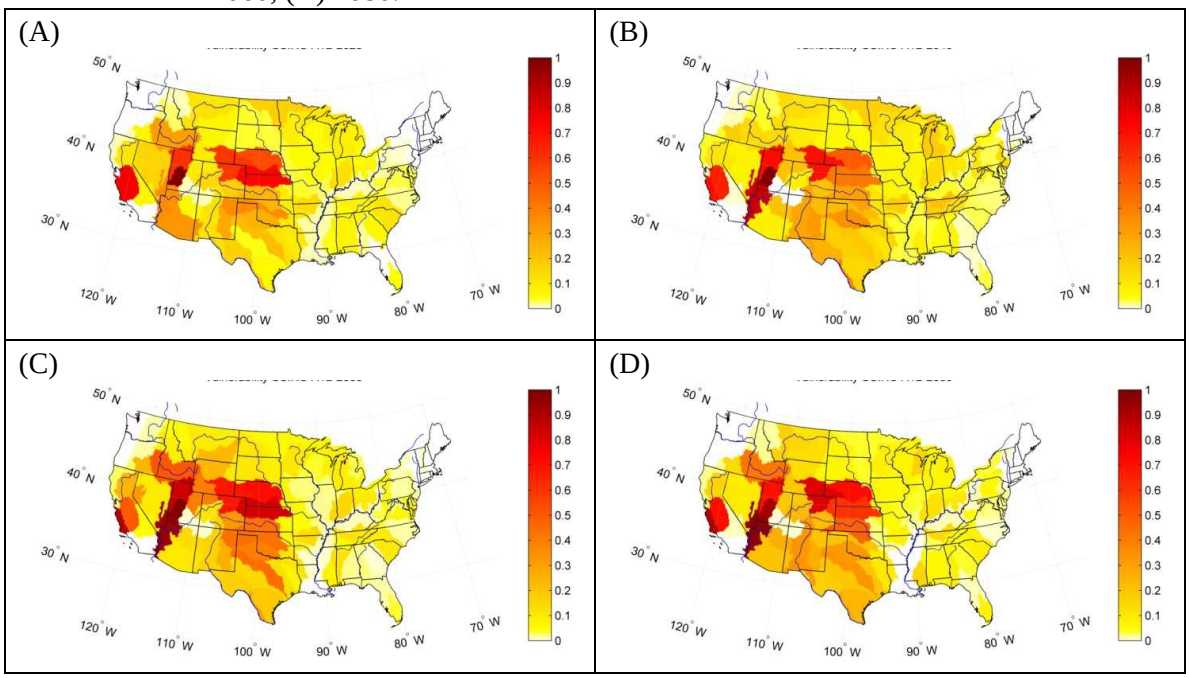


Fig. D.8 Future vulnerability for the CSIRO/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

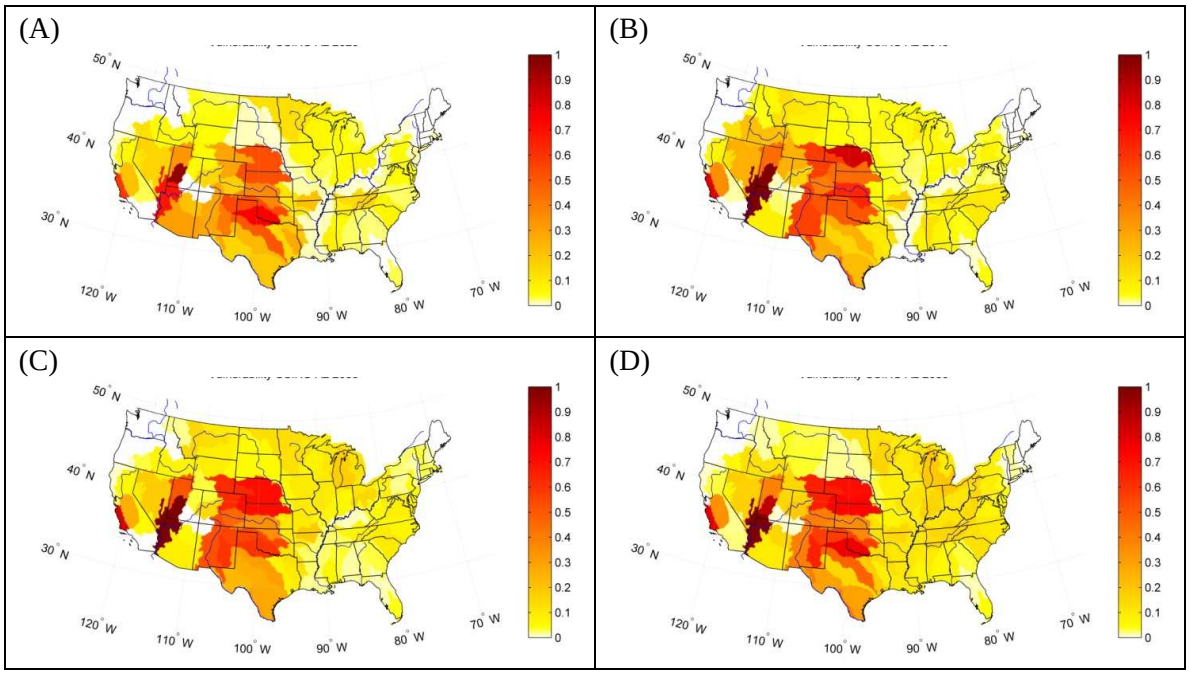


Fig. D.9 Future vulnerability for the CSIRO/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

