WHERE SHOULD FINE-RESOLUTION HETEROGENEITY BE CAPTURED IN LAND SURFACE MODELS?

By

RYAN EDWARD HULL

A thesis submitted in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE IN CIVIL ENGINEERING

WASHINGTON STATE UNIVERSITY
Department of Civil and Environmental Engineering

DECEMBER 2016

To the Faculty of Washington State University:

The members of the Committee appointed to examine the thesis of RYAN EDWARD HULL find it satisfactory and recommend that it be accepted.

Jennifer Adam, Ph.D., Chair

Nicholas Engdahl, Ph.D.

Ming-Liang Liu, Ph.D.

Joseph Vaughan, Ph.D.

ACKNOWLEDGEMENT

I owe a great debt of gratitude to many people for the completion of this thesis, far more than I could possibly mention in this small space, but I specifically want to acknowledge the people who have worked with me most often and directly.

Thanks to:

My advisor, Dr. Jenny Adam: You have been a source of ceaseless encouragement, wonderful advice, and honest feedback. Your mentorship has set me up for success, now and in the future.

Dr. Ming-Liang Liu: Thank you for helping me through all of the miscellaneous problems I encountered trying to work with RHESSys, for your help generating maps from my results, and for your time spent on my thesis committee.

Dr. Engdahl and Dr. Vaughan: Thank you for agreeing to serve on my thesis committee. The early feedback you provided helped me to clarify the purpose of this thesis. One of the few regrets I have is not seeking out your input earlier and more often.

Dr. Naomi Tague, whose conversations inspired the research topic for this thesis and whose advice helped keep it moving forward.

All of my fellow grad students who worked in Jenny's lab or shared office space with me. Every single one of you helped me complete this thesis. Some of you helped by giving suggestions on how to run RHESSys or adjust model parameters. For others it was just a smile when I walked in the room. Regardless, thank you all for being there for me.

WHERE SHOULD FINE-RESOLUTION HETEROGENEITY

BE CAPTURED IN LAND SURFACE MODELS?

Abstract

by Ryan Edward Hull, M.S.
Washington State University
December 2016

Chair: Jennifer Adam

Climate change and variability and its associated impacts are influenced by the interplay between land and atmospheric processes. Quantifying these effects is one of the roles of land surface models, such as macroscale hydrologic models, which requires an understanding of the underlying physics as well as access to appropriate input data and adequate computational tools. Currently, large-scale hydrologic simulations are limited by the availability of both high-resolution gridded datasets and computational power so there is a need to critically consider when and where to apply high-resolution models. Despite the inherent nonlinearity of most hydrologic processes, their complex interactions can sometimes lead to nearly linear response curves. These areas of linear response can be upscaled with minimal increase in the overall error as long as lateral interactions, such as soil moisture redistribution between fine-scale spatial elements, is not a critical factor for the aggregate response. In this study we address three questions: Is it possible to aggregate over large climate gradients? Does accounting for fine-scale lateral redistribution impact watershed-scale response over climate gradients? And how can we use this information to inform upscaling efforts.

First, we ran a series of simulations using a hydro-ecological model (RHESSys),

independently scaling temperature (from -5 to +5 ˚C) and precipitation (25 – 175%). Then we computed a linear regression at each grid cell for evapotranspiration (*ET*) and net surface and subsurface flow ($subsurface_{in}$ + $runoff_{in}$ – $subsurface_{out}$ – $runoff_{out}$; *Q*) over different sized ranges of climate perturbations and used the square of Pearson's correlation coefficient as a measure of linearity. To test the impact of explicit lateral connectivity between grid cells, we repeated the above analysis with lateral redistribution disabled and compared the results. Finally, we generated "upscalability" recommendations per bioclimatic zone according to the distribution of linearity scores and found that whether upscaling is possible for a given response variable depends on the size of the climate gradient and the process being studied. These results have applications for upscaling existing models and datasets, and identifying the appropriate process scale when initializing new models.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction:

The impacts of climate change can be enhanced or dampened depending upon the response of the hydrologic cycle and the co-incident land-atmosphere interactions (e.g. soil moisture acting as a store of water and energy, Seneviratne et al. 2010). Simulating these impacts is scale-dependent in terms of both modeling architecture and input data. Data at some resolutions are unavailable at large spatial scales because techniques to acquire them have not been developed or they are very difficult to collect in a cost-effective way. Advancements in remote sensing techniques have progressively improved the quality of spatial datasets, and computational models are advancing to take advantage of them, however a Grand Challenge call has been issued to focus the research community's efforts toward a larger jump in hydrologic modeling capabilities through the collection of global, hyper-resolution datasets and development of the computation tools to work with them (Wood et al. 2011, Clark et al. 2015). However, as model and data resolutions increase, practical concerns such as computing time and data storage limits increase too. There is therefore a need for critical analysis of when and where hyper-resolution models are most valuable.

An intermediate step between the current modeling paradigm and full hyper-resolution modeling would focus efforts using the highest-resolution only on areas where such techniques are most important, while using less expensive (in terms of both computational costs and data acquisition) techniques elsewhere. Some modeling architectures do not require uniform resolution among the model components or input data, so it is possible to selectively apply hyper-resolution. The key, then, is to identify the ideal resolution for modeling hydrologic processes relevant to the research

question. The shapes of the curves of model response variables, as a function of the input variable that is being aggregated, hold one key to this distinction. If a response curve is highly nonlinear at fine resolution, it will be susceptible to aggregation errors at reduced resolutions and capturing fine-scale heterogeneity will be important. On the other hand, if the response curve is linear, it may be possible to upscale, or aggregate, to a lower resolution without the introduction of additional error, assuming the spatial organization of the elements does not also affect the shape of the response curve. The lateral connectivity of fine-scale spatial elements is often ignored in hydrologic studies, although it can be non-trivial (Tague and Peng, 2013). If this fine-scale connectivity is also important to the overall response, then a fine model resolution must be used regardless of the shape of the response curves.

In the long term, the goal is to be able to upscale catchment-scale models to the regional- or higher-scale as efficiently as possible while retaining maximum model resolution where it is really needed. Accomplishing that will require a thorough understanding of which types of heterogeneity are important, and at which scales, to capture the large-scale, aggregate hydrologic response. Many hydrologic processes are nonlinear in nature which intuitively suggests that high-resolution modeling should reduce error, however interacting and competing effects result in complex behavior that can be approximately linear. This behavior can then be exploited by selectively reducing model resolution. The primary goal of this study is determine if we can aggregate over areas of large climate gradients. We do this by testing the linearity of evapotranspiration and net lateral flow (runoff plus subsurface flow) in an eco-hydrologic model as a function of changes in temperature and precipitation. These are done by bioclimatic zones to explore the role that land surface parameters (e.g., vegetation and land use) play in upscaling

potential. We also test the role of explicit lateral connectivity (through the redistribution of soil moisture) between the model's fine-scale spatial elements.

2 Background – Heterogeneity in Hydrologic Models:

2.1 Debates Over Heterogeneity:

Wood et al. (2011) called on the research community to focus resources on the development of hyper-resolution (<100 m resolution) hydrologic datasets, improved models capable of utilizing these data, and the development of computational systems necessary to run these models. They argue that the key to greater understanding of hydrologic systems lies in the explicit representation of as much detail as possible, in terms of both representation of heterogeneity and modeling of hydrologic processes. The call for hyper-resolution models and data sets, then, is driven by the assumption that higher resolution implies better representation of all types of spatial heterogeneity.

Beven and Cloke (2012) respond to Wood et al. with two primary criticisms: failure to acknowledge epistemic uncertainty in process representation and observations; and the problem that regardless of scale there will always be sub-grid heterogeneity in natural systems. They do not argue that improving process representation and data resolution are a poor use of effort, but instead argue that until techniques have been developed to account for problems that exist with current-scale data the gains from hyper-resolution will not be as likely. Both Wood et al. and Beven and Cloke's arguments acknowledge the need for improved representation of hydrologic

processes, a better understanding of how error is introduced using current techniques, and how heterogeneity in the natural environment is accounted for.

2.2 Types of Heterogeneity in Hydrologic Models:

Heterogeneities are attributes of the environment that vary with space or time, and can be grouped broadly into three different categories as they pertain to hydrologic models: land parameters, meteorological forcings, and lateral connectivity. Each of these will be described below.

Land Parameterization. These parameters represent the vegetation, soil, land use, and topographic characteristics that control land surface processes. In a watershed, the dominant vegetation type, a land surface parameter, often varies from its upper reaches to the drainage point. If the watershed is modeled using a very high resolution grid, most grid cells will be dominated by a single vegetation class, each cell effectively representing a homogeneous spatial unit. As the cell size increases, however, they might encompass more than one vegetation type. If these cells are assumed to be homogeneous, parametric information is lost, but if we know how much of the grid cell's land area is of each vegetation type it is possible to account for this in model calculations using an area-weighted statistical distribution.

Meteorological Forcings. Meteorological data needed to drive land surface models are often limited to temperature, precipitation, and wind speed; with radiation and humidity variables calculated from these input variables using empirical relationships (e.g., Thornton and Running, 1999). While all of these forcing variables have their own inherent limitations for driving fine-resolution hydrologic models, for this discussion we focus on precipitation. Precipitation

forcings for distributed hydrologic models are usually derived by interpolating between point measurements and applying corrections for errors at the point scale (e.g., for wind-induced undercatch; Adam and Lettenmaier 2003) or due to gridding (e.g., over complex terrain where orographic effects have a strong influence on precipitation; Daly et al. 2008 and Adam et al. 2006). However, with modern remote sensing, atmospheric modeling, and reanalysis technologies (e.g., Kidd et al. 2012) it is increasingly becoming possible to generate high-resolution gridded land surface model forcings datasets in locations that are not possible with *in situ* observations alone. As these techniques improve, we will become less limited by the resolution and accuracy of these input forcings. In this study we are assuming that high-resolution meteorological data will eventually become available that do not limit high-resolution hydrologic modeling applications, and that model resolution is our limiting factor. Currently, however, there are limitations in both data and model resolutions.

Lateral Connectivity. Another type of spatial heterogeneity is the horizontal (i.e., lateral) interconnection between grid cells. Lateral connectivity is important for obvious mechanisms such as subsurface flow and runoff and is responsible for much of the transport of nutrients and water in a system (Stieglitz et al., 2003), but it is also relevant for problems like fire spread and beetle forest kill, and can be directly influenced by human land use (Callow and Smettem, 2009). Several studies have assessed the effect of lateral connectivity in hydrologic models. Tague and Peng (2013) specifically tested the impacts of enabling lateral connectivity using a distributed eco-hydrologic model (RHESSys; described later in this paper) in a forested catchment and found that overall the aggregate ET response was roughly 10% higher with connectivity enabled. Maxwell and Condon (2016) conducted a study on the effects lateral hydrologic flows have on the partitioning of fluxes between evaporation and transpiration at continental scales, and found

that lateral flows resulted in a highly significant increase in the share of water lost due to transpiration. They hypothesize that this is due to the lateral flow's ability to stabilize soil moisture levels which can increase moisture in the root zone where it is available for transpiration but deep enough to avoid direct evaporation. We hypothesize that lateral flow will play a more significant role in semi-arid regions where there exists a closer balance between water availability and potential ET.

2.3 Representation of Heterogeneity in Current Hydrologic Models:

Distributed, process-based hydrologic models use varying degrees of vertical and lateral complexity to represent heterogeneity (Newman 2014). "Integral" (Todini 1998) land surface models (LSMs) were originally developed primarily for land-atmosphere interaction studies and use statistical distributions to represent sub-grid heterogeneity within explicitly-located but large, independent grid cells, and so have high vertical complexity but little horizontal complexity. There is no representation of fine-scale lateral connectivity in integral models. Examples include the VIC model (Liang, 1994), CLM (Oleson et al. 2010) and NoahMP (Niu et al. 2011).

