



THESIS
3
5579



This is to certify that the
dissertation entitled

Predicting the Impacts of Land Use and Climate on Regional-
Scale Hydrologic Fluxes

presented by

Anthony D Kendall

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Environmental Geosciences

A handwritten signature in dark ink, appearing to read "David Hyndman", written over a horizontal line.

Major Professor's Signature

A handwritten date "5/10/2009" in dark ink, written over a horizontal line.

Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
MAY 28 2013		
11-28-10		

**PREDICTING THE IMPACTS OF LAND USE AND CLIMATE ON
REGIONAL-SCALE HYDROLOGIC FLUXES**

By

Anthony D Kendall

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Environmental Geosciences

2009

ABSTRACT

PREDICTING THE IMPACTS OF LAND USE AND CLIMATE ON REGIONAL-SCALE HYDROLOGIC FLUXES

By

Anthony D Kendall

Climate and land use change are altering the terrestrial hydrologic cycle locally, regionally, and globally. Stream discharge and runoff from local to continental scales are increasing in much of the world, even as peak runoff timing shifts because of earlier snow melt. Regional precipitation patterns are shifting, changing the quantity and distribution of moisture. Land use and land cover changes due to both human and natural activities further alter hydrologic fluxes, changing the balance between runoff and evapotranspiration.

Managing water resources in the face of these changes requires a predictive understanding of their impacts at all scales. Most decisions about land use occur at local scales, smaller than individual watersheds. On the other hand, water withdrawal regulations are typically implemented at the statewide scale. These decisions then impact regional water budget and fluxes in large basins, often influencing the flows entering other states or nations.

Little is known about the quantitative influence of land use and cover types on terrestrial hydrologic fluxes, including groundwater recharge and evapotranspiration. At the plot scale, the processes governing biosphere-hydrosphere interactions have been intensively studied. But the tools necessary to examine fluxes at larger scales are not yet fully developed.

This dissertation presents the Integrated Landscape Hydrology Model (ILHM) designed to simulate fine-resolution terrestrial hydrologic fluxes over large domains. It is

a process-based, fully-distributed model that is parameterized solely using physically-measurable parameters. ILHM can be used with little calibration to determine the water budgets of regional watersheds. An application to the Muskegon River Watershed (MRW) in lower Michigan showed uncertainties in water budget estimates on the order of 10%.

ILHM also provides hourly values of a host of hydrologic fluxes and storages, and here is used to understand the relative influences of regional climate variability, land use, and soil texture on groundwater recharge and evapotranspiration (ET) in the MRW. Off-growing season lake-effect precipitation patterns strongly influence groundwater recharge, but not ET, resulting in nearly 50% more recharge near the lake shore than further inland. Soil textures play a very important role in governing both recharge and ET, with fine textures reducing recharge relative to sandy soils by nearly 40% and increasing upland ET by 30%. Variability between land use classes in sandy soils is almost 20% of upland ET and 45% of recharge.

This work is dedicated to my wife Cheryl, whose tireless support included countless hours of assistance in the field and expertise in the lab, and to my daughter Lydia, whose life began after Chapter 3 was written and has since helped her Daddy make it all the way through the end of Chapter 5.

ACKNOWLEDGMENTS

I would first like to thank my advisor, David Hyndman. His unwavering support during the development of ILHM helped me slog through some very difficult problems. His understanding of the physical processes that drive terrestrial hydrologic fluxes proved far more valuable than any textbook. Most importantly, however, has been his constant encouragement of my education and career. Had it not been for his suggestion now seven years ago that I pursue a graduate degree in the Geosciences, this Dissertation would be very different indeed.

Though I have already dedicated this work to my wife Cheryl, she deserves a place here as well. Her skills both in the field and the lab allowed me to accomplish far more in the early part of my degree than I would have otherwise.