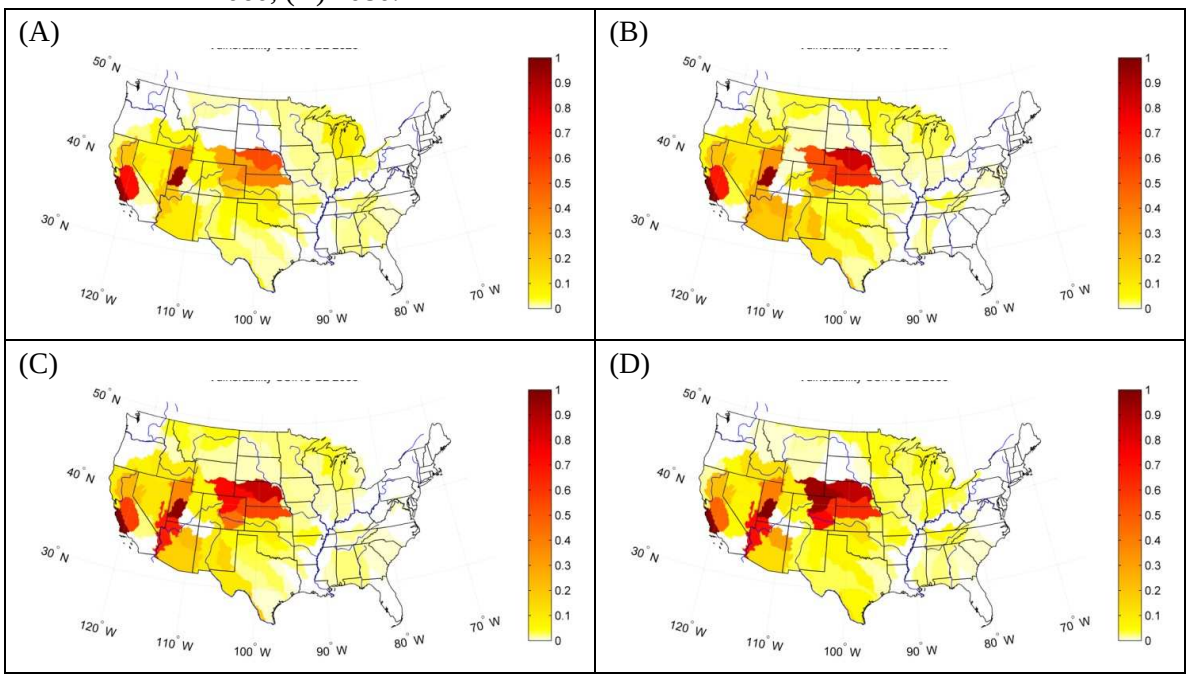


Fig. D.10 Future vulnerability for the CSIRO/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

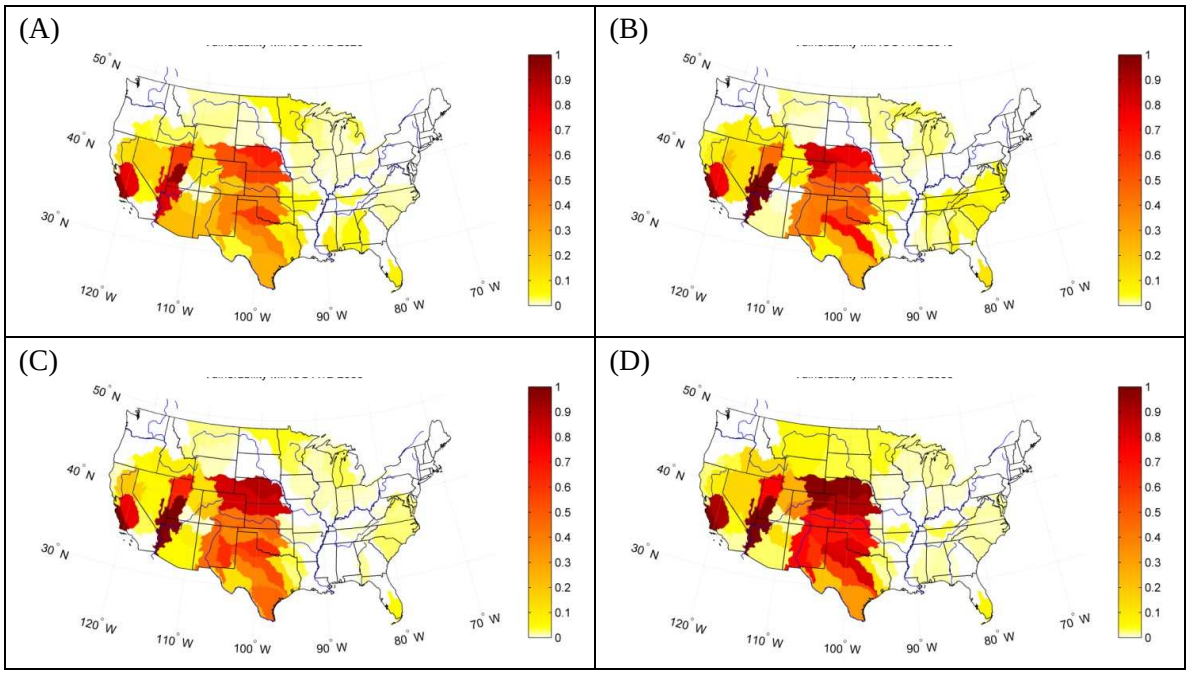


Fig. D.11 Future vulnerability for the MIROC/A1B scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

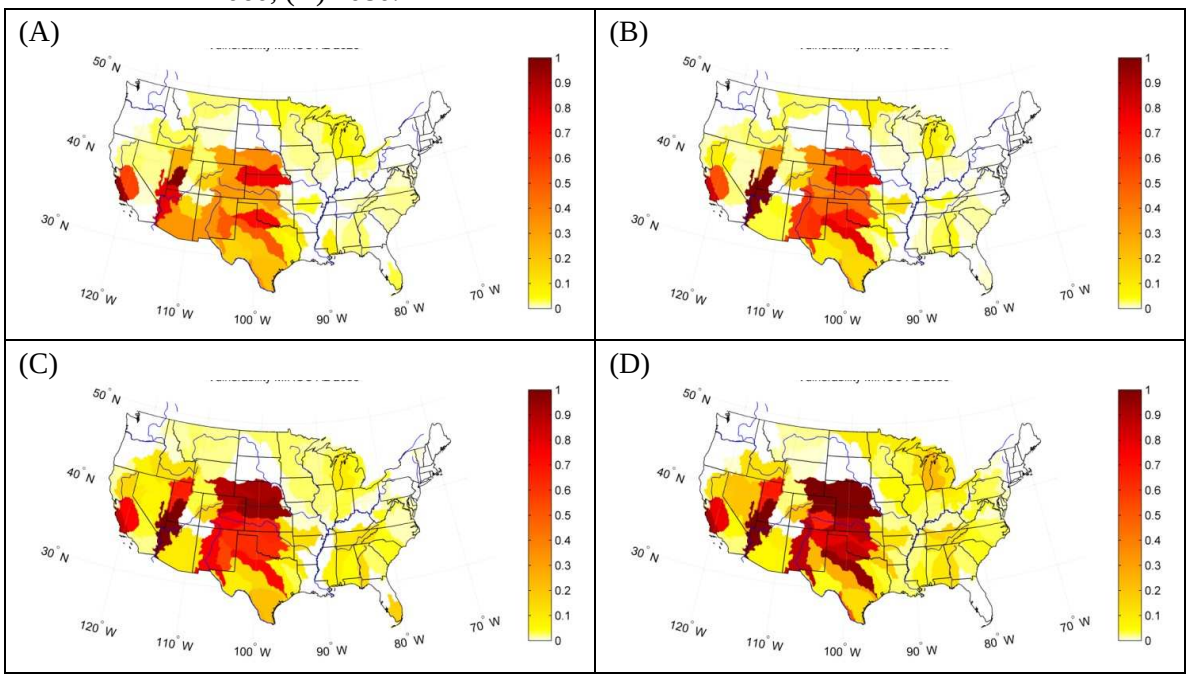


Fig. D.12 Future vulnerability for the MIROC/A2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