"Differential" (Todini 1998) hydrologic models use much smaller, explicitly-located spatial elements which interact directly with one another. Thus, lateral connectivity is captured in a spatially-explicit manner. Examples include the Distributed Hydrology Soil Vegetation Model (DHSVM; Wigmosta et al. 1994) and the Regional Hydro-Ecologic Simulation System (RHESSys; Tague and Band, 2004) and). For a more detailed list and a subjective rating of the relative vertical and horizontal complexities of various models, see the summary in Newman

(2014).

It is important to note that both of these approaches lie on a continuum that is scale dependent, each with its own tradeoffs. Very fine grid cells are spatially explicit and contain less subgrid heterogeneity than very coarse grid cells, but they come with increased complexity and computational cost. Very coarse grid cells may still accurately represent certain types of subgrid heterogeneity and require less computational time, but sacrifice fine-scale spatial relationships. With the right computational framework, it is possible to utilize a middle-ground approach with flexible-resolution spatial elements that maintain lateral connectivity at moderate scales but also incorporate statistical subgrid heterogeneity. This results in a model that preserves some of the lateral connectivity among spatial elements without the full computational cost of a hyper-resolution implementation. Some newer modeling platforms have been developed that incorporate lateral connectivity into a larger-scale framework and, in many of these, spatial units can be of any shape or size; however, spatial resolution does not vary across the domain. Examples of these include models include HydroBlocks (Chaney et al. 2016), WRF-Hydro (Gochis et al. 2013), and Hybrid-3D (Hazenberg et al. 2015). There is need for frameworks that also combine subgrid variability with lateral connectivity (as with the SUMMA model; Clark et al. 2015) and have the flexibility to allow spatial resolution to vary across the domain, such that the appropriate scales are used in each location.

2.4 Need for a Flexible Resolution (Across Space) Modeling Framework:

Heterogeneity is an inherent aspect of all natural systems. No two study areas are identical, and

no single site is absolutely homogeneous. Whether particular types of heterogeneities can be ignored is a function of the process being studied, the degree of heterogeneity, and the spatial and temporal scales involved. Historically, hydrologic studies were done over relatively small spatial areas so heterogeneity could be estimated by direct observation. However, as methodologies have improved, so have the extents of the regions being studied (Eagleson, 1994), with attempts to run simulations at high resolution over global scales becoming common (e.g. Wood et al., 2011; Bierkens et al., 2015). Despite the rapid growth and increasing availability of computational power, access is still an issue (Khan et al., 2014) and inconsistent data availability can make the use of hyper-resolution models infeasible.

Traditional calibration methods applied to a hyper-resolution model at the regional or broader scale are typically intractable with current technologies, though alternatives are being developed, such as the regional parameterization approach of Kumar et al. (2013). Another possible path to reduce computational load is to develop a modeling framework that allows for flexibility in how and at what scales spatial heterogeneity is represented, even varying across a single region. This allows computational resources to be focused in locations where they are most needed. Depending on the research question, maximum resolution and explicit representation of spatial heterogeneity may neither be necessary nor desirable. By differentiating those sub regions where hyper-resolution is valuable from those where it is unnecessary, we can selectively reduce computational load by using a flexible modeling framework capable of running at variable resolutions. SUMMA (Clark et al. 2015a,b) is one notable example of this framework.

As part of the BioEarth Project (Adam et al. 2014), we are implementing “Upscaled-RHESSys”.

RHESSys does not use rectangular grids but, like the SUMMA model and an adaptation of CLM for sub-basins (SCLM; Tesfa et al. 2014), uses irregular elements (called “patches” in the RHESSYs model) as determined by the user. Upscaled-RHESSys (see Figure 1) implements sub-patch variability using distributions of land cover, soil, and topographic parameters; we call these “aspatial patches”. In short, Upscaled-RHESSys combines the advantages of both integral and differential models because lateral connectivity between the upscaled patches, which are still explicitly located in space, is retained. Therefore, land parameter heterogeneity can be accurately represented at any resolution using aspatial patches; and heterogeneity in meteorological forcings, such as precipitation and temperature, can be explicitly captured at the scale of the patch and implicitly captured through the use of lapse rates across elevation bands within the aspatial patches. Explicit lateral connectivity and redistribution, components often neglected in large-scale studies, are captured at the resolution of the finest-scale spatial elements. If aspatial patches with sub-grid heterogeneity are used to upscale patch sizes, then some of the spatial organization and connectivity from the fine-scale data is lost, but the explicit lateral connectivity remains at the same scale as the spatially-explicit patches.

These flexible-resolution frameworks are critical for efficiently representing heterogeneities in hydrologic models while minimizing computational requirements, particularly when a large number of modeling scenarios are needed for the experimental design. The key when modeling and focusing on any of these types of heterogeneity is finding an optimal resolution that minimizes error and maximizes computational efficiency. Scale transition theory can be used to understand where and what type of heterogeneity is needed given a specific scientific question.

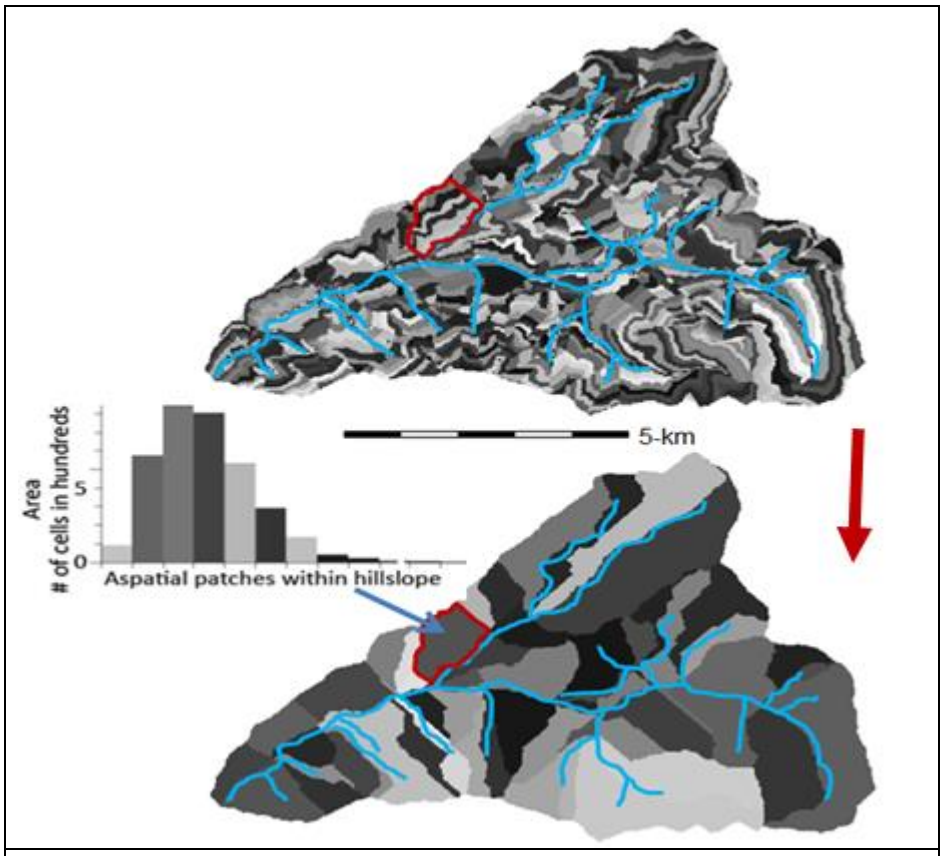


Figure 1: Traditional RHESSys implementation in which fine-resolution, explicitly-located patches with fine-scale laterally-connected elements (top), as compared to Upscaled-RHESSys in which aspatial patches are used to represent land parameter heterogeneity within a patch and where lateral connectivity is maintained between the spatially-explicit patches at a coarser resolution (bottom).

3 Linearity and Upscaling:

The appropriate model scale depends on the research question and the processes being modeled, and might vary across the study area. Generally speaking, higher model resolution means improved model accuracy. However, even if hyper-resolution datasets become widely available, it may not be desirable for a model's finest spatial units to match that resolution because higher

resolution entails increased computational requirements. When using computationally-intensive models, it is important to find a reasonable balance between better resolution and lower computational cost. In this study we consider how evapotranspiration and net lateral flow change in response to perturbations in temperature and precipitation. We seek ways to identify where finer resolution simulations are the most important and where lower resolutions can be used. To that end, we consider the three following questions, noting that this same approach is generalizable to other inquiries by changing the response variables and forcing function (see Figure 2 for a more detailed depiction of the generalized upscaling criteria):

1: Can we aggregate over areas of large climate gradients?

2: Does accounting for fine-scale lateral redistribution impact watershed-scale response over climate gradients?

3: How can this information inform hydrologic model upscaling?

Statistical upscaling, or aggregation, is the science of moving from smaller to larger; in this case, moving from high spatial resolution data to some lower resolution. The goal is to minimize aggregation error and maintain the fine-scale interconnectivity where it is needed. The types of heterogeneity described earlier influence upscaling in different ways. 1) Land parameter heterogeneity can be accounted for in both spatially explicit and aspatial modes (through parameter distributions), therefore lending itself well to upscaling. 2) Lateral connectivity, on the other hand, requires an explicit representation of spatial relationships and cannot be upscaled if fine-scale connectivity is an important contributor to the aggregate model response. 3) With respect to meteorological forcings, there is a mixed effect. Upscaling of patch size also results in

upscaling of forcing data as meteorological forcings are homogenous over the patches (unless relationships between aspatial elements and forcings are implemented, such as through temperature lapse rates which are often problematic given their dynamic nature; e.g., Minder et al. 2010). Therefore, whether or not upscaling results in aggregation error depends on the linearity of the model output in response to perturbations of model forcings (Chesson 2012, Melbourne and Chesson 2005, and shown in Figure 3). Some researchers are working on ways to quantify and reduce this error using statistical methods (e.g. Kure et al., 2011).

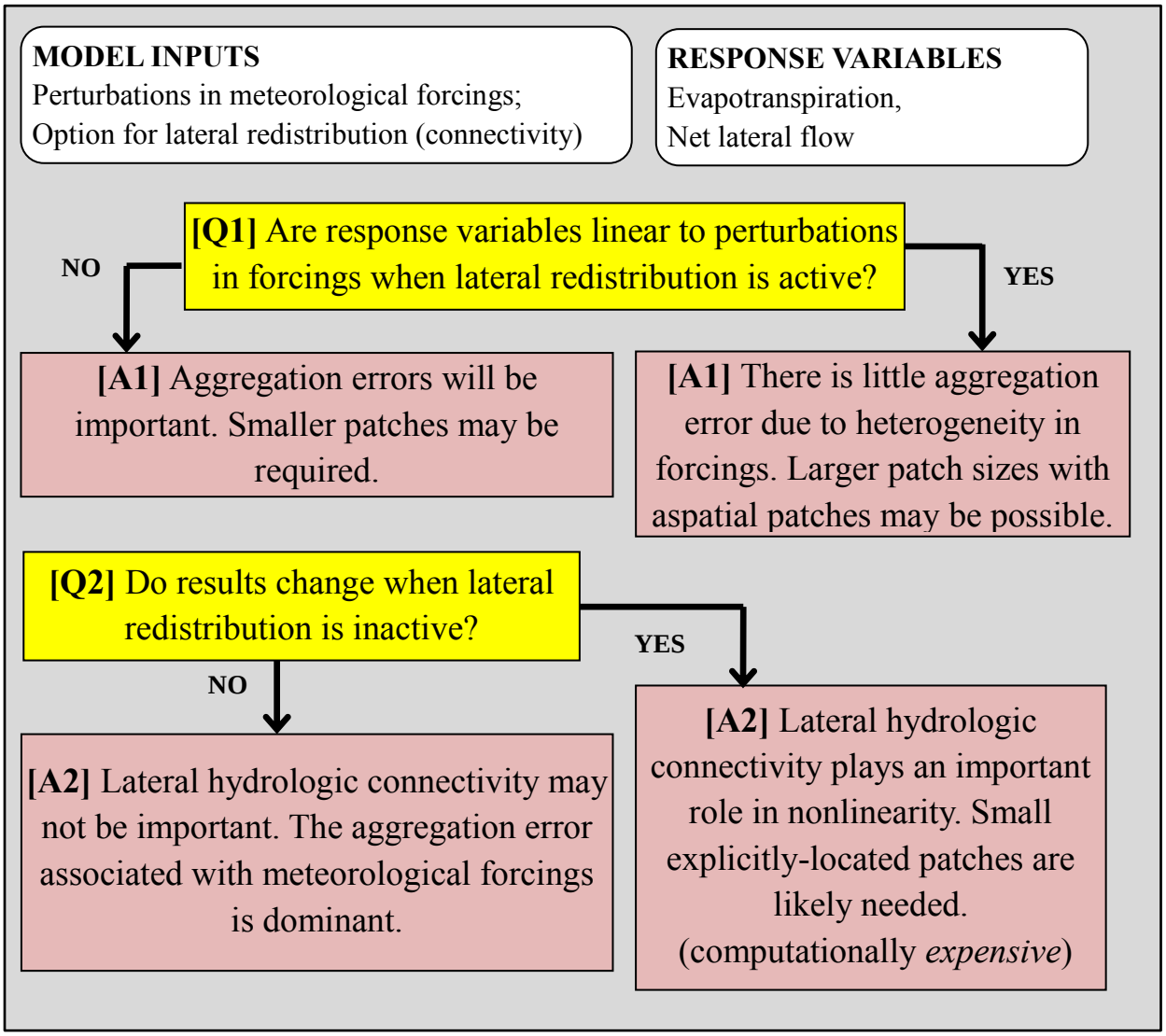


Figure 2: Decision tree for whether upscaling is possible and how heterogeneity of different types should be incorporated. If a model responds linearly to changes over some forcing variable and lateral redistribution is not significant then upscaling is p