There are a number of graduate students, and some undergraduates, who have lent their talents as well. In particular, Dush Jayawickreme and Nick Welty provided both field assistance and helped to develop early portions of ILHM code. Others, including Chris May, Matt Spansky, Jeff Eggleston, and Jason Bernstein lent numerous hours to collect data in the field.

Finally, my funding has come from a number of sources, including grants from the National Science Foundation, Great Lakes Fisheries Trust, and MSU's Center for Water Sciences. Additionally, MSU and the Department of Geological Sciences provided support through teaching assistantships, fellowships, and scholarships.

TABLE OF CONTENTS

LIST OF TABLES		viii
LIST OF FIGURES		xi
1 Introduction		1
1.1 Integrated Hydrologic Modeling		2
1.2 Description of Chapters		5
2 Examining Watershed Processes Using Spectral Analysis Methods Including the Scaled-Windowed Fourier Transform		7
2.1 Abstract		7
2.2 Introduction		8
2.3 Datasets and Study Area		10
2.4 Methods		12
2.5 Results and Discussion		31
2.6 Conclusions		46
3 Evaluating Temporal and Spatial Variations in Recharge and Streamflow Using the Integrated Landscape Hydrology Model (ILHM)		49
3.1 Abstract		49
3.2 Introduction		50
3.3 Methods		53
3.4 Results		70
3.5 Discussion and Conclusions		79
4 ILHM Development, Validation, and Simulation of Regional-Scale Stream Discharge		84
4.1 Introduction		84
4.2 The Integrated Landscape Hydrology Model		88
4.3 Model Domain and Input Data Preparation		103
4.4 Uncalibrated Model Results and Discussion		122
4.5 Conclusions		131
5 Influence of Land Use, Climate, and Soils on Recharge and Evapotranspiration Across a Regional Watershed		133
5.1 Introduction		133
5.2 Methods and Study Site		136

5.3	Model Results and Discussion	143
5.4	Conclusions	168
APPENDICES		170
A	ILHM Model Development	170
A.1	Evaporation and Transpiration	170
A.2	Infiltration, Percolation, Throughflow, and Exfiltration	177
A.3	Runoff Routing	180
B	ILHM Regional Upscaling	182
B.1	Surface Domain	182
B.2	Deep Unsaturated Zone	195
B.3	Saturated Zone	196
BIBLIOGRAPHY		197

LIST OF TABLES

2.1	List of the nominal time periods used in each seasonal analysis, and of the years omitted from the average unit responses. Data were from 1990-2001.	19
2.2	Slope break locations and β values for Figure 2.4.	34
3.1	List of adjustable parameters in the UEB Snowmelt Model and their assumed values	64
3.2	List of calibrated parameters for the unsaturated and saturated groundwater models.	71
4.1	Input climate and LAI data sources. Start and end dates are generalized. Locations of gage data are shown in Figure 4.4.	105
4.2	Table of observations and validation data. Location abbreviations “MR” refers to “Muskegon River”. The “Area” column refers to the surface watershed area.	106
4.3	Wetland depths assigned based on NWI classifications. The depths are based on a sinusoidal water table fluctuation of ~ 2.5 m.	109
4.4	Table of optimized constants for the NDVI $\rightarrow$ LAI mapping function 4.12, and coefficient of determination R^2 . Open water and barren are assumed to have $LAI = 0$ in ILHM.	117
4.5	Total Nash-Sutcliffe efficiency and average annual flux error values for each of the 6 USGS gages. Efficiency values are calculated using log-transformed simulated and gaged values. Average annual flux error is $sim - obs$ divided by the catchment area of each gage. The final column is average annual error normalized by average annual precipitation for that catchment.	130
5.1	Input climate and LAI data sources with the start and end dates of each source. Locations of gage data are shown in Figure 5.1.	140