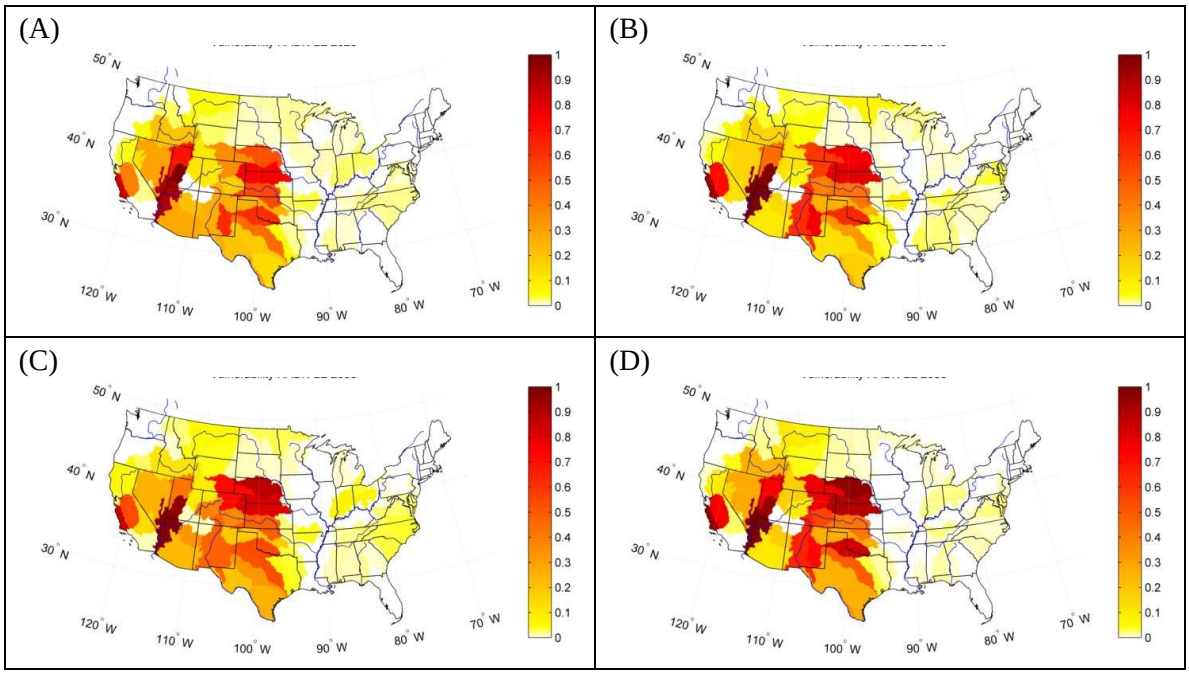


Fig. D.13 Future vulnerability for the HADN/B2 scenario for: (A) 2020; (B) 2040; (C) 2060; (D) 2080.

Appendix E A Probabilistic Framework for Assessing Vulnerability to Climate Variability and Change: The Case of the US Water Supply System

E.1 Abstract

We introduce a probabilistic framework for vulnerability analysis and use it to quantify current and future vulnerability of the US water supply system. We also determine the contributions of hydro-climatic and socio-economic drivers to the changes in projected vulnerability. For all scenarios and global climate models examined, the US Southwest including California in the west and the southern Great Plains to the east was consistently found to be the most vulnerable. For most of the US, the largest contributions to changes in vulnerability come from changes in evapotranspiration, followed by those from changes in precipitation. Projected increases in demand have comparatively minor effect on changes in vulnerability. Changes in vulnerability from projected changes in the variances of precipitation, evapotranspiration and demand are of the same magnitude or larger than those from changes in the corresponding means over most of the US, except in central California and in southern Florida.

E.2 Vulnerability

The vulnerability of a system is a function of its ability to respond to (*i.e.*, cope with, adapt to) inherently variable stressors. Because the magnitude of the stresses and the capacity to withstand them are uncertain, vulnerability should be quantified probabilistically and depends on the joint probability distribution function of capacity and stresses. We define vulnerability as the probability that a critical system threshold, itself a function of both the capacity and the stressors of the system, will be crossed (Korchendorfer and Ramirez 1996).

For a water supply system, this probabilistic character implies that vulnerability depends on the mean, variance, and co-variance of water supply and water demand. More importantly, it implies that to address questions about future changes in vulnerability, it is not sufficient to quantify the effects of changes in the mean values of hydro-climatic and socio-economic variables of interest—it is necessary also to quantify the effects of changes in the inherent variability of those variables (see Appendix D).

We quantify the vulnerability of a water supply system as the probability that water demand (D) will exceed water supply (S). Equivalently, defining water surplus, Z, as $S–D$, vulnerability is the probability that Z is negative. For a given variance of surplus, vulnerability increases as the mean surplus μ_z decreases; and for a given mean surplus, vulnerability increases as surplus variance, σ^2_z, increases if the mean surplus is positive, or it decreases otherwise. The dependence of vulnerability on the first two moments of the probability distribution of surplus may also be captured in terms of a reliability ratio, the ratio of the mean surplus to the corresponding standard deviation, $\beta = \mu_z/\sigma_z$. The reliability ratio quantifies in units of the standard deviation how far from the critical threshold of zero the surplus of a given location is, with locations of smaller reliability ratios being more vulnerable.

E.3 Water Supply and Water Demand

Our approach to estimating water supply and demand begins with analysis of historical records and with preparation of two alternative scenarios of socioeconomic changes and greenhouse gas emissions corresponding to the IPCC's SRES A1B and A2 scenarios (Nakicenovic, Alcamo et al. 2000; Zarnoch, Cordell et al. 2010). Future hydro-climatic variables for each scenario were obtained from downscaled projections of three global climate models, CGCM, CSIRO, and MIROC (Price, McKenney et al. 2006). Water supply was estimated as water yield, Y (the difference between precipitation, P, and actual evapotranspiration, E) plus inflows subjected to management via reservoir storage and diversion. Water yield was estimated

for 5x5 *km* cells with a detailed physically based, statistical-dynamical hydrologic model driven by hydro-climatic variables (see Appendix A). Water demand was estimated as the net amount of water depletion that would occur if water supply were no more limiting than it has been in the recent past, with water depletion (withdrawal minus return flow) projected for each of six categories of water use as a product of a water use driver (*e.g.*, population, irrigated area) and a water use efficiency factor (*e.g.*, domestic withdrawals per capita, irrigation water depth) while taking into account expected trends in water use efficiency as affected by climate change (see Appendix B). Because comprehensive data on available deep aquifer storage were lacking, groundwater mining was not included in water supply.