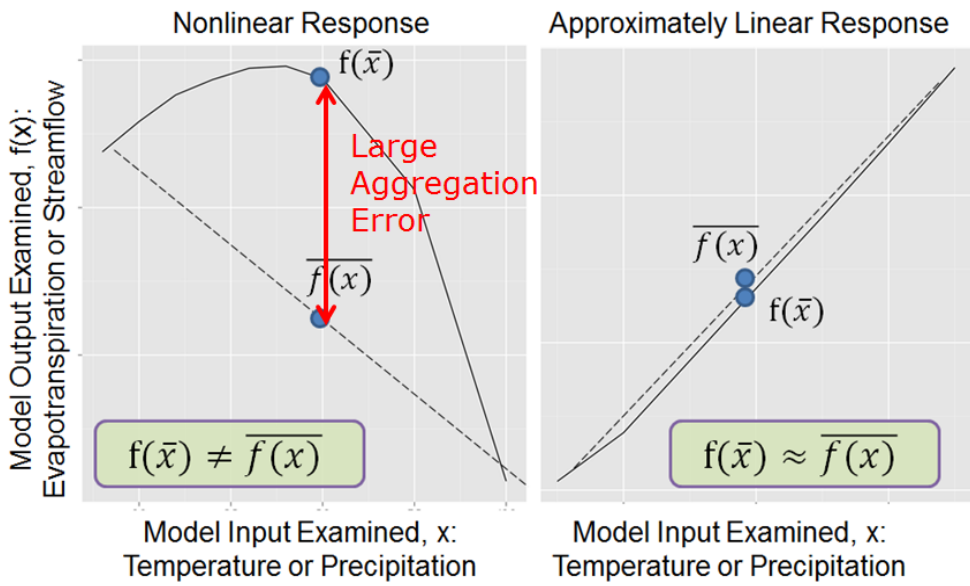


Figure 3: Differences in aggregation error for nonlinear vs. linear response curves between model input, x, and model output, f(x). Based on Rastetter et al. (1992).

However, we note that hydrologic processes interact in complex ways and hypothesize that under certain circumstances these interactions could result in nearly-linear, large-scale aggregate response curves of certain variables for relevant ranges of temperature and precipitation. The more linear the response curve over a range of drivers, the less aggregation error will accrue when upscaling a region encompassing those drivers. In this study, we identify regions where the model response is approximately linear and in so doing differentiate between the areas where hyper-resolution models may be necessary (non-linear response) from those where they may not (linear response).

4 Research Methods:

4.1 Study Area:

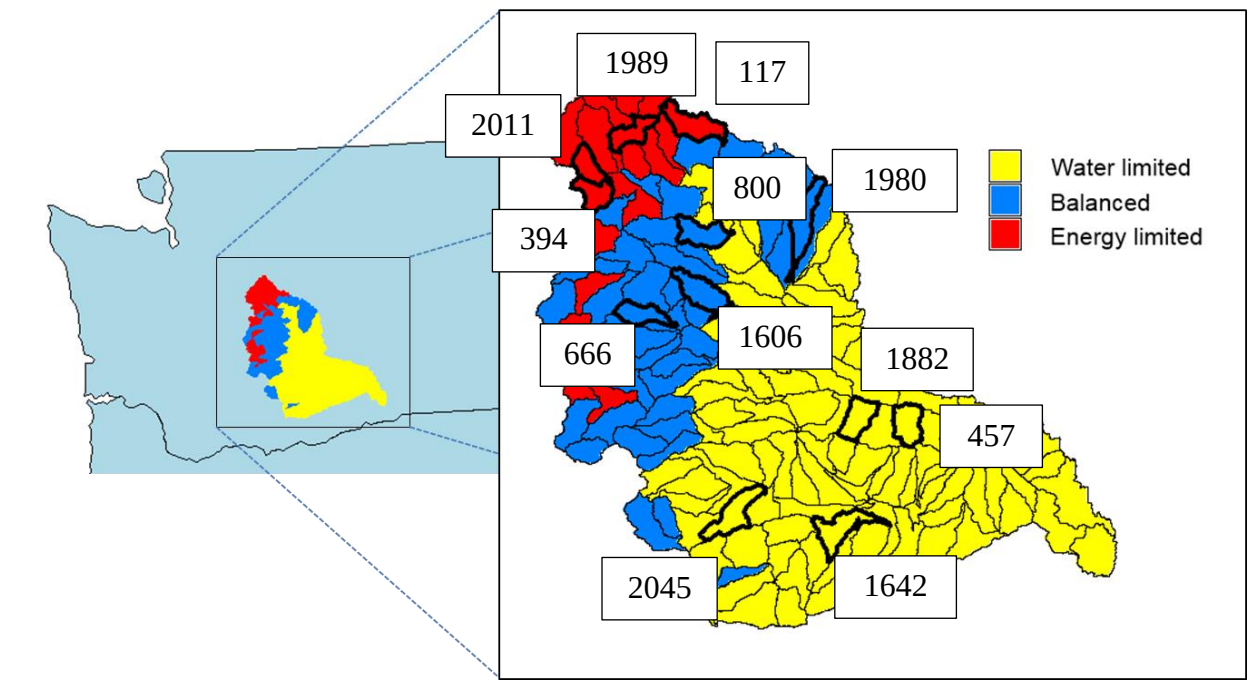


Figure 4: Yakima River Basin study area aridity index (defined as potential ET / precipitation at the annual scale). Study subbasins are shown in bold.

This study was conducted on subbasins selected from within the Yakima River basin (YRB). The YRB is located in central Washington State, and encompasses climates from cool and wet in the Cascade Mountains at its western extent, to hot and very dry on its eastern side. The cool, wet areas are dominated by evergreen forests, and the dry regions are shrub- and grassland. We used the United States Geological Survey's HUC12 watershed boundaries to delineate our study watersheds, as the size of these delineated watersheds works well with RHESSys and our research goals. Other studies (e.g. James et al. 2007) have suggested that the importance of fine-scale lateral connectivity may vary based on the climate of the study area, and particularly the

antecedent soil moisture conditions during precipitation events, so to test the effects of climate on linearity, twelve subbasins were selected at random, four from each of the aridity index (mean annual *potential ET/precipitation*, averaged over 30 years of historical data; PET is a RHESSys output, computed using the Penman-Monteith method) groups. Detailed information regarding the selected subbasins can be found in the maps and tables in Appendix B.

Table 1: Aridity Index Classification Rules

	Aridity Index
Water-Limited	< 0.8
Balanced	0.8 – 2
Energy-Limited	> 2

4.2 Hydrologic Model:

The Regional Hydro-Ecologic Simulation System (RHESSys; Tague and Band 2004) was used for all simulations. RHESSys is a distributed, hydro-ecological model originally based on CENTURY (Parton et al. 1996), a soil and vegetation model, and the arbitrarily shaped, explicitly-connected spatial unit representation of the DHSVM model (Wigmosta et al. 1994). It is a catchment-scale model, with the operational assumption that most precipitation runoff will exit the watershed within 24 hours. Spatial features are represented using a tiered system with scale-relevant parameters specified at each tier, and the model runs appropriate calculations for each level. The levels are referred to as “world”, “basin”, “hillslope”, “zone”, “patch”, and “strata”. Elements at each level contain one or more elements of the next lower level. There is only one world object which defines the simulation extent and contains one or more basins,

which encompass one or more hillslopes, and so on. Strata are unique in that they always have the same extent as their parent patch but represent different vertical layers of vegetation. This makes it possible to model multiple canopy types at different elevations from the land surface. Every patch is connected through a drainage network to other patches so water and nutrients can be routed over the land surface to the basin outlet, which allows patch-scale representation of nutrient transport dynamics. In this study, patches and strata are treated as homogeneous units; i.e., to isolate the effects of meteorological forcings and lateral connectivity in driving land surface response to climate perturbations, we are not implementing aspatial patches.

There are three primary calibration parameters in RHESSys: *k*, the saturated hydraulic conductivity; *m*, the rate at which *k* changes with depth; and soil depth. *m* and *k* control how fast water travels between the soil layers, and the depth controls the size of each of the layers. Consider a calibrated RHESSys instance. If we adjust the soil depth of this calibrated model, the resulting hydrograph will have changed, however we can compensate for the change to some extent by adjusting m and k such that water will move through the soil more quickly or slowly. Practically speaking, reducing the soil depth has a similar effect as increasing *k* and reducing *m.* In the same way, increasing the depth is like reducing *k* and increasing *m.* Because this study deals with the shape of the response curves and not their absolute magnitude, and because of the time intensity in calibrating 12 sub-basins, we instead run experiments in which we document that a robust calibration would not alter our conclusions; i.e., we show that varying soil depth multipliers (0.5x, 1.0x, and 1.5x) does not significantly alter the shapes of the response curves. Proper model calibration will still be important when running future upscaling experiments (see Section 7.1).

4.3 Data Sources:

The model input data we used are all publicly available. Digital elevation data are sourced from the National Elevation Dataset at 3 arc second resolution, courtesy of the U.S. Geological Survey. Land use and vegetation data are derived from the 2011 National Land Cover Dataset product (Homer et al. 2015). The vegetation data were reclassified as Evergreen, Deciduous, Grass, or Unvegetated to match the four default RHESSys vegetation types (see Appendix A for specific type mappings). Both evergreen and deciduous species of shrubs grow in the study area; however, they are not differentiated in our land cover dataset. Because generating a new "shrub" vegetation parameter set to account for this was beyond the scope of this study, we assumed all shrubs are evergreens. One of the primary differences between evergreen and deciduous vegetation types is that evergreens transpire year-round, whereas deciduous trees and shrubs stop transpiring in the autumn after their leaves fall. Therefore, this assumption that all shrubs are evergreens would be of significantly greater concern if we were looking at seasonal patterns. At annual scales, however, this effect should be muted.

Running RHESSys requires daily meteorological data including precipitation, minimum temperature and maximum temperature data. For historical weather information (daily maximum and minimum temperatures, relative humidity, precipitation, and wind speed), we applied the 4-km gridded product developed by Abatzoglou (2013) for the period of 1979-2012, which is a combination of *in situ* observations and regional reanalysis data. Soil classifications are made by soil texture and the data are derived primarily from the STATSGO2 database provided by the

Natural Resources Conservation Service, with additional details incorporated from Maurer et al. (2002) and Elsner et al. (2010).