5.2	Tables of multiple regression models for annual watershed-average A) Upland ET, and B) Upland DP. Columns include the root-mean square residual between the multiple regression model and ILHM, along with the correlation coefficient R^2 , and the model parameters. Note, the final five years of the simulation were excluded from this analysis because DP seems to behave exceptionally during this period compared the rest of the simulation.	150
5.3	Cross-correlation values between monthly climate indices.	151
5.4	Watershed-average annual mean, standard deviation, and coefficient of variation, c_v , of hydrologic fluxes.	154
5.5	Watershed-averaged mean, standard deviation, and coefficient of variation, c_v , of upland water budget components.	154
5.6	Average annual upland ET and DP by soil texture in agricultural areas between 75-150 km from the lake shore along the prevailing wind direction.	159
5.7	Average annual upland ET and DP by land use in sandy soils 75-150 km from the lake shore along the prevailing wind direction.	162
5.8	Nested linear multiple-regression models for upland ET. A) displays the model intercepts, coefficients, correlation coefficients, and root-mean square residuals ($RMSR$) between the ILHM results and the linear models, listed in B) are the ranges of values for each linear model parameter, C) shows the influence of each parameter relative to the mean field (the intercept value, abbreviated <i>Intcpt.</i>), calculated as the product of the range of each parameter and the model coefficients from part A. P_{GS} , P_{NGS} , T_{GS} and T_{NGS} are growing season and non-growing season precipitation and temperature, respectively. FC is soil field capacity, and all land use classes are abbreviated.	165
5.9	Nested linear multiple-regression models for upland DP. A) displays the model intercepts, coefficients, correlation coefficients, and root-mean square residuals between the ILHM results and the linear models, listed in B) are the ranges of values for each linear model parameter, C) shows the absolute value of the relative influence of each parameter, calculated as the product of the range of each parameter and the model coefficients from part A. Column headers as in Table 5.8.	166

A.1	Parameters used in calculating PET	172
A.2	Land-Use parameters	174
A.3	Soil properties	178
B.1	Land use dependent parameters in ILHM	186
B.2	Depression capacity, $S_{DEP,max}$ [$m \times 10^3$]. Values are from <i>Manfreda et al.</i> (2005), assuming no soil-texture dependence, and taking values for clay soil textures.	187
B.3	Soil-texture dependent parameters. Texture name abbreviates S, L, Si, C, and M are sand, loam, silt, clay, and muck, respectively. If multiple sources are listed for a particular parameter, values obtained from multiple sources were averaged	188
B.4	Table listing physical constants	189
B.5	Table listing albedo and emissivity parameters, and sources	189
B.6	Table of domain-independed physical parameters and sources	190
B.7	Table of domain-dependent physical parameters, values are explained in Chapter 4	190
B.8	Lookup table of unsaturated zone wetting front velocities.	195
B.9	Table of preliminarily-calibrated saturated hydraulic conductivities for deep sediments	196

LIST OF FIGURES

Images in this dissertation are presented in color.

2.1	Map of the Muskegon River and Grand Traverse Bay watersheds, with an inset map of Michigan showing their locations. Groundwater wells with transducers are marked with triangles, and the Evart gauge sub-basin is shaded in grey.	11
2.2	SWFT results for a test function, $f(x)$. a) Curves representing the successive superposition of the three sinusoids with different frequencies, and b) the SWFT scalogram. The shaded contours represent amplitudes between 0.5 and 1.0 as shown in the colorbar. The dashed line represents the boundary beyond which the SWFT scalogram is undefined (where half the window width is greater than the available number of points within the dataset).	26
2.3	a) Comparison of the heuristic snowmelt model to the UEB model, b) the difference between the two (heuristic-UEB) and the cumulative difference.	30
2.4	Composite plot of stream discharge, well-head relaxation, and precipitation spectra. Grey lines plot the raw FT spectra, solid lines are the binned-mean amplitudes, and dashed lines give the upper 95% confidence intervals. a) An overlay of the spectra in parts c-e and their linear fits, b) NEXRAD precipitation spectrum for 4/1/2004-11/30/2004, not plotted on part a, c) stream discharge spectra for 10/1/1997-9/30/2004, d) well-head relaxation for well B13 from 7/1/2003-7/1/2005, and e) watershed-available precipitation from 1/1/1990-1/1/2001. Note that, in part c the limited resolution makes it appear as if much of the stream discharge spectrum is above the confidence limit, however fewer than 5% of points exceed the limit in a spectrum with over 130,000 points.	32