We estimate the vulnerability of the water supply system over the 21st century for 98 basins, called Assessment Sub-Regions (ASRs) (U.S. Water Resources Council 1978), which make up the contiguous 48 states of the US. The natural water flows and existing water diversions between ASRs result in 18 water networks, three multi-ASR networks (of 69, 10, and four ASRs) and 15 single-ASR networks that drain to the ocean or into Canada, or are closed basins. We use a hydrologic network model to route water and simulate water management in each network (Labadie, Pineda et al. 1984). The simulations provide annual values of water flows between ASRs, reservoir storage and evaporation levels in each ASR, and water assigned to each demand, all of which depend both on climate and the following priorities: (1) in-stream flow requirements, (2) trans-ASR diversions, (3) consumptive water uses, and (4) reservoir storage. These priorities recognize the importance of guaranteeing a minimal amount of water for environmental and ecosystem needs before water is diverted for other uses, and allow trans-ASR diversions to occur before within-ASR diversions[6]. Because storage is assigned the lowest priority, water is stored in

[6] With this set of priorities we ignore detailed operating rules that could condition some trans-ASR diversions, perhaps allowing some within-ASR demands to be met before the trans-ASR diversion is fully satisfied. For example, our priorities satisfy the full diversion from the Sacramento Basin to southern California via the State Water Project before local diversions in the Sacramento Basin are met.

a given year only after all the demands reachable by a reservoir are satisfied. The probability distribution functions of *S*, *D*, and *Z* and of their corresponding hydro-climatic drivers are estimated from the output of the hydrologic network model simulations over the period 1953 to 2090 (see Appendix D). We assess current vulnerability over the period 1986-2005 and future vulnerability over the four 20-year periods centered at 2020, 2040, 2060 and 2080.

E.4 Current Vulnerability

The water supply system for much of the US west of the Mississippi river is vulnerable under current hydro-climatic and socio-economic conditions (Fig. E.1A). However, only a few areas show vulnerability values exceeding 0.05 at the ASR scale[7], and they tend to be those that rely heavily on mining of groundwater. The current low levels of vulnerability gradually increase in both degree (probability of shortage) and areal extent, especially in the southwestern US, as seen for the CGCM/A1B hydro-climatic and socio-economic projection by comparing Fig. E.1A with Fig. E.2C and Fig. E.2D.

E.5 Sensitivity of Vulnerability to Climate Change

The future change in vulnerability depends on the sensitivity of vulnerability to unit changes in the mean and variance of *P*, *E*, and *D* and on the future changes in *P*, *E*, and *D*. With respect to the means, vulnerability decreases as mean *P* (and supply) increases (Fig. E.2A), and increases as mean *E* (and demand) increases (Fig. E.2B). The magnitude of this sensitivity tends to be larger in those areas of the western US exhibiting reliability ratios less than two (Fig. E.1B). These areas, in addition to being vulnerable under the current conditions, are more prone to large changes in vulnerability than other areas for the same change in *P*, *E*, *S* or *D*. In other words,

[7] Modeling at the ASR scale will fail to detect some local shortages, as in the case with Atlanta, which, because it is located in the upper reaches of its ASR, is upstream of most of the water available within the ASR itself.

these areas are vulnerable because their mean surplus is close to the critical threshold of zero, and because they are more sensitive to unit changes in mean surplus.

As the variances of precipitation and evapotranspiration increase so does the vulnerability of water supply over most of the US (Fig. E.2B, C and D). However, where mean surplus is negative, increases in the variance of *P* or *E* lead to decreases in vulnerability. Currently this occurs only in the Sevier Lake basin in Utah. However, the condition is projected to occur in the future under some GCM/scenario combinations in the most arid parts of the Southwest. This negative sensitivity of vulnerability in the arid Southwest to unit changes in variance of *P* and *E* is also true with respect to changes in the variance of *S* and *D* (see Appendix D).

E.6 Future Vulnerability

Decreases in yield from the current period to 2060 are projected for most of the US except in parts of the Southwest where yield is projected to increase (Fig. E.3C). Although both precipitation (Fig. E.3A) and actual evaporation (Fig. E.3B) are projected to decrease in the Southwest, *E* is projected to decrease more than *P* in some basins leading to a small increase in yield. For the rest of the country, both *P* and *E* are projected to increase but *E* increases more, leading to a decrease in yield. As a result, supply is projected to decrease for most of the US, except in isolated basins of the Southwest and the southern Plains where it is projected to increase (Fig. E.3E). Water demand is projected to increase generally over the entire US, but especially in the lower Mississippi river basin largely as a result of growth in irrigated agriculture (Fig. E.3D).

Vulnerability is projected to change as a result of changes in the probability distribution functions of supply and demand. Changes in mean *D* from the current period to 2060 are projected to induce increases in vulnerability in the Southwest, central California, and the southern Great Plains, and to have essentially no effect over the rest of the country (Fig. E.3G). Changes in the variance of *D* are projected to have negligible effects on vulnerability. The combined vulnerability changes resulting from changes in mean and variance of *P*, *E*, and thus *S*

lead to increases in vulnerability (Fig. E.3F) over the central and southern Great Plains and in central California, and to decreases in the Sacramento basin, Nevada, and Arizona. Although not shown (see Appendix D), vulnerability is projected to decrease as a result of decreases in the variance of *S* for most of the US except in the southern Great Plains (Eastern Colorado, Kansas, Arkansas and Texas), and in New Mexico and Arizona. Finally, changes in vulnerability resulting from projected changes in the variances of *P*, *E* and *D* are of the same magnitude or larger than those from changes in the corresponding means over most of the US, except notably in central coastal California and the San Joaquin river basin and in southern Florida where the opposite is true (Fig. E.3H).

The above assessment is based on the projections of the CGCM/A1B combination, and thus fails to account for any uncertainty about the projections. Composite maps of the maximum (Fig. E.4A) and minimum (Fig. E.4B) values of vulnerability from among the projections by the MIROC, CSIRO, and CGCM models under scenarios A1B and A2 show that although there is general agreement that the water supply system of the southwestern US is the most vulnerable to hydro-climatic variability and socio-economic changes, there is also a great deal of disagreement about the degree of changes in vulnerability. The disagreement is greatest in the central and southern Great Plains, the Rio Grande basin, the lower Colorado River basin, the San Joaquin river basin in California, and southern Idaho. (See Appendix D for further detail on how the projections differ among the scenario/model combinations.)

E.7 Concluding Remarks

These results assume no modifications to the physical structure of US water networks. In addition, in-stream flow requirements and trans-ASR diversions were set constant, thereby ignoring possible future changes in surface water redistribution. Indeed, it is the purpose of this assessment to point to those locations where adaptation (e.g., enlarged trans-basin diversion capacity or within-basin water transfers and enhanced water conservation) will be most needed.

Because all simulations project a steady decline in the water levels of the main Southwestern reservoirs, indicating that water scarcity is primarily due to demand-supply imbalance rather than to insufficient storage capacity, increasing storage capacity within existing ASR networks does not appear to be a successful adaptation strategy.

Contrary to a prior global scale conclusion (Vörösmarty, Green et al. 2000) and in concert with a recent US study (Roy, L. et al. 2010), we find that future increases in the vulnerability of the US water supply will depend more on changes in water yield than on growth in water demand. Total water use in the US has leveled off in recent years (Kenny, Barber et al. 2009) as irrigated area in the West has diminished and the efficiency of water withdrawals in nearly all sectors has improved (Brown 2000). Although climate change will increase water demand, future water use efficiency improvements will mitigate that impact so that overall increases in desired water use are expected to be modest in comparison with the effect of climatic changes on water yield and thus on water supply.