The extents and types of land parameters assigned to each patch were generated by layering the each of the land parameter rasters (vegetation, soil, elevation, and land use) on top of each other and using the resolution of the highest resolution raster as the base resolution and extent for each patch, then by stepping through the other rasters within that patch's extent and selecting the type that was most prevalent. Our highest resolution raster was approximately 90m, so that is the initial size for our patches. Adjacent cells that were identical in all land parameters were collected into a single patch, meaning many patches are only 90m, but some are much larger in areas that are consistently more homogeneous.

4.4 Scenarios:

In this study we are specifically looking at heterogeneity that is driven by meteorological forcings by examining model output response curves to perturbations in these forcings. In addition, we assess the impact of lateral flow routing on the linearity of these response curves. Because both precipitation and temperature can vary over spatial scales at which we might like to upscale, we conducted experiments over each individually, for multiple minimum and maximum temperature shifts (-5 ˚C to +5 ˚C shift, in 1 ˚C increments), and for multiple rescaling of daily precipitation volumes (25% to 175%, in 25% increments). Future studies may wish to consider more robust methods of adjusting these values as minimum and maximum daily

temperatures are unlikely to rise and fall in lock step, and precipitation volume can change with the number and duration of precipitation events in addition to their intensity; however, this level of analysis is beyond the scope of this study. We ran RHESSys for both scenarios over all twelve subbasins, once with lateral routing enabled and once with lateral routing disabled to see if such lateral flows have a significant effect on the linearity of the response curves. In all cases we disabled RHESSys' dynamic vegetation module. This means there was no plant growth or death, and no changes in the carbon and nitrogen stores at each patch, but it also saved us from having to conduct a lengthy spin up process on each simulated subbasin. Transpiration continues to occur in static vegetation mode, and hydrologic stress on the plants is accounted for via the stomatal conductance processes.

Linearity was defined to be the square of the Pearson correlation coefficient (R^2) between the mean annual values of a response variable and the amount of temperature shift or precipitation rescaling. We calculated R^2 values for temperature shifts (in °C) as follows: -2 to +2; -3 to +3; -4 to +4; and -5 to +5. For example, when quantifying the linearity of the response of evapotranspiration (ET) to temperature perturbations from -2 to +2 °C, five values were used to calculate R^2: mean annual ET at temperature perturbations of -2, -1, 0, 1, and 2 °C. R^2 values for precipitation rescales (in %) were calculated as follows: 75 to 125; 50 to 150; and 25 to 175. An R^2 value near one is highly linear and one near 0 is highly nonlinear, with 0.75 defined as our linearity threshold. We also considered the slope of the regression line and the absolute magnitude of the change to identify patches that were insensitive to the changing meteorological forcings.

Each patch was assigned to a bioclimatic zone by a combination of aridity index and vegetation type. The patch-scale linearity and slope values for each subbasin were then aggregated by bioclimatic zone to generate a subbasin-scale response assessment: linear, insensitive, mixed, or nonlinear using the following rules. The first matching rule is applied. Figure 5 illustrates how this classification was performed.

- Result: Slope over the maximum perturbation range (+/-5 *C, 25% - 175% precipitation) ~= 0.

  Assessment: Insensitive

- Result: Lower 25th percentile (of all patch-specific responses) R^2 value over the maximum perturbation range is greater than the linearity threshold.

  Assessment: Linear (e.g. evergreen in Fig. 5)

- Result: Lower 25th percentile (of all patch-specific responses) R^2 value over the maximum perturbation range is less than the linearity threshold, but the lower 25th percentile R^2 value over the smallest perturbation range is above the linearity threshold.

  Assessment: Mixed (e.g. grassland in Fig. 5)

- Result: All other responses.

  Assessment: Nonlinear (e.g. deciduous in Fig. 5)

Finally, the subbasin-scale classifications were aggregated into a single recommendation per bioclimatic zone indicating that overall the response for that zone type is consistently linear, sometimes linear, or never linear.

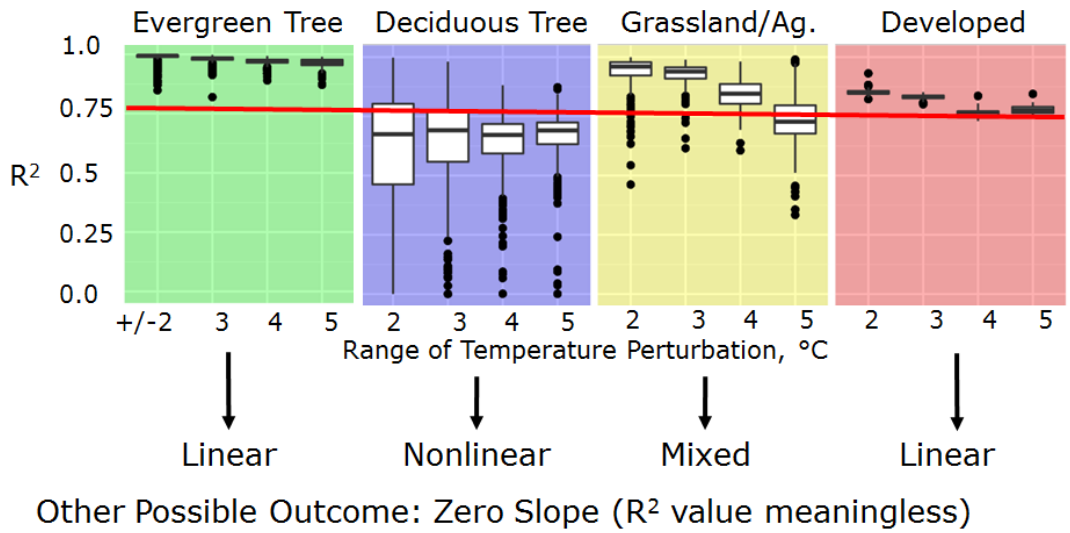


Figure 5: Example box plots showing the different linearity classes. The horizontal red line at 0.75 indicates the linearity threshold. In this example the Evergreen Tree plot's R^2 value at +/- 5 ˚C is above the linearity threshold and has a non-zero slope (not shown) so its classification is "linear." The Grassland plot falls below the threshold at +/-5 ˚C but is above it at +/-2 ˚C, so the response is "mixed." The Deciduous Tree plot's values all lie below the threshold, so it is classified as "nonlinear."

4.5 Hypothesis Formulation:

We focused on mean annual evapotranspiration (ET) and net patch outflow (Q) along the gradients of both meteorological perturbations. For the purpose of this study, net patch outflow is defined to be the difference between lateral hydrologic inputs and lateral outputs of both surface runoff and subsurface flows.

Land surface heterogeneity affects fluxes to the atmosphere which then impact atmospheric dynamics. Evapotranspiration (ET) is one of these important hydrologic fluxes, determined by climatic conditions, vegetation types, and soil moisture levels. ET is limited by the availability of water and energy. Increasing the temperature adds energy to the system, while increasing

precipitation adds available water. According to the Penman-Monteith equation, potential ET (PET) increases linearly with temperature, but actual ET can only increase up to the amount of available water. In wet basins that are energy-limited, we therefore know that PET will increase linearly with temperature, and we expect actual ET to follow PET. Dry basins, on the other hand, are water-limited and should show minimal change in ET as the temperature changes. As precipitation volume changes, we expect to the see the opposite behavior: linear ET increases with increasing precipitation in dry basins, and minimal change in wet basins because actual ET is already close to PET. In regions with balanced water-energy terms, the results are harder to predict. Lateral flows could have a strong impact in these regions as well, because excess water in some patches could be routed to and evaporated or transpired by connected patches with a water deficit.

Mean annual net patch outflow (Q) is a measure of the volume of water routed between adjacent patches and ultimately controls how much water is available at a patch's location. Water is routed into patches from adjacent upslope patches via overland and subsurface flows (Q_{in}). After all the other hydrologic processes are accounted for, any excess water is routed to downslope patches (Q_{out}). The net volume (Q) is $Q_{in} – Q_{out}$. In RHESSys such lateral flows are accounted for after ET, so in very dry basins net Q will be zero because overland flow will be very low, and any subsurface flow is likely below the root zone and will continue to be routed downslope. If it is non-zero, with constant precipitation and in the absence of competing effects, Q should respond to temperature perturbations with the same magnitude but opposite sign as ET.

We predict that as T is perturbed, Q will change linearly for wet regions and not change in dry ones. When P varies in dry basins, where all or most of the available water is initially lost as ET, we do not expect to see a linear response. In these cases, the net lateral flow will increase (from zero) only when there is enough precipitation to result in runoff or additional water percolating to the subsurface layers. Less precipitation will have no effect on the net lateral flow. In wet areas Q should change linearly with P, and in balanced basins we expect the results to be mixed.

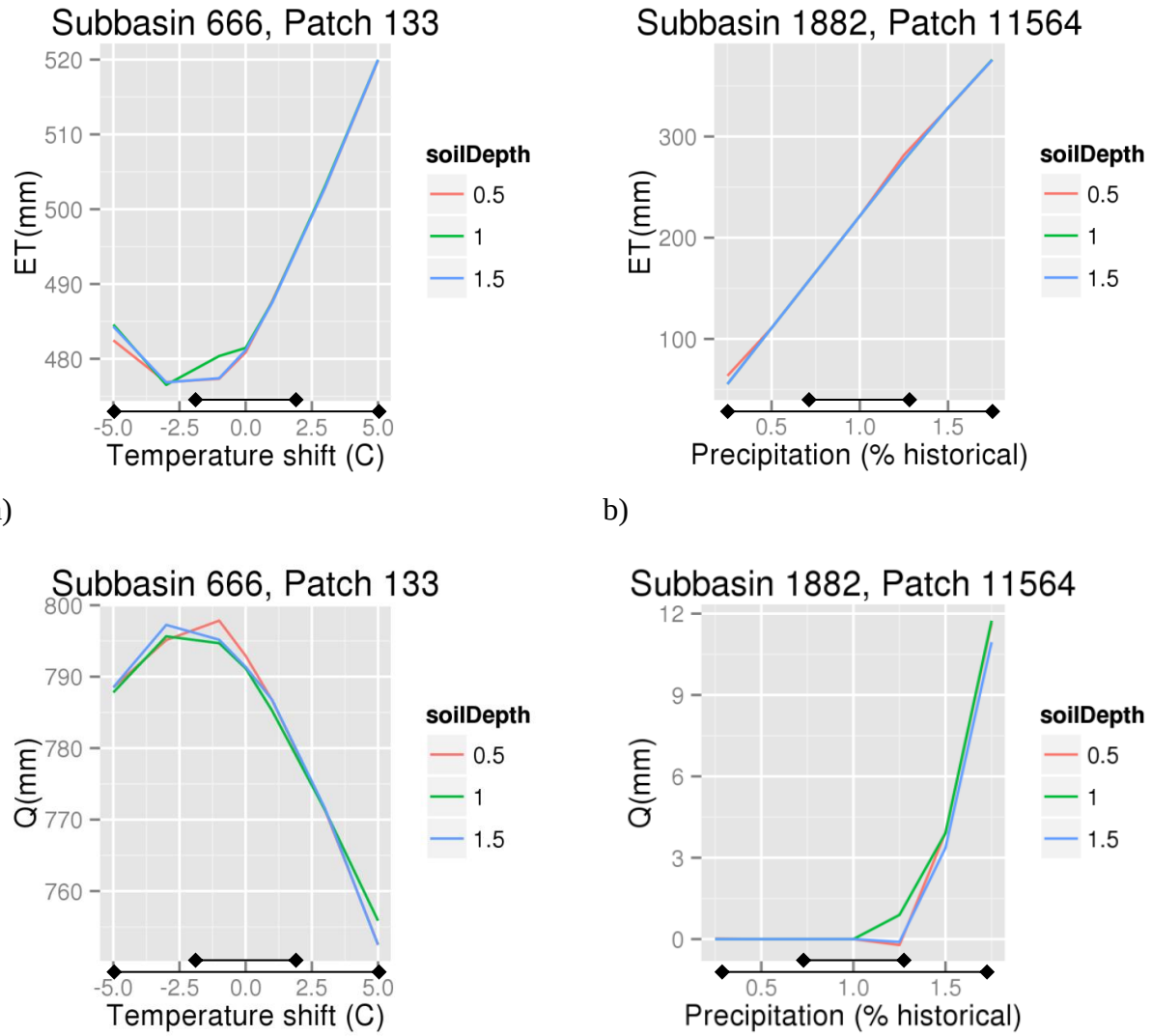
Table 2: Summary of linearity predictions by biome and climate driver. 1=linear, 2=no change, 3=unclear trend, 4=nonlinear

	(PET/P)		ET	Q
Water-limited	> 2	Temp	2	2
		Precip	1	4
Balanced	0.8 – 2	Temp	3	3
		Precip	3	3
Energy-limited	< 0.8	Temp	1	1
		Precip	2	1

5 Results:

5.1 Preliminary Checks:

By visual inspection of the response curves of several randomly selected patches (Figure 6), we see that 0.5x, 1.0x, and 1.5x soil depth multipliers typically do not change the shape of the response curves. Even the magnitude of the change is small. However, because hydrologic systems are complex, changing the soil depth does occasionally result in unexpected response curve patterns, for example the patch shown in Figure 7. Other than the outliers, the overall distribution of R^2 values stays nearly constant when changing the soil depth, as shown for one subbasin in Figure 8. We are confident that our assumption of constant response curve shape without model calibration is acceptable.



a) b)

c) d)

Figure 6: Typical patch response curves using three different calibration soil depth multipliers. Plots a) and b) are typical response curves for ET, while c) and d) show normal results for Q. Patches were selected at random. The black bars at the bottom of each plot represent the largest and smallest ranges used to calculate the regression R^2 scores.