2.5	SWFT of stream discharge at the USGS gage in Evart, MI. a) Stream discharge time-series divided by the area of the watershed above this gage, b) Scale-averaged amplitudes, A_p , c) Filled contour scalogram of the amplitude (A) vs. both period (P) and time (τ), along with the location of the amplitude-weighted mean period, Pt (white line), d) Time-averaged amplitude spectrum, At	38
2.6	SWFT of watershed-available precipitation at Big Rapids, MI. a) watershed-available precipitation time-series, b) Scale-averaged amplitudes, A_p , c) Filled contour scalogram of the amplitudes (A) vs. both period (P) and time (τ), along with the location of the amplitude-weighted mean period, Pt (white line), Note, white areas have amplitudes $< \log(\text{amplitude}) = -3.5$, d) Time-averaged amplitude spectrum, At	40
2.7	SWFT of stream discharge at Evart from 9/1/1991 to 8/31/2004. a) Stream discharge time-series, b) Filled contour scalogram of the amplitude (A) vs. both period (P) and time (τ). The year labels are centered at January 1.	41
2.8	SWFT of groundwater-head fluctuations in Well B10 in the Grand Traverse Bay Watershed. a) Well-head fluctuation time-series, b) Filled contour scalogram of the amplitudes (A) vs. both period (P) and time (τ).	43
2.9	Stream unit response hydrographs. a) Unit response functions obtained using data from 9/1/1999 to 8/31/2000, and b) Seasonal unit response functions using selected years (see Table 2.1) from 1990-1999.	44
2.10	Plots of convolved stream discharge from 09/01/1999 to 8/31/2000 using the annually- and seasonally-derived unit hydrographs (initial modeled discharge matched to stream discharge) on top of the measured stream discharge (a), and on top of the measured stream discharge. b) Plot of the percentage difference between each model and the measured discharge $(\text{modeled-measured})/(\text{measured}) \times 100$	45

3.1	On the left is a map for the Cedar Creek watershed (shaded) along with the groundwater contributing area to Cedar Creek (dashed outline). Also displayed on this map are locations of two stream gages, nine discharge measurement cross sections, and residential drinking water wells located within the watershed. On the right, a map of the lower peninsula of the state of Michigan shows the Cedar Creek watershed within the greater Muskegon River watershed. The boundary of a regional groundwater model of the Muskegon River is shown in bold on the state map. The three precipitation and climate gage locations are Hesperia (H), and Fremont (F).	54
3.2	Static GIS datasets that were integrated into the ILHM framework. From left to right, the datasets are land use/cover from IFMAP, quaternary geology from (<i>Farrand and Bell</i> , 1982), SSURGO soil textures, and land surface elevation from the NED 26.5 m DEM.	55
3.3	Time series data inputs to ILHM for 2003 and 2004: A) LAI of forest and agricultural land covers; B) monthly rain and snow along with daily average temperature, a horizontal line at 0°C is included as a visual aid; C) weekly averaged windspeed and solar flux; D) inset of hourly windspeed and solar flux data from 6/21/04 through 6/24/04; E) weekly averaged relative humidity (R.H) and temperature; F) inset of hourly R.H. and temperature from 6/21/04 through 6/24/04. . . .	60
3.4	Conceptual model and simplified box diagram of the ILHM.	62
3.5	Plot of measured discharge versus hydraulic radius in streams in the greater Muskegon River Watershed and adjacent watersheds. The hydraulic radius calculation assumes rectangular cross-section geometry. The power-law fit to the data produced a correlation coefficient, R^2 , of 0.77.	70
3.6	Uncalibrated simulated vs. observed groundwater heads plotted on top of a 1:1 line. Observations are from 96 residential wells installed in the Cedar Creek Watershed between 1990 and 2004 (see locations in Figure 3.1).	72
3.7	A & C) Upper and B & D) Lower Cedar Creek stream discharges with simulated values shown in white and gage values in black calculated based on stage discharge relationships. Manually measured discharge values are shown as black circles.	73