Our finding of greater vulnerability in the Southwest and central and southern Great Plains of the US is in keeping with other large-scale assessments (Hurd, Leary et al. 1999; Vörösmarty, Green et al. 2000; Kenny, Barber et al. 2009). Importantly, our work adds an accounting for reservoir storage, trans-basin diversions and routing of water among basins; a more comprehensive effort to project future desired water use; and a probabilistic approach to vulnerability. This probabilistic methodology can be applied to any vulnerability analysis, and is the only methodology that both accounts for the probabilistic character of the drivers and allows for explicit inclusion of thresholds.

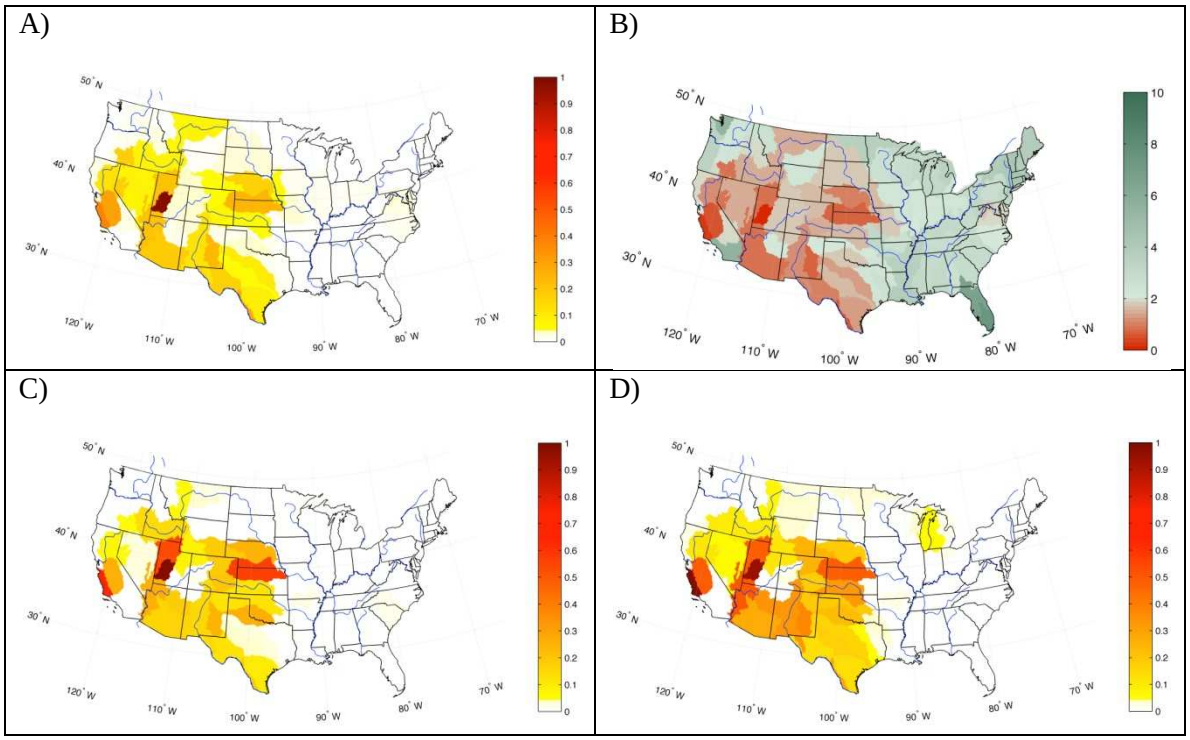


Fig. E.1 Current vulnerability (A) and reliability ratio (B), and future vulnerability projected for 2040 (C) and 2060 (D) by the CGCM model under the A1B SRES scenario.

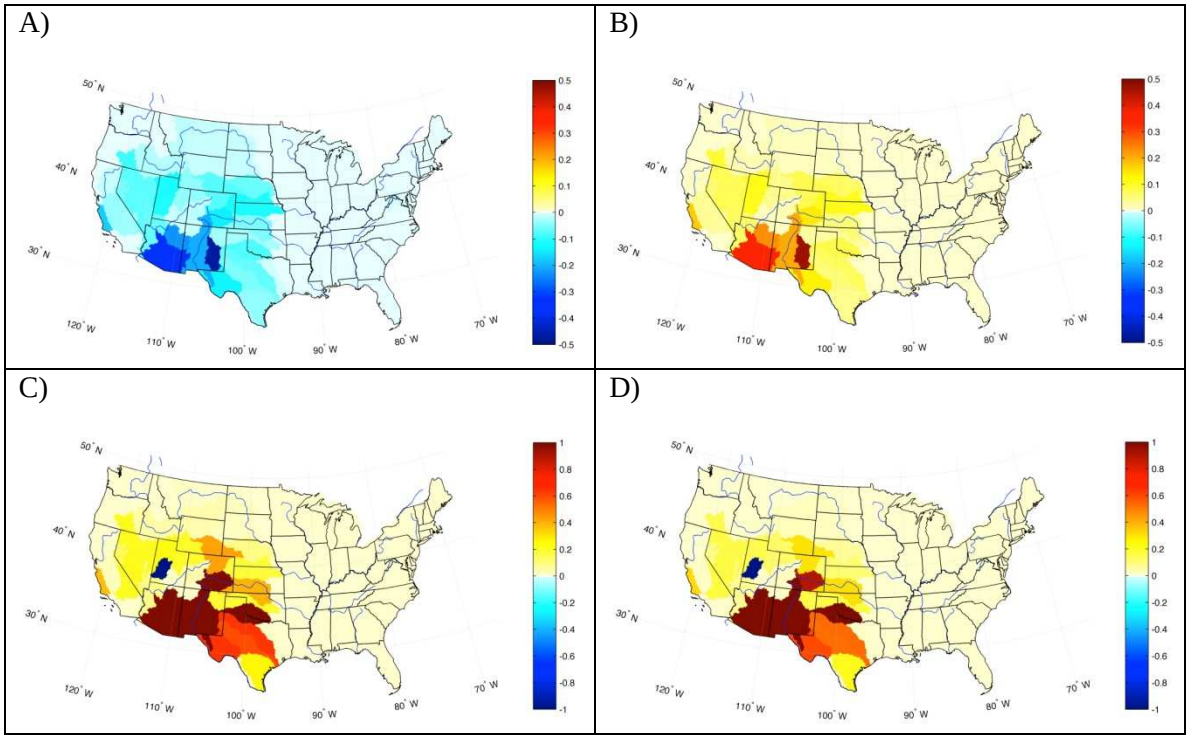


Fig. E.2 Current sensitivity of vulnerability to unit changes in mean precipitation (A), mean evapotranspiration (B), standard deviation of precipitation (C), and standard deviation of evapotranspiration (D) [cm^{-1}].