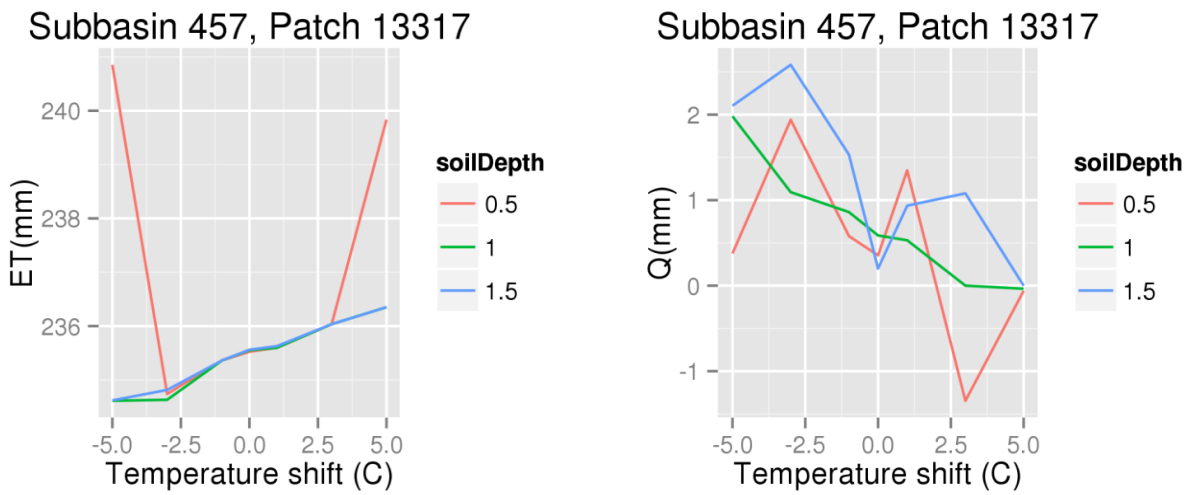


Figure 7: Atypical patch response curves. Note the small variation in absolute magnitude.

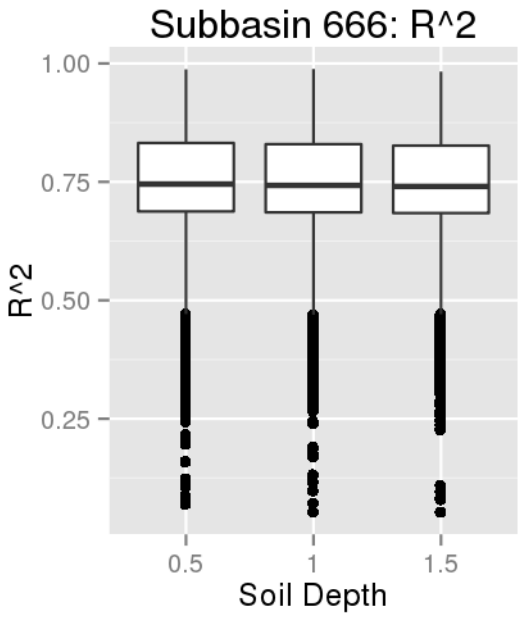


Figure 8: The distribution of majority of linearities in Subbasin 666 does not change with soil depth, though there is some variation in the outliers.

5.2 ET as a Function of T:

Most patches in dry bioclimate zones showed no response to changes in temperature. Some of the evergreen patches had a slight, linear response, while the pattern in deciduous patches was unclear. In the balanced aridity index zones, evergreens were consistently linear while everything else was unclear. In the energy-limited zones, evergreens have a linear response, deciduous returns the only nonlinear result in the entire study, grasses are unclear, and unvegetated areas show linear change. Lateral routing had an effect on the deciduous zones, though the effect was mixed with a positive effect on the linearity of one of the dry subbasins and a negative effect on the linearity of both of the wet subbasins.

Table 3: ET(T) summary of linearity results. 1=linear, 2=no response, 3=sometimes linear, 4=nonlinear.

Subbasin ID		Aridity	Evergreen				Deciduous				Grass				Unvegetated			
			R_on		R_off		R_on		R_off		R_on		R_off		R_on		R_off	
457	1882	water-limited	1	2	1	2	-	-	-	-	2	2	2	2	2	2	2	2
1642	2045		2	1	2	1	1	3	3	3	2	2	2	2	-	2	-	2
800	666	balanced	1	1	1	1	-	3	-	3	3	3	3	3	3	3	3	3
1606	1980		1	1	1	1	-	-	-	-	3	3	3	1	3	3	3	3
394	117	energy-limited	1	3	1	3	-	-	-	-	3	3	3	3	1	1	1	1
2011	1989		1	1	1	1	4	3	3	1	3	3	3	3	1	1	1	1

5.3 ET as a Function of P:

Scaling precipitation had a linear effect on ET on all of the vegetated patches, regardless of the aridity class or vegetation type. In the unvegetated patches it had no effect in the dry subbasin, mixed impact in the balanced ones, and mostly linear impact in the wet subbasins. Explicit

lateral routing had no impact on the linearity of these results.

Table 4: ET(P) summary of linearity results. 1=linear, 2=no response, 3=sometimes linear, 4=nonlinear.

Subbasin ID		Aridity	Evergreen				Deciduous				Grass				Unvegetated			
			R_on		R_off		R_on		R_off		R_on		R_off		R_on		R_off	
457	1882	water-limited	1	1	1	1	-	-	-	-	1	1	1	1	2	2	2	2
1642	2045		1	1	1	1	1	1	1	1	1	1	1	1	-	2	-	2
800	666	Balanced	1	1	1	1	-	1	-	1	1	1	1	1	2	1	2	1
1606	1980		1	1	1	1	-	-	-	-	1	1	1	1	3	3	3	3
394	117	energy-limited	1	1	1	1	-	-	-	-	1	1	1	1	1	3	1	3
2011	1989		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

5.4 Q as a function of T:

The linearity of the Q response overall is much more complicated than ET. In dry evergreen deciduous patches, Q is mixed if lateral routing is on, but linear or unchanging if not. Dry grass and unvegetated patches show no sensitivity to T. Evergreen patches in the wet and balanced subbasins are mostly linear if lateral routing is off, but mixed if it is enabled. Balanced deciduous, grass, and unvegetated patches are all unclear. Wet deciduous patches have a more extreme spread, having subbasins classifications showing nonlinear and uncertain if routing is one, and nonlinear and linear when lateral routing is disabled.

Table 5: Q(T) summary of linearity results. 1=linear, 2=no response, 3=sometimes linear, 4=nonlinear.

Subbasin ID		Aridity	Evergreen				Deciduous				Grass				Unvegetated			
			R_on		R_off		R_on		R_off		R_on		R_off		R_on		R_off	
457	1882	water-limited	3	2	1	2	-	-	-	-	2	2	2	2	2	2	2	2
1642	2045		3	3	2	1	3	3	1	3	2	2	2	2	-	3	-	2
800	666	Balanced	1	3	1	1	-	3	-	3	3	3	3	3	3	3	3	3
1606	1980		1	3	1	1	-	-	-	-	3	3	3	1	3	3	3	3
394	117	energy-limited	1	3	1	3	-	-	-	-	3	3	3	3	3	3	1	3
2011	1989		1	3	1	1	4	3	4	1	3	3	3	3	3	3	3	3

5.5 Q as a function of P:

Only the dry subbasins showed any mixed linearity scores. Evergreens were split evenly between linear and unclear results, deciduous and unvegetated biomes were linear, and grass was not consistently uncertain. All biomes in the balanced and energy-limited subbasins responded linearly. Explicit lateral routing did not affect the linearity classification results.

Table 6: Q(P) summary of linearity results. 1=linear, 2=no response, 3=sometimes linear, 4=nonlinear.

Subbasin ID		Aridity	Evergreen				Deciduous				Grass				Unvegetated			
			R_on		R_off		R_on		R_off		R_on		R_off		R_on		R_off	
457	1882	water-limited	3	3	3	3	-	-	-	-	3	3	3	3	1	1	1	1
1642	2045		1	1	1	1	1	1	1	1	3	3	3	3	-	1	-	1
800	666	Balanced	1	1	1	1	-	1	-	1	1	1	1	1	1	1	1	1
1606	1980		1	1	1	1	-	-	-	-	1	3	1	3	1	1	1	1
394	117	energy-limited	1	1	1	1	-	-	-	-	1	1	1	1	1	1	1	1
2011	1989		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

6 DISCUSSION:

The final linearity/upscalability recommendations are based on the widest range of data points available in our simulations, and are summarized in Tables 6 and 7. Bioclimatic zones labeled "Y" satisfied the linearity criteria over the widest range of values tested (-5 to +5 °C temperature shift, 25 to 175% precipitation rescaling). For the most part, the linearity results are as hypothesized. Many of the "Maybes," such as ET when varying temperatures in dry, deciduous basins, are linear over smaller bracket ranges, but become less linear as the range of included perturbations becomes broader. It should be emphasized that whether or not these "Maybes" can be upscaled depends upon the expected range of climatic conditions. If upscaling over a region

with a small range in Temperature or Precipitation, it is more likely to be nearly linear, any aggregation error will remain low, and upscaling may be possible. As patch size, and thus the possible range of forcing values, increases the more the linearity or nonlinearity of the response curves will matter.

Table 7: Summary of results from representative subbasins: Potential for ET upscaling (Yes/Maybe/No). Yellow indicates one step deviation from predictions, blue indicates two steps of deviation.

Aridity		Predicted	Evergreen		Deciduous		Grass		Unvegetated	
			R_on	R_off	R_on	R_off	R_on	R_off	R_on	R_off
Water-limited	Temp	Y	Y	Y	M	M	Y	Y	Y	Y
	Precip	Y	Y	Y	Y	Y	Y	Y	Y	Y
Balanced	Temp	M	Y	Y	M	M	M	M	M	M
	Precip	M	Y	Y	Y	Y	Y	Y	M	M
Energy-limited	Temp	Y	Y	Y	N	M	M	M	Y	Y
	Precip	Y	Y	Y	Y	Y	Y	Y	Y	Y

PET changes linearly with temperature, so we expect actual ET to change somewhat linearly as well. The model results demonstrate this behavior except in the deciduous zones and areas of water-energy limitation balance. Few patches were consistently nonlinear across all tested ranges of perturbations. The “Maybe” linearity scores in the balanced aridity index subbasins were linear for small perturbation ranges, but became less linear as the ranges increased. Lateral routing affected the linearity predictions of only the wet, deciduous patches, but these are uncommon in the study area. Much of the YRB is a candidate for upscaling ET over temperature gradients, particularly in the lower and higher elevation regions. The complex pattern seen in the upper reaches of the YRB prediction maps for ET as a function of temperature (Figure 9) is due to the “Maybe” result of the wet grasslands found in the drainages at higher elevations.