3.8	Uncalibrated simulated vs. observed flows for the Upper and Lower Cedar Creek gages plotted on top of a 1:1 line from approximately 20 manual discharge measurements taken at each of the upper and lower Cedar Creek gage locations.	75
3.9	ILHM-simulated average annual recharge for the Cedar Creek watershed. Annual total precipitation averaged approximately 83 cm during the period of simulation.	76
3.10	ILHM-simulated average annual precipitation excess runoff for the Cedar Creek watershed. Annual total precipitation averaged approximately 83 cm during the period of simulation.	77
3.11	A) Monthly average deep percolation (bars) in the Cedar Creek watershed plotted with the ratio of percolation to total precipitation for each month (stars). B) Monthly difference in deep percolation between forested and agricultural land covers.	78
3.12	Average monthly evaporation (E), transpiration (T), and evapotranspiration (E+T) plotted for both forested and agricultural landscape as stacked bars along with the difference (Forest-Agriculture) between the two land-use types plotted as circles.	80
4.1	Graphic depiction of the spatial discretization scheme used in ILHM.	90
4.2	ILHM generalized process flow diagram, focusing on surface processes. Dashed outlines indicate optional items.	92
4.3	Detail of hydrologic fluxes calculated by ILHM for upland (left) and wetland (right) portions of cells.	94
4.4	Map of the Muskegon River Watershed and surrounding model boundary. In part A, major municipalities along the drainage network are highlighted. Part B shows the locations of input gages and observation data locations.	104
4.5	Static GIS inputs for ILHM expanded model boundary	108

4.6	Plot of annual model-averaged solar radiation, from 1980 through 2007. There are two primary periods, 1) only cloudiness observations are available for shortwave radiation modeling, 2) cloudiness observations along with limited solar radiation data are available.	113
4.7	Plot of monthly model-averaged precipitation from daily interpolated NCDC precipitation gages, and NEXRAD radar from 1996 - 2008. The quality of the NEXRAD data can be differentiated into three periods, 1) all monthly observations are a poor fit to gage data, 2) NEXRAD provides a good match to months with liquid precipitation only, and 3) NEXRAD data and gage data match closely in almost all months.	115
4.8	Plot of monthly model-averaged LAI from three sources: 1) 8 km GIMMS AVHRR data, converted from NDVI to LAI, 2) 1 km AVHRR data, uncorrected, converted from NDVI to LAI, and 3) MODIS LAI data.	120
4.9	Simulated and observed (gage) hydrographs for three USGS stream gages, A) Muskegon River at Evert, B) Clam River at Vogel Center, and C) Bear Creek near Muskegon.	124
4.10	Plots of simulated and observed streamflows and groundwater head, observation data are described in Table 4.2. A) Observed daily discharge at the Evert, MI gage vs. unbiased residual $(sim - obs) / obs$, B) observed discharge from 105 points across the MRW during low-flow conditions vs. simulated baseflow , C) observed groundwater head vs. simulated head, and D) annual median and ± 1 standard deviation of the head residuals $sim - obs$. All horizontal and diagonal lines are reference lines, not linear fits.	126
4.11	Plot of water-year (Oct - Sep) Nash-Sutcliffe efficiency of the log-transformed simulated and gage streamflow values for all six USGS gages. "MR" refers to "Muskegon River".	128
4.12	Plots of stream discharge error $sim - obs$ for all six USGS gages normalized by the area of each gage catchment. Plotted in A) are water-year error values, and B) are average monthly error values across the 27 water years in the simulation.	129