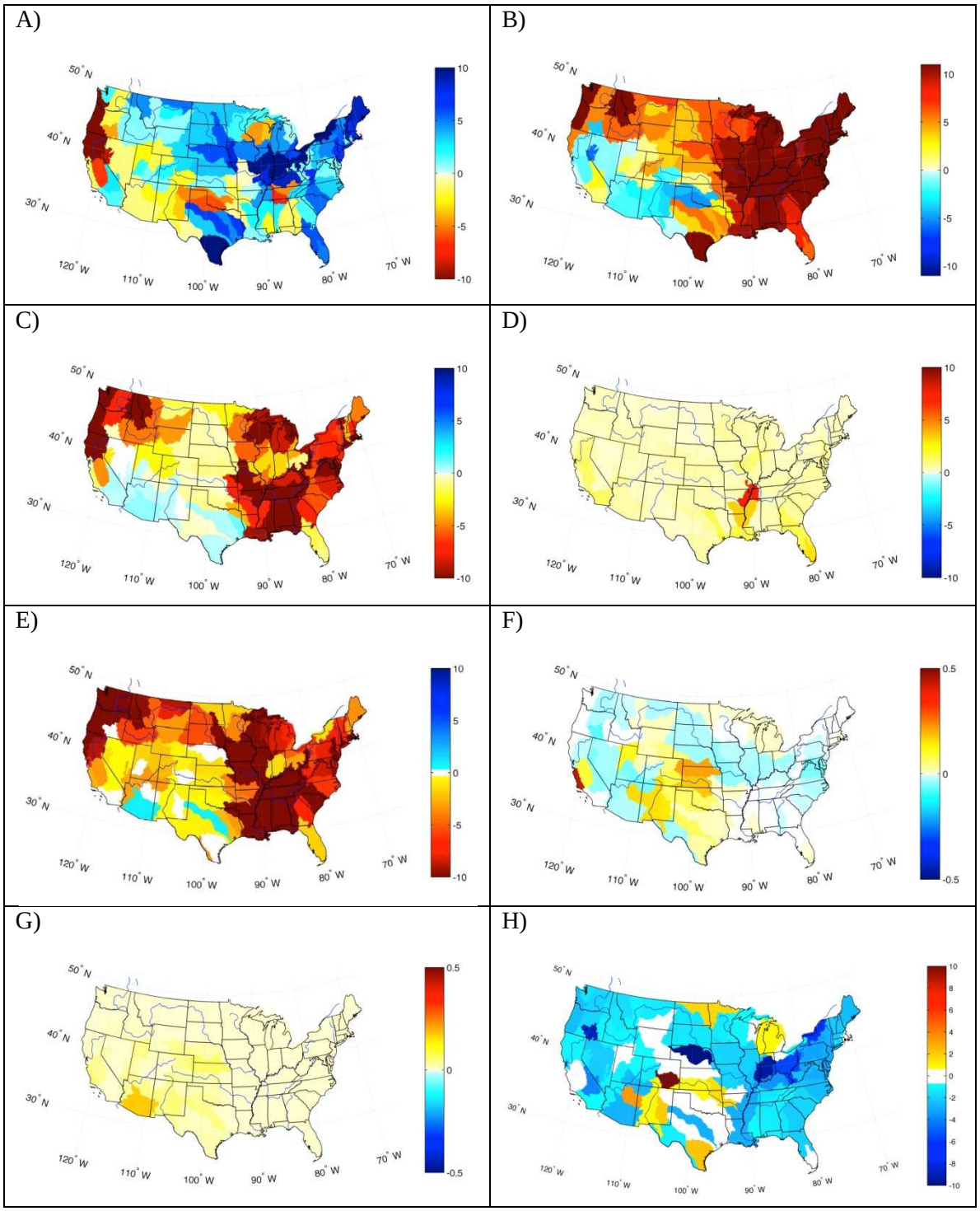


Fig. E.3 Projected changes from the current period to 2060 in mean precipitation (A), mean evapotranspiration (B), mean yield (C), mean demand (D), and mean supply (E) [cm], and changes in vulnerability resulting from changes in supply (F) and demand (G), projected by the CGCM model under the A1B scenario. Ratio of contribution to total change in vulnerability by changes in the variance of surplus to those by changes in the mean of surplus (H).

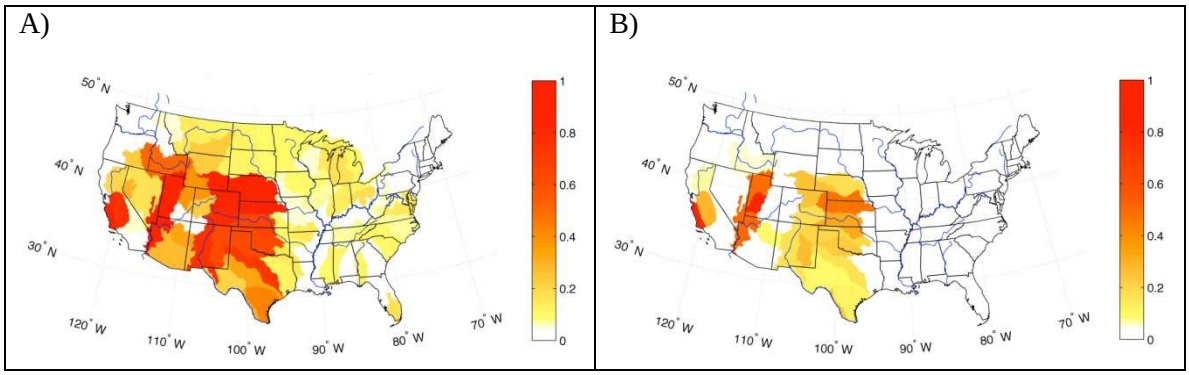


Fig. E.4 Composite for 2060 of maximum (A) and minimum (B) vulnerability projected by six GCM/SRES-scenario combinations.

REFERENCES

Abu Rizaiza, O. S. (1991). "Residential water usage: A case study of the major cities of the western region of Saudi Arabia." Water Resources Research **27**(5): 667-671.

Barbier, N., P. Couteron, et al. (2008). "Spatial decoupling of facilitation and competition at the origin of gapped vegetation patterns." Ecology **89**(6): 1521-1531.

Blaikie, P., T. Cannon, et al. (1994). At Risk: Natural Hazards, People's Vulnerability, and Disasters. London, Routledge.

Boaler, S. B. and C. A. H. Hodge (1962). "Vegetation stripes in Somaliland." Journal of Ecology **50**: 465-524.

Borgogno, F., E. D'Odorico, et al. (2009). "Mathematical models of vegetation pattern formation in ecohydrology." Reviews of Geophysics **47**.

Bromley, J., J. Brouwer, et al. (1997). "The role of surface water redistribution in an area of patterned vegetation in south west Niger." Journal of Hydrology **198**: 1-29.

Brooks, R. H. and A. T. Corey (1966). "Properties of porous media affecting fluid flow." Journal of Irrigation and Drainage Engineering-ASCE **IR2**: 61-88.