ET as a function of precipitation is surprisingly linear in all cases except for unvegetated patches in balanced subbasins. We expected it to be so in the wet and dry basins, but in the balanced basins we did not anticipate this level of linearity because competing effects are amplified when the aridity index is near 1.

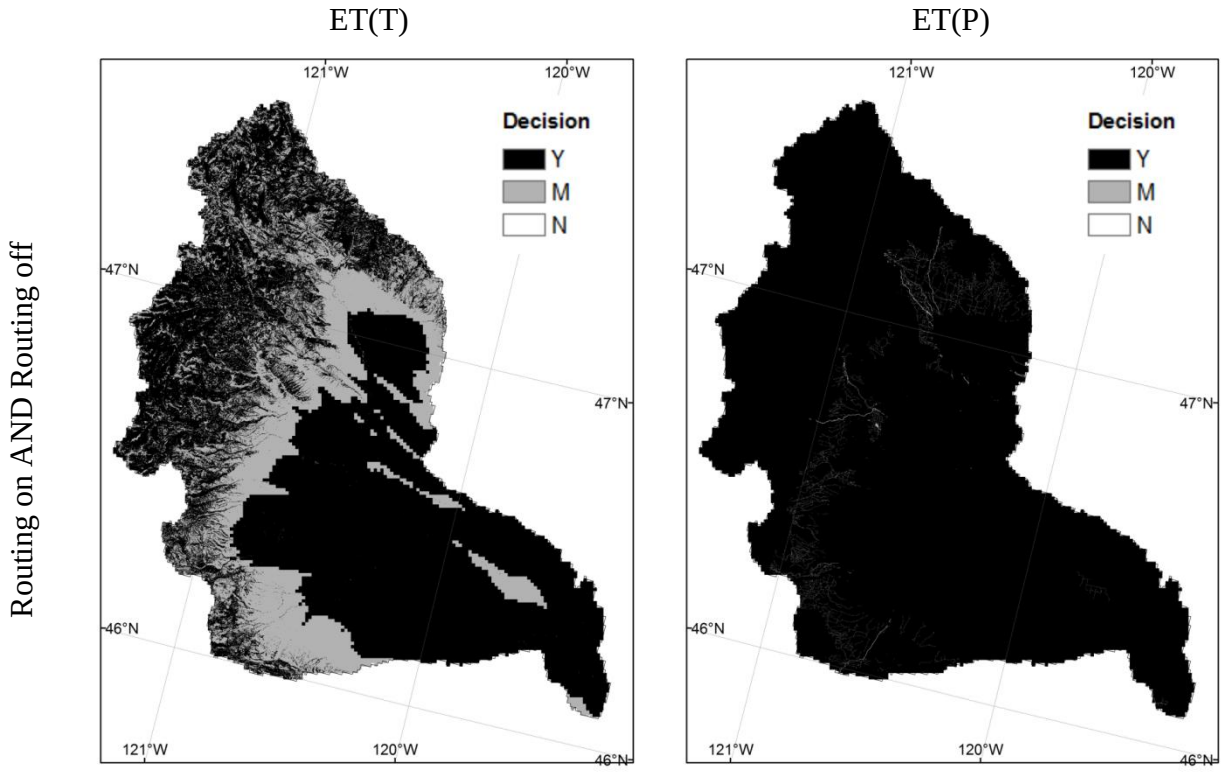


Figure 9: Linearity predictions for ET as a function of T applied over the whole Yakima River basin. There is no difference between results with lateral routing on and lateral routing disabled except for the deciduous case which is very uncommon in this so only one image is included.

Figure 10: Linearity predictions of ET as a function of precipitation perturbation over the Yakima River basin. There is no difference between linearity predictions with lateral routing on and lateral routing disabled. The only areas not linear are the unvegetated patches in balanced subbasins.

Table 8: Summary of results from representative subbasins: Potential for Q upscaling (Yes/Maybe/No). Yellow indicates one step deviation from predictions, blue indicates 2 steps.

Aridity		Predicted	Evergreen		Deciduous		Grass		Unvegetated	
			R_on	R_off	R_on	R_off	R_on	R_off	R_on	R_off
Water-	Temp	Y	M	Y	M	M	Y	Y	Y	Y
limited	Precip	N	M	M	Y	Y	M	M	Y	Y
Balanced	Temp	M	M	Y	M	M	M	M	M	M
	Precip	M	Y	Y	Y	Y	Y	Y	Y	Y
Energy-	Temp	Y	M	Y	N	M	M	M	M	M
limited	Precip	Y	Y	Y	Y	Y	Y	Y	Y	Y

The linearity of Q as a function of temperature perturbation is not as consistent as the ET response. In very dry subbasins, there is no net patch outflow to begin with, and increasing the temperature will not change this. This has a strong impact on the linearity of the response curves depending on how close to zero the net outflow is to begin with. If the basin is exceedingly dry, the resulting curve is flat because most of the Q values are at or near zero. If the basin is exceptionally wet, Q will drop with temperature inversely proportionally to ET. With lateral flow routing enabled, patches with excess water can contribute moisture to dry neighboring patches resulting in more complicated and less linear flow patterns.

The only patches where Q was not consistently linear with precipitation perturbations was in the evergreen and deciduous patches of the dry basins. In wet areas, Q is linear with precipitation change, which is as expected, because adding or removing water from a region that is already energy limited should result in more or less runoff. In the driest basins there is minimal patch outflow under normal conditions so making these patches drier has no effect; however, there is some point at which the volume of precipitation will overcome this deficit and result in positive net patch outflow. The shape of this curve may or may not be straight enough to satisfy the linearity requirements depending on how many of the data points are at or near zero and how

rapidly the lateral outflow increases (see Figure 11). Explicit lateral flow routing did not change the net patch lateral flow when precipitation was perturbed.

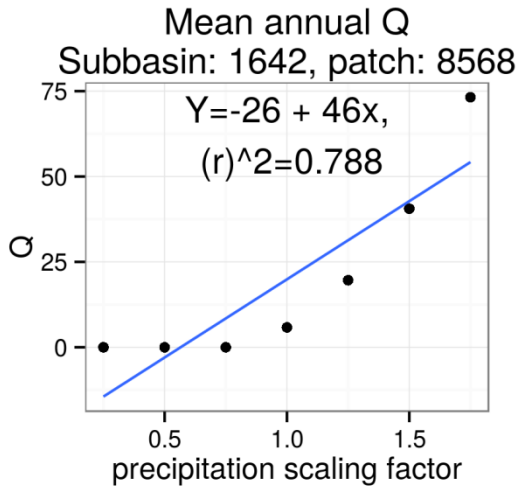


Figure 11: Q as a function of precipitation for a patch in a dry basin. There is no net patch outflow when precipitation is reduced, but the R^2 value can still be above the 0.75 linearity threshold.

Q(T)

Q(P)

Routing Off

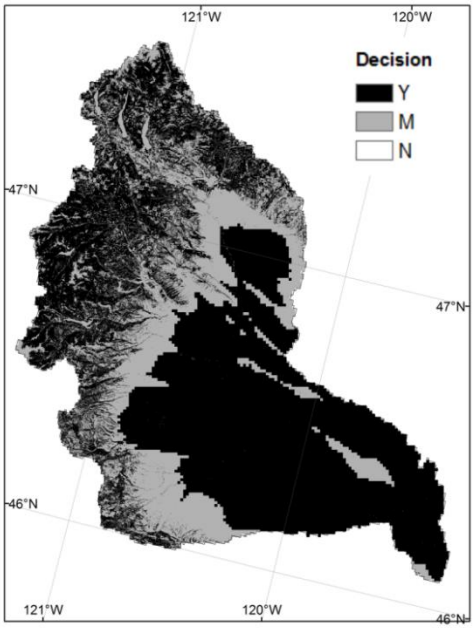


Figure 12: Linearity predictions of Q as a function of temperature, explicit lateral routing disabled. Note that this map is nearly identical to the map of linearity predictions of ET(T).

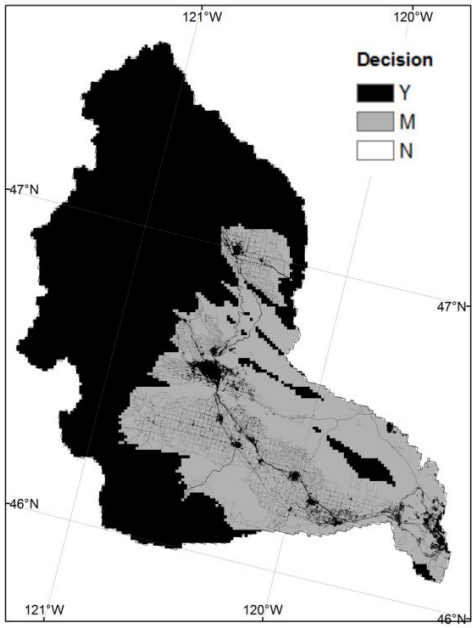


Figure 13: Linearity predictions of Q as a function of precipitation. There is no difference between linearity predictions with lateral routing on and lateral routing disabled.

Routing On

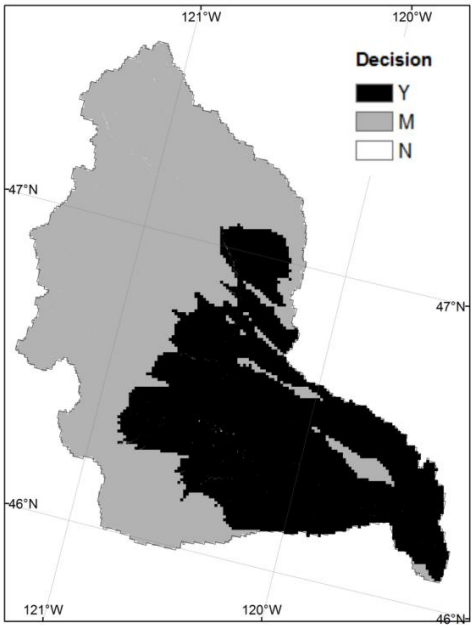


Figure 14: Linearity predictions of Q as a function of temperature, explicit lateral routing enabled. The high elevation linearities are noticeably reduced from

<blank>

previous prediction maps.

In light of these results, we return to the original research questions. First, can we aggregate over areas of large climate gradients? The answer to that depends on the size of the gradient, the process being studied, and the bioclimatic conditions. Except in cases where the response curve is perfectly linear, aggregating over small gradients will result in less aggregation error than over large ones. This is most clearly demonstrated by the large number of "Maybe" results in the upscaling recommendations, which by definition mean the results were linear over a small gradient but not over a larger one. The more linear is a function's response curve, the larger the gradient that can be safely aggregated. Because many meteorological forcings vary with elevation, this means the potential size of the upscaled patches is likely to be larger in regions of low topographical relief than in complex, mountainous terrain. (However, with aspatial patches implemented, the dependency of temperature forcings with elevation can be somewhat captured by lapsing temperature with aspatial elements for elevation.) The bioclimatic zone also has an effect on upscalability, with certain vegetation classes, e.g. evergreens, being more likely to have a linear response than others, such as unvegetated areas.

Does accounting for fine-scale lateral redistribution impact the watershed-scale response over climate gradients? Somewhat, but this depends particularly on the process being studied. In our experiments ET showed no sensitivity to fine-scale lateral routing, but Q did, especially in the more topographically complex northern areas of the Yakima River basin. With lateral redistribution disabled, excess water is routed from each patch directly out of the basin. However, enabling explicit lateral routing means the excess water is routed according to the complexity of the terrain. Down slope patches in mountainous terrain, like we see in the wet

study subbasins, are going to be heavily influenced by upslope drainage.