5.1	Map of the Muskegon River Watershed and surrounding model domain boundary, with an inset showing the location of the model domain within lower Michigan, USA. Filled symbols on the lower map are climate gage locations for data listed in Table 5.1.	137
5.2	Maps of select ILHM input datasets. Sub-figures A and B show distributions of major and minor land use classes, C illustrates the dominant soil textural classes, and D plots the average annual precipitation across the model domain.	138
5.3	Monthly values of select climate and leaf area index (LAI) inputs. The heavy solid line is the median monthly value during the simulation period, which is surrounded by a shaded range indicating the 5% and 95% values. Plotted in A) monthly temperatures, B) monthly precipitation, C) monthly average solar radiation, and D) monthly average LAI for four land use classes.	139
5.4	Maps of total cell mean annual modeled A) evapotranspiration (ET), B) mean annual modeled deep percolation (DP), C) coefficient of variation (ratio of standard deviation to the mean) of annual ET, D) coefficient of variation, c_v of annual DP. Note the color scales differ for each sub-figure.	146
5.5	Plots of annual watershed-averaged fluxes for the MRW. On the left, watershed-average ET, storage, and runoff for the MRW are plotted, while on the right upland ET, overland flow, root-zone storage, and DP within upland landunits are shown. Parts A & B plot annual fluxes in cm. Those quantities, divided by annual precipitation, are plotted in parts C & D. Note, for storage, negative values indicate release of water from storages in the model.	148
5.6	Average monthly fluxes of A) watershed ET, runoff (overland + ground-water discharge), and storage, and B) upland ET, overland flow, root-zone storage, and DP within upland landunits.	152
5.7	Filled area plot of watershed-average monthly evaporation and transpiration components for the MRW. Shown are upland transpiration, wetland transpiration, soil evaporation, canopy evaporation, snow and ice sublimation, and open water evaporation.	153

5.8	Filled value plots of average monthly A) precipitation, B) upland evapotranspiration (ET), and C) upland deep percolation (DP) vs. distance from Lake Michigan along the prevailing wind direction. All values are averages of only cells within the model domain containing greater than 75% deciduous forests, with sand as the dominant texture. Values are averaged over the entire model period.	156
5.9	Plots of monthly average upland ET and DP across generalized soil texture classes as a percent of water-year (Oct - Sep) precipitation. Values are from cells within the model domain containing greater than 75% agricultural land use between 75-150 km from the lake shore along the prevailing wind direction.	158
5.10	Plots of monthly average upland ET and DP across land use classes as a percent of water-year precipitation. Values are from cells within the model domain located between 75-150 km from the lake shore along the prevailing wind direction with either sands (A & B) or loams (C & D) as the dominant soil texture class both vertically and horizontally.	161
5.11	Maps of linear regression modeled annual average A) Upland ET and B) Upland deep percolation (DP) for 1980 - 2007. Parts C) and D) show mean root mean square residuals (RMSR) between the ILHM annual fluxes and the linear regression modeled values of ET and DP respectively.	167

CHAPTER 1

Introduction

Climate and land use change are altering the terrestrial water cycle worldwide (*Labat et al.*, 2004). Effects of these changes vary regionally in response to heterogeneous forcings, landscape properties, and anthropogenic influences. Impacts of these changes are felt across a variety of scales, from hectares to continents. These impacts include changes to soil moisture, groundwater recharge, evaporation and transpiration, and stream discharge.

Changes to the soil moisture regime may be the single most economically-disruptive impact of climate warming. Increased evapotranspirative demand coupled with a longer growing season and shifting seasonality and quantity of precipitation will alter the productivity of non-irrigated agricultural. In many regions, including the mid-western US, growing season precipitation is projected to decline, even while annual totals increase (*Christensen et al.*, 2007).

Increased demand for irrigation will stress highly-allocated water resources. In much of the western US, net water demand greatly exceeds annual supplies (*Vorosmarty et al.*, 2000). Elsewhere, much of what is not allocated to human use provides stream discharge and surface water levels to ecosystems adapted to the status quo. Irrigation alters the water balance of a watershed by increasing evapotranspiration, and thereby reducing groundwater levels and stream discharge.