Brown, T. C. (2000). "Projecting U.S. freshwater withdrawals." Water Resources Research **36**(3): 769-780.

Brutsaert, W. H. (1967). "Some methods of calculating unsaturated permeability." Trasaction of the ASAE(400-404).

California Department of Water Resources (1998). California Water Plan Update. Sacramento, CA, California Department of Water Resources.

Callaway, R. M., R. W. Brooker, et al. (2002). "Positive interactions among alpine plants increase with stress." Nature **417**: 844-848.

Casenave, A. and C. Valentin (1992). "A runoff capability classification system based on surface features criteria in semi-arid areas of West Africa." Journal of Hydrology **130**: 231-249.

Chan, S. O. and R. L. Bras (1979). "Urban storm water management: distribution of flood volumes." Water Resources Research **15**(2): 371-382.

Colorado Water Conservation Board (1998). River Basin Facts, Colorado Division of Water Resources.

Colorado Water Conservation Board. (2010). "CDSS memoranda." from http://cdss.state.co.us/DNN/.

D'Odorico, E., F. Laio, et al. (2006). "Patterns as indicators of productivity enhancement by facilitation and competition in dryland vegetation." Journal of Geophysics Research **111**.

Daly, C., R. P. Neilson, et al. (1994). "A statistical-topographic model for mapping climatological precipitation over mountainous terrain." Journal of Applied Meteorology **33**(2): 140-158.

Dunkerley, D. and K. Brown (1999). "Banded vegetation near Broken Hill, Australia: significance of surface roughness and soil physical properties." Catena **37**: 75-88.

Eagleson, P. S. (1978). "Climate, soil and vegetation.1. Introduction to water balance dynamics." Water Resources Research **14**(5): 705-712.

Eagleson, P. S. (1978). "Climate, soil, and vegetation.1. Introduction to water-balance dynamics." Water Resources Research **14**(5): 705-712.

Eagleson, P. S. (1978). "Climate, soil, and vegetation. 2. Distribution of annual precipitation derived from observed storm sequences." Water Resources Research **14**(5): 713-721.

Eagleson, P. S. (1978). "Climate, soil, and vegetation. 3. Simplified model of soil-moisture movement in liquid-phase." Water Resources Research **14**(5): 722-730.

Eagleson, P. S. (1978). "Climate, soil, and vegetation. 4. Expected value of annual evapotranspiration." Water Resources Research **14**(5): 731-739.

Eagleson, P. S. (1978). "Climate, soil, and vegetation. 5. Derived distribution of storm surface runoff." Water Resources Research **14**(5): 741-748.

Eagleson, P. S. (1978). "Climate, soil, and vegetation. 6. Dynamics of annual water-balance." Water Resources Research **14**(5): 749-764.

Eagleson, P. S. (1978). "Climate, soil, and vegetation. 7. Derived distribution of annual water yield." Water Resources Research **14**(5): 765-776.

Eagleson, P. S. (1978). "Climate, soli and vegetation .2. Distribution of annual precipitation derived from observed storm sequences." Water Resources Research **14**(5): 713-721.

Eagleson, P. S. (1978). "Climate, soli and vegetation .3. Simplified model of soil moisture movement in liquid phase." Water Resources Research **14**(5): 722-730.

Eagleson, P. S. (1978). "Climate, soli and vegetation .4. Expected value of annual evapotranspiration." Water Resources Research **14**(5): 731-739.

Eagleson, P. S. (1978). "Climate, soli and vegetation .5. Derived distribution of storm surface runoff." Water Resources Research **14**(5): 741-748.

Eagleson, P. S. (1978). "Climate, soli and vegetation .6. Dynamics of annual waer balance." Water Resources Research **14**(5): 749-764.

Eagleson, P. S. (1978). "Climate, soli and vegetation .7. Derived distribution of annual water yield." Water Resources Research **14**(5): 765-776.

FAO (2001). Lecture notes on the major soils of the world.

Grayman, W. M. and P. S. Eagleson (1971). Evaluation of radar and raingage systems for flood forecasting. Cambridge, Massachussets Institute of Technology. **138**.

Greig-Smith, P. (1979). "Pattern in Vegetation." Journal of Ecology **67**: 755-779.

HilleRisLambers, R., M. Rietkerk, et al. (2001). "Vegetation patter formation in semi-arid grazing systems." Ecology **82**(1): 50-61.

Holmgren, M., M. Sheffer, et al. (1997). "The interplay of facilitation and competition in plant communities." Ecology **78**(7): 1966-1975.

Hurd, B., N. Leary, et al. (1999). "Relative regional vulnerability of water resources to climate change." Journal of the American Water Resources Association **35**: 1399–1409.

Hutson, S. S., N. L. Barber, et al. (2004). Estimated use of water in the United States in 2000. Reston, VA, U. S. Geological Survey**:** 46.

Hydrosphere Data Products (1996). Hydrodata USGS Daily Values, Volume 8.0. Boulder, CO, Hydrosphere Data Products, Inc.

Joyce, L. A., D. T. Price, et al. (in preparation 2). Climate change scenarios for the United States: a technical document supporting the Forest Service 2010 RPA Assessment. Fort Collins, CO, Rocky Mountain Research Station, US Forest Service.

Joyce, L. A., D. T. Price, et al. (in preparation 1). High resolution statistical interpolation of IPCC AR4 GCM climate scenarios for the coterminous USA and Alaska: a contribution to the USDA Forest Service 2010 national renewable resource assessment. Fort Collins, CO, Rocky Mountain Research Station, US Forest Service.

Kelly, P. M. and W. N. Adger (2000). "Theory and practice in assessing vulnerability to climate change and facilitating adaptation." Climatic Change **47**(4): 325-352.

Kenny, J. F., N. L. Barber, et al. (2009). Estimated use of water in the United States in 2005. Reston, VA, U. S. Geological Survey**:** 52.

Kittel, T. G. F., N. A. Rosenbloom, et al. (1996). The VEMAP Phase I Database: An Integrated Input Dataset for Ecosystem and Vegetation Modeling for the Conterminous United States**:** CDROM and World Wide Web

Kittel, T. G. F., N. A. Rosenbloom, et al. (1995). "The VEMAP integrated database for modeling United States ecosystem/vegetation sensitivity to climate change." Journal of Biogeography **22**(4-5): 857-862.

Klausmeier, C. A. (1999). "Regular and irregular patterns in semi-arid vegetation." Science **284**: 1826-1828.

Kochendorfer, J. P. (2005). A monthly, two-soil-layer statistical-dynamical water balance model for ecohydrologically focused climate impact assessments. Civil Engineering. Fort Collins, Colorado State University**:** 287.

Korchendorfer, J. P. and J. A. Ramirez (1996). Integrated hydrological, ecological, economic modeling for examining the vulnerability of water resources to climate change. North American Water and Environmental Congress '96, New York.

Krug, W. R., W. A. Gebert, et al. (1989). Preparation of average annual runoff map of the United States, 1951-80. Madison, Wisconsin, U.S. Geological Survey**:** 414.