Finally, how can this information inform hydrologic model upscaling? Bloschl and Sivapalan (1995) lay out a framework that is useful here. They describe three different types of scales in hydrological modeling: process, observational, and model. Process scale is the natural scale at which a process occurs, for example the process scale of runoff during a storm event would include the storm's spatial extent (on the order of kilometers), and its temporal extent (minutes to days). Observational scale refers to measurement density. Point measurements tend to be dense temporally but sparse spatially, like precipitation measurements taken using rain gages. Model scale is a function of both the process and observational scales, and computational considerations. This framework provides a way for modelers to identify the process scale given the research question under investigation. This process scale may be driven by heterogeneity in land parameterization, meteorological forcings, or governed by lateral processes that redistribute moisture and nutrients. Ideally, we want to match the model scale to that of the process scale. By understanding these process scales, a flexible-scale hydrologic model can be implemented to capture each of these scales in the most appropriate way for each specific location. For example, if ET response to precipitation perturbations is linear for small perturbation ranges, then this perturbation range may dictate the scale of the model resolution in that area and land parameter heterogeneity can be captured through aspatial patches. However, if lateral flow causes a response to become nonlinear, then we know that explicitly-located patches at the finer-resolution are necessary at this location. Finally, the assumption behind matching model scale to that of process scale is that the scale of the observational data used as model input is fine enough to support the finest model scale that is required. In many cases, observation scale will be the

limiting factor.

It is also important to note that different process models are relevant only at a specific range of scales. Modeling pore-scale ground water flow requires solving the Navier-Stokes equations, but if you increase the size of the study area enough, Darcy's law becomes the more appropriate computational method. The type of "upscaling" we are discussing in this paper assumes that we don't make one of these scale jumps between process models. Even if every patch in several adjacent watersheds had a linear response in some response variable, we would not be able to combine these watersheds if they violate the foundational assumption in RHESSys that water is capable of draining from the basin in 24 hours, so in a sense this is the upper limit on how far RHESSys can be upscaled, because any larger and it would require a new process model.

7 Future directions and Study Limitations:

7.1 Future Directions

The obvious next step in this study is to apply and test these recommendations by generating patches that are large enough to encompass climate gradients in the areas we identified as upscalable. This is important because the recommendations still need to be verified by comparing the model performances of a baseline and upscaled watershed, in terms of both hydrologic simulation accuracy and model computational time. In this case a full calibration will be required because the magnitude of the response curves matter.

To enhance the usefulness of these results, it will also be important to completely characterize the meteorological ranges of linear response. The large number of "maybes" in our upscaling recommendations means that the results were linear over the interval of +/-2 ˚C but not the larger

interval of +/-5 C, and no thorough analysis of intervals between +/-2 ˚C and +/-5 ˚C has been completed. A more complete characterization will help identify grid sizes closer to the optimal process scale. If we were to consider sub-annual time-scales for shorter simulations, we would additionally have to calculate the linearities on a per-season basis because some processes are very season-dependent, such as ET (e.g. because transpiration ceases when deciduous plants lose their leaves as well as other temperature- and day-length- driven factors).

We initialized the model's carbon and nitrogen stores from satellite-derived leaf area index and used a static vegetation model to avoid much of the computationally expensive spin up and to simplify the final analysis (i.e., so we would have fewer processes impacting our simulated results.) Running in static mode meant that there was no change from year to year in the amount of available carbon and nitrogen and no change in the vegetation types. Under prolonged changes to temperature or precipitation the vegetation dynamics could change, and because ET is strongly influenced by these dynamics, this could change the linearity of the response curves and thus the upscaling recommendations. Even with static vegetation mode enabled, climate adaptation and mitigation studies can be done by manually altering the vegetation types in each patch. Additionally, for future study areas with a high areal percentage of vegetation that does not fit clearly into the basic vegetation classes distributed with RHESSys, creating new parameter sets for the classes that don't fit ideally would result in more accurate overall results and would also allow us to dig deeper and test if/how seasonal effects affect linearity scores.

The full analysis we've conducted is based on the linearity of the change of the means of each response variable. It is interesting to consider the linearity of the annual extremes (e.g. 90th

percentile) as well, because many planning and policy issues are driven by the worst-case scenario, not the average case. Consider an assessment of flood risk. If we want to apply this aggregation method in studies assessing peak stream flow under differing conditions, we would need to be sure that our aggregation technique does not unduly dampen or enhance the peak signal.

7.2 Study Limitations:

This study describes a method for identifying areas where it may be possible to upscale a hydrologic model by aggregating over regions having linear ET and Q curves with changes in temperature and precipitation, however, the linearity threshold itself needs to be calibrated. We selected 0.75 based on what “looked” linear in a graph, but the actual aggregation error introduced by using different thresholds still needs to be quantified and optimized.

8 Conclusion:

The grand challenges we face represent the gaps in our scientific knowledge that offer the greatest potential for scientific, social, cultural and environmental impact. Hyper-resolution modeling is one of hydrology's grand challenges and it promises to change the way we understand water in the environment, but it is not a magic bullet for solving all of the issues hydrologic modelers face. Hyper-resolution modeling requires exponentially more storage space and computational time, and hyper-resolution data collection will require major advances in observational techniques and efforts. However, if we can selectively identify those regions where hyper-resolution will be the most useful we will be able to focus our resources on those areas and progress faster in addressing scientific questions that rely on these models.

Distributed models are in a great position to utilize hyper-resolution data because of the flexibility of the scale at which it can be used, and because they are capable of fully representing multiple types of spatial heterogeneity. Future improvements to RHESSys will give it the best of both the spatially explicit and the aspatial statistical representations of heterogeneity, which means it can be used efficiently by varying the data resolution depending on the needs of the research question, mixing hyper- and standard-resolution data into a single model configuration.

Using the RHESSys model, we have established a method for identifying areas where upscaling may be possible for specific hydrologic response variables while minimizing aggregation error by looking at the linearity of response curves over meteorological forcings. We showed that in

almost all of the tested scenarios, ET and Q response curves were linear over a small range of temperature and precipitation values, while only some maintained that linearity over a wider range. How much upscaling is possible depends on the linearity of the response curve and the range of expected climatic conditions. Our results show that upscaling should be possible in some areas, but they can also be used to inform how and where different types of spatial heterogeneity should be captured across our model domain. More work remains to be done, particularly in the validation of these results. An upscaled model needs to be compared against a baseline, fine-scale one to test the recommendation and verify that this process of identifying upscalable regions works and quantify how much error it causes. Recommendations for a variety of meteorological forcing ranges would provide the most flexibility and accuracy when determining the appropriate model resolution for a given bioclimatic zone in a RHESSys model, so more work remains to be done to fully characterize the response curve linearities.

We have contributed to the hyper-resolution grand challenge discussion by showing that, over some biomes and climates, upscaling over meteorological drivers should be possible with minimal aggregation error, and conversely, by identifying those regions where increasing the resolution of our data and models will not improve our modeling results. This frees up limited resources to be spent in those regions that benefit from it the most.

In the future, knowing the answers to scientific questions that can be addressed by land surface models, such as sustainable management of our fresh water supplies and climate feedback effects, will become more important than ever. The better we understand these issues of scale

and connectivity, the better our models will become, and the better it will be to address these questions and preserve our water for future generations.

This project was supported in part by the BioEarth project, funded by the Department of Agriculture, National Institute of Food and Agriculture grant number 2011-67003-30346, and by the Wisdm project, funded by the Department of Agriculture, National Institute of Food and Agriculture Grant Number 2012-67003-19805 through the Water Sustainability and Climate program, and the State of Washington Water Research Center Grant Number G11AP20113 through the United States Geological Survey 104(b) funding to the Water Resource Research Institutes.

Works Cited:

Abatzoglou J. T. " Development of gridded surface meteorological data for ecological applications and modeling " *International Journal of Climatology*. (2011): 10.1002/joc.3413.

Beven, Keith J., and Hannah L. Cloke. "Comment on 'Hyperresolution Global Land Surface Modeling: Meeting a Grand Challenge for Monitoring Earth's Terrestrial Water' by Eric F. Wood et Al." *Water Resources Research* 48, no. 1 (2012): W01801. doi:10.1029/2011WR010982.

Bierkens, Marc F. P., Victoria A. Bell, Peter Burek, Nathaniel Chaney, Laura E. Condon, Cédric H. David, Ad de Roo, et al. "Hyper-Resolution Global Hydrological Modelling: What Is Next?" *Hydrological Processes* 29, no. 2 (January 15, 2015): 310–20. doi:10.1002/hyp.10391.

Callow, J. N., and K. R. J. Smettem. "The Effect of Farm Dams and Constructed Banks on Hydrologic Connectivity and Runoff Estimation in Agricultural Landscapes." *Environmental Modelling & Software* 24, no. 8 (August 2009): 959–68. doi:10.1016/j.envsoft.2009.02.003.

Chesson, Peter. "Scale Transition Theory: Its Aims, Motivations and Predictions." *Ecological Complexity* 10 (June 2012): 52–68. doi:10.1016/j.ecocom.2011.11.002.

Clark, Martyn P., Ying Fan, David M. Lawrence, Jennifer C. Adam, Diogo Bolster, David J. Gochis, Richard P. Hooper, et al. "Improving the Representation of Hydrologic Processes in Earth System Models." *Water Resources Research* 51, no. 8 (August 1, 2015): 5929–56. doi:10.1002/2015WR017096.

Clark, M. P., B. Nijssen, J. Lundquist, D. Kavetski, D. Rupp, R. Woods, E. Gutmann, A. Wood, L. Brekke, J. Arnold, D. Gochis, and R. Rasmussen. "A unified approach to process-based hydrologic modeling. Part 1: Modeling concept." *Water Resources Research* (2015a). doi: 10.1002/2015WR017198.

Clark, M. P., B. Nijssen, J. Lundquist, D. Kavetski, D. Rupp, R. Woods, E. Gutmann, A. Wood, D. Gochis, R. Rasmussen, D. Tarboton, V. Mahat, G. Flerchinger, and D. Marks. "A unified approach for process-based hydrologic modeling: Part 2. Model implementation and example applications." *Water Resources Research* (2015b). doi: 10.1002/2015WR017200

Daly, Christopher, Michael Halbleib, Joseph I. Smith, Wayne P. Gibson, Matthew K. Doggett, George H. Taylor, Jan Curtis, and Phillip P. Pasteris. "Physiographically Sensitive Mapping of Climatological Temperature and Precipitation across the Conterminous United States." *International Journal of Climatology* 28, no. 15 (December 1, 2008): 2031–64. doi:10.1002/joc.1688.

Eagleson, Peter S. "The Evolution of Modern Hydrology (from Watershed to Continent in 30 Years)." *Advances in Water Resources* 17, no. 1–2 (January 1994): 3–18. doi:10.1016/0309-1708(94)90019-1.

Elsner, M. M., L. Cuo, N. Voisin, J. S. Deems, A. F. Hamlet, J. A. Vano, K. E. B. Mickelson, S.-Y. Lee, and D. P. Lettenmaier (2010), Implications of 21st century climate change for the hydrology of Washington State, *Clim. Chang.*, 102(1-2), 225–260, doi:10.1007/s10584-010-9855-0.

James, April L., and Nigel T. Roulet. "Investigating Hydrologic Connectivity and Its Association with Threshold Change in Runoff Response in a Temperate Forested Watershed." *Hydrological Processes* 21, no. 25 (December 1, 2007): 3391–3408. doi:10.1002/hyp.6554.

Khan, Urooj, Narendra Kumar Tuteja, Hoori Ajami, and Ashish Sharma. "An Equivalent Cross-Sectional Basis for Semidistributed Hydrological Modeling." *Water Resources Research* 50, no. 5 (2014): 4395–4415. doi:10.1002/2013WR014741.

Kidd, C., P. Bauer, J. Turk, G. J. Huffman, R. Joyce, K.-L. Hsu, and D. Braithwaite. "Intercomparison of High-Resolution Precipitation Products over Northwest Europe." *Journal of Hydrometeorology* 13, no. 1 (August 26, 2011): 67–83. doi:10.1175/JHM-D-11-042.1.

Kumar, R., B. Livneh, and L. Samaniego. "Toward Computationally Efficient Large-Scale Hydrologic Predictions with a Multiscale Regionalization Scheme." *Water Resources Research* 49, no. 9 (September 2013): 5700–5714. doi:10.1002/wrcr.20431.