Labadie, J. W., A. M. Pineda, et al. (1984). Network analysis of raw water supplies under complex water rights and exchanges: Documentation for Program MODSIM3. Fort Collins, CO, Colorado Water Resources Institute, Colorado State University**:** 94.

Lefever, R. and O. Lejeune (1997). "On the origin of tiger bush." Bulletin of Mathematical Biology **59**(2): 263-294.

Leprun, J.-C. (1999). "The influences of ecological factors on tiger bush and dotted bush patterns along a gradient from Mali to northern Burkina Faso. ." Catena **37** 25-44.

Linacre, E. T. (1977). "A simple formula for estimating evaporation rates in various climates using temperature data alone." Agricultural Meteorology **18**: 409-424.

Litke, D. W. and C. L. Appel (1989). Estimated use of water in Colorado, 1985. Denver, CO, U.S. Geological Survey**:** 157.

Lobell, D. B. and G. P. Asner (2002). "Moisture Effects on Soil Reflectance." Soil Science Society of America **66**: 722-727.

Mooty, W. B. and H. H. Jeffcoat (1986). Inventory of interbasin transfers of water in the eastern United States. Tuscaloosa, AL, U.S. Geological Survey**:** 47.

Mooty, W. S. and H. H. Jeffcoat (1986). Inventory of interbasin transfers of water in the easter United States. Tuscaloosa, AL, U.S. Geological Survey.

Muchiri, P. W. (2007). Climate of Somalia. Nairobi, FAO-SWALIM.

Nakicenovic, N., J. Alcamo, et al. (2000). Emissions scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge, UK, Cambridge University Press**:** 599.

Petsch, H. E., Jr. (1985). Inventory of interbasin transfers of water in the western conterminous United States. Lakewood, CO, U.S. Geological Survey**:** 45.

Phillip, J. R. (1969). The theory of infiltration. Advances in hydroscience. V. T. Chow. New York, Academic. **5:** 215-296.

Postel, S. L. (2000). "Entering an era of water scarcity: the challenge ahead." Ecological applications **10**(4): 941-948.

Price, D. T., D. W. McKenney, et al. (2006). High-resolution climate change scenarios for North America. Sault Ste. Marie, ON, Canadian Forestry Service.

Rechid, D., T. J. Raddatz, et al. (2009). "Parameterization of snow-free land surface albedo as a function of vegetation phenology based on MODIS data and applied in climate modelling." Theoretical and Applied Climatology **95**: 245-255.

Rietkerk, M., M. C. Boerlijst, et al. (2002). "Self-organization of vegetation in arid ecosystemns." The American Naturalist **160**(4): 524-530.

Rietkerk, M. and J. Van de Koppel (2008). "Regular pattern formation in real ecosystems." Trends in Ecology and Evolution **23**(3).

Roy, S. B., C. L., et al. (2010). Evaluating sustainability of projected water demands under future climate change scenarios. T. T. Inc. Lafayette, Tetra Tech Inc.

Saco, P. M., G. R. Willgoose, et al. (2007). "Eco-geomorphology of banded vegetation patterns in arid and semi-arid regions." Hydrology and Earth System Sciences Discussions **11**: 1717-1730.

Sailor, D. J. (2001). "Relating residential and commercial sector electricity loads to climate--evaluating state level sensitivities and vulnerabilities." Energy **26**(7): 645-657.

Sailor, D. J. and A. A. Pavlova (2003). "Air conditioning market saturation and long-term response of residential cooling energy demand to climate change." Energy **28**(9): 941-951.

Scanlon, T. M., K. K. Caylor, et al. (2007). "Positive feedbacks promote power-law clustering of Kalahari vegetation." Nature **449**.

Schneider, S. H., S. Semenov, et al. (2007). Assessing key vulnerabilities and the risk from climate change. Climate Change 2007: Impacts, Adaptation

and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK, Intergovernmental Panel on Climate Change**:** 779-810.

Sherrat, J. A. (2005). "An analysis of vegetation stripe formation in semi-arid landscapes." Journal of Mathematical Biology **51**(2): 183-197.

Slack, J. R. and J. M. Landwehr (1992). Hydro-climatic data network (HCDN): A U. S. Geological Survey streamflow data set for the United States for the study of climate variations, 1874-1988, U. S. Geologic Survey.

Solley, W. B., C. F. Merk, et al. (1988). Estimated use of water in the United States in 1985. Denver, CO, U. S. Geological Survey.

Solley, W. B., R. R. Pierce, et al. (1993). Estimated use of water in the United States in 1990. Denver, CO, U. S. Geological Survey.

Solley, W. B., R. R. Pierce, et al. (1998). Estimated use of water in the United States in 1995. Denver, CO, U. S. Geological Survey.

Tennant, D. L. (1976). "Instream flow regimens for fish, wildlife, recreation and related environmental resources." Fisheries **1**(4).

Thiéry, J., J.-M. d'Herbès, et al. (1995). "A model for simulating the genesis of banded patterns in Niger." Journal of Ecology **83**: 497-507.

U.S. Bureau of Census. (2004). "U.S. interim projections by age, sex, race, and Hispanic origin: 2000-2050." Released March 18, 2004, from http://www.census.gov/population/www.projections/usinterimproj/.

U.S. Water Resources Council (1978). The nation's water resources 1975-2000. Washington, D.C., U. S. Government Printing Office.

Valentin, C., J. M. d'Herbès, et al. (1999). "Soil and water components of banded vegetation patterns." Catena **37**(1-24).

Venema, J. H. (2007). Land resources assessment of Somalia. Nairobi, FAO-SWALIM.

Vörösmarty, C. J., E. M. Douglas, et al. (2005). "Geospatial Indicators of Emerging Water Stress: An Application to Africa." Ambio **34**(3).

Vörösmarty, C. J., P. Green, et al. (2000). "Global Water Resources: Vulnerability from Climate Change and Population Growth." Science **289**: 284-288.

Wang, K., P. Wang, et al. (2005). "Variation of surface albedo and soil thermal paramenters with soil moisture content at a semi-desert site on the western Tibetan plateau." Boundary-Layer Meteorology **116**: 117-129.

Weib, M. and J. Alcamo (2011). "A systematic approach to assessing the sensitivity and vulnerability of water availability to climate change in Europe." Water Resources Research **47**.

Worral, G. A. (1959). "The Butana grass patterns." Journal of Soil Science **10**: 34-53.

Worral, G. A. (1960). "Patchiness in vegetation in the northern Sudan." Journal of Ecology **48**: 107-117.

Zarnoch, S. J., H. K. Cordell, et al. (2008). Projecting county-level populations under three climate change future scenarios for the 2010 RPA Assessment. Athens, GA, Internet Research Information Series (IRIS).

Zarnoch, S. J., H. K. Cordell, et al. (2010). Projecting county-level populations under three climate change future scenarios for the 2010 RPA Assessment. Ashville, NC, Southern Research Station, U.S. Forest Service. **SRS-128**.