Kure, S., M. L. Kavvas, N. Ohara, and S. Jang. "Upscaling of Coupled Land Surface Process Modeling for Heterogeneous Landscapes: Stochastic Approach." *Journal of Hydrologic Engineering* 16, no. 12 (December 2011): 1017–29. doi:10.1061/(ASCE)HE.1943-5584.0000163.

Liang, X. "A Two-Layer Variable Infiltration Capacity Land Surface Representation for General Circulation Models." *Water Resour. Series*, TR140 (May 1994), 208 pp., Univ. of Washington, Seattle.

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen: "A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States", *J. Climate* 15 (2002): 3237-3251.

Minder, Justin R., Philip W. Mote, and Jessica D. Lundquist. "Surface Temperature Lapse Rates over Complex Terrain: Lessons from the Cascade Mountains." *Journal of Geophysical Research: Atmospheres* 115, no. D14 (July 27, 2010): D14122. doi:10.1029/2009JD013493.

Melbourne, Brett A., and Peter Chesson. "Scaling up Population Dynamics: Integrating Theory and Data." *Oecologia* 145, no. 2 (September 1, 2005): 178–86. doi:10.1007/s00442-005-0058-8.

Rastetter, Edward B., King Anthony W., Cosby Bernard J., Hornberger George M., O'Neill Robert V., and Hobbie John E. "Aggregating Fine-Scale Ecological Knowledge to Model Coarser-Scale Attributes of Ecosystems." *Ecological Applications* 2, no. 1 (1992): 55-70.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at http://websoilsurvey.nrcs.usda.gov/. Accessed 2015.

Stieglitz, Marc, Jeff Shaman, James McNamara, Victor Engel, Jamie Shanley, and George W. Kling. "An Approach to Understanding Hydrologic Connectivity on the Hillslope and the Implications for Nutrient Transport." *Global Biogeochemical Cycles* 17, no. 4 (December 1, 2003): 1105. doi:10.1029/2003GB002041.

Tague, C. L., and L. E. Band. "RHESSys: Regional Hydro-Ecologic Simulation System-An Object-Oriented Approach to Spatially Distributed Modeling of Carbon, Water, and Nutrient Cycling." *Earth Interactions* 8 (2004): 19.

Tague, Christina, and Hui Peng. "The Sensitivity of Forest Water Use to the Timing of Precipitation and Snowmelt Recharge in the California Sierra: Implications for a Warming Climate." *Journal of Geophysical Research: Biogeosciences* 118, no. 2 (2013): 875–87. doi:10.1002/jgrg.20073.

Tesfa, T. K., H.-Y. Li, L. R. Leung, M. Huang, Y. Ke, Y. Sun, and Y. Liu. "A Subbasin-Based Framework to Represent Land Surface Processes in an Earth System Model." *Geosci. Model Dev.* 7, no. 3 (May 20, 2014): 947–63. doi:10.5194/gmd-7-947-2014.

Thornton, P. E., and S. W. Running. "An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation." *Agric. For. Meteor.*, 93 (1999): 211–228.

Wood, Eric F., Joshua K. Roundy, Tara J. Troy, L. P. H. van Beek, Marc F. P. Bierkens, Eleanor Blyth, Ad de Roo, et al. "Hyperresolution Global Land Surface Modeling: Meeting a Grand Challenge for Monitoring Earth's Terrestrial Water." *Water Resources Research* 47, no. 5 (2011): W05301. doi:10.1029/2010WR010090.

Appendix A: Model configurations and data reclassifications

National Land Cover Database (NLCD) classification to RHESSys category mappings:

Evergreen forest, Shrub/scrub = Evergreen

Deciduous forest, mixed forest, woody wetlands = Deciduous

Grassland/herbaceous, sedge/herbaceous, pasture/hay, emergent herbaceous wetland = Grass

All others: Nonveg

Soil label categories (based on NRCS soil texture triangle):

1 Clay

2 Silty Clay

3 Silty Clay Loam

4 Sandy Clay

5 Sandy Clay Loam

6 Clay Loam

7 Silt

8 Silty Loam

9 Loam

10 Sand

11 Loamy Sand

12 Sandy Loam

Appendix B: Subbasin Statistics and Maps

Table B1: Study subbasin physical characteristics

	Vegetation class Percentage (%)					Elevation (m)	
Mean annual precip (mm)	Evergreen	Deciduous	Grass/ Ag	Unvegetated	Soil Type	Min	Max
1384	80.6	0.1	14.5	4.8	Clay	834	2180
2026	89.5	1.1	3.5	5.9	Loamy Sand	678	1665
251	77.7	0	20.8	1.5	Silty Loam	460	1248
1118	86.6	2.6	6.3	4.5	Loam	612	1962
816	85.6	0.7	7.5	6.1	Loam	518	1529
924	93.9	0.3	4.1	1.7	Loam	669	1866
225	77.1	2.2	19.4	1.3	Sandy Loam	197	661
246	74.7	0.1	23.6	1.6	Silty Loam	363	1214
611	60.3	0.2	34.3	5.2	Loam	440	1954
1657	89.6	1.2	6.1	3.1	Clay	677	1874
1991	83.9	5.6	3.1	7.5	Loamy Sand	668	1563
316	71.7	1.7	22.1	4.5	Silty Loam	245	1263

			Mean Daily Temp (˚C)	
Study ID	HUC12 ID	Aridity Class	Low	High
117	170300010201	Energy-limited	-0.5	9.6
394	170300010302	Energy-limited	1.1	9.8
457	170300031102	Water-limited	3.3	14.9
666	170300020203	Balanced	0.7	11.0
800	170300010507	Balanced	1.5	12.7
1606	170300010602	Balanced	0.8	11.6
1642	170300030807	Water-limited	3.9	17.9
1882	170300030202	Water-limited	3.1	15.3
1980	170300010403	Balanced	1.3	12.2
1989	170300010104	Energy-limited	0.6	10.3
2011	170300010304	Energy-limited	1.2	10.3
2045	170300030605	Water-limited	3.0	16.8

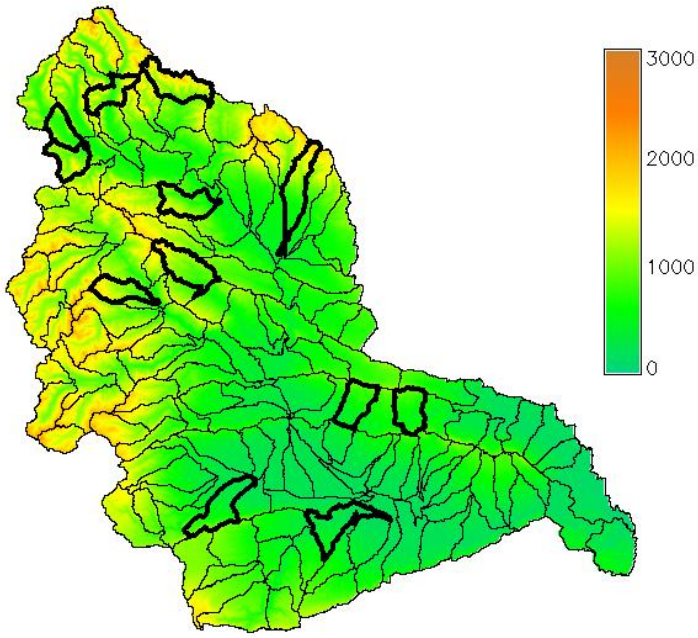


Figure B1: Yakima River basin elevation map (m)

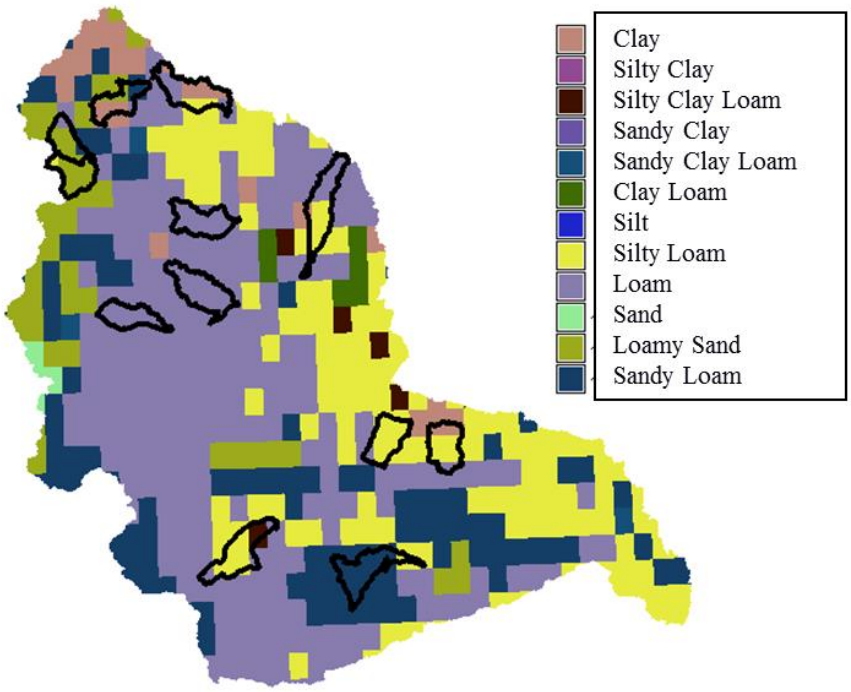


Figure B2: Yakima River basin soil types.

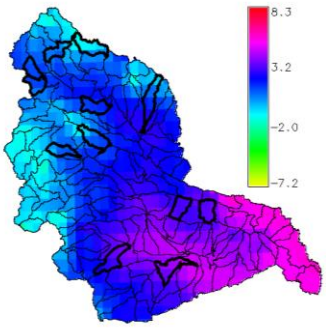


Figure B3: Yakima River Basin mean minimum temperature (˚C)

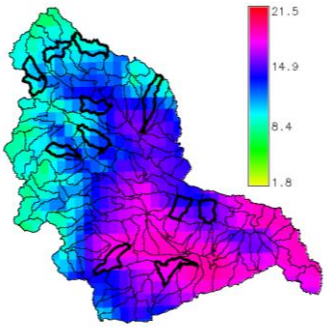


Figure B4: Yakima River Basin mean maximum temperature (˚C)

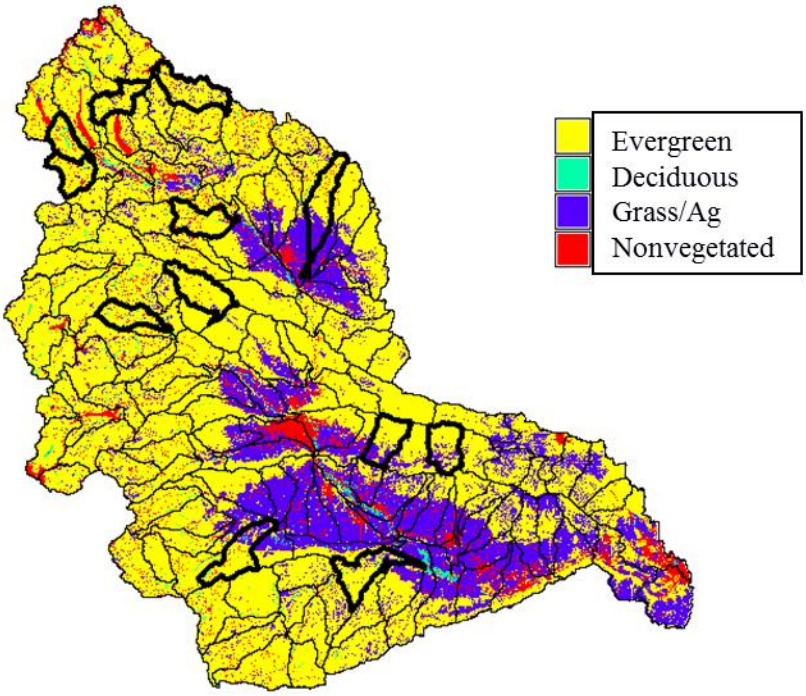


Figure B5: Yakima River basin vegetation classifications.