Agricultural practices are not the only land use changes that disrupt terrestrial wa-

ter cycle. Reforestation and deforestation, urbanization, draining and channelization of wetlands, and impoundment of streams all affect the balance between ET, runoff, and storage. Many of these land use changes occur simultaneously, particular within larger basins, though the character of the aggregate varies regionally.

Our ability to quantitatively forecast the impacts of climate and land use change is limited due primarily to the complexity of the interactions among landscape processes, and their characteristic spatial and temporal scales. Precipitation, temperature, solar radiation, and other climatic factors in the hydrologic cycle vary globally, regionally, and more locally due to orographic and lake effects. Discharge within a large river system is the aggregate of the response of each individual square meter within that basin, whose infiltration capacities, land cover, and relief vary considerably over hectare scales. Beneath the surface, groundwater systems of varying complexity and heterogeneity transmit fluxes to surface water bodies, with varying degrees of influence depending on the hydrogeologic setting.

1.1 Integrated Hydrologic Modeling

To develop predictive capabilities, we must first understand how climate and land use impact terrestrial hydrology. One of the primary tools to understand environmental processes is long-term monitoring. Unfortunately, except for stream discharge, very few hydrologic time series are available for more than a few decades. Critical information, such as plot-scale soil moisture, groundwater levels, wetland inundation extents, shortwave solar radiation, plant phenological development, soil temperature, evapotranspiration, and many others are collected in very few locations and generally over short periods only. As a result, direct water-balance calculations generally require estimation of one or several terms of the water budget. The result is that critical landscape and climate input heterogeneities are not resolved.

Process-based models can be used both to understand the behavior of complex systems and as a predictive tool to forecast impacts of changes. Recent advances in processing capabilities, and the general availability of GIS data, makes the task of simulating the terrestrial water cycle at the relevant spatial and temporal scales feasible. There are dozens of widely-available hydrologic models, but many do not directly simulate physical processes such as evaporation, transpiration, snow accumulation and melt, soil moisture vertical redistribution, groundwater recharge, and streamflow. Many do, however, simulate physical processes, though few meet the following set of requirements:

1. The code must be open-source, so that process descriptions in the model can be modified and examined
2. The model must simulate key hydrologic fluxes, including overland flow runoff and routing, groundwater recharge, surface and groundwater evapotranspiration, snow processes, and soil moisture redistribution.
3. Surface, near-surface, and ground water processes must all be explicitly described, not simply by linear reservoir routing.
4. Fine-resolution, large-domain modeling must be possible and feasible, both computationally and practically. This includes being able to run on modest equipment (desktop machines or small clusters), and incorporate GIS inputs to facilitate rapid development.

Historically, hydrologic models have been developed for application at a narrow range of scales. Land surface models beneath global and regional climate models run at hourly timesteps over hundreds of square kilometer cells, but typically do not include comprehensive groundwater modules. Watershed models operate on catchments at a variety of scales while lumping the processes within those catchments into a collection

of calibratable parameters. Plot-scale and one-dimensional root zone models simulate a host of physical processes with vertical layers as small as a few centimeters. Fully-coupled hydrologic models describe small catchments with dozens of vertical layers and cell or mesh sizes as small as a few square meters.

None of these classes of models are well suited for simultaneously describing the locale, regional, and continental impacts of land use and climate change. More recently, a new class of integrated hydrologic model has arisen, coupling one-dimensional surface process and three dimensional saturated zone models. These have mostly adapted existing watershed models, some fully- or semi-distributed, others lumped parameter, and coupled them with existing groundwater modeling codes. The heritage of the separate models varies widely, but all are an attempt to address the issue of fine-resolution, large-scale terrestrial hydrologic modeling.

When my research began in 2004, few of these models were available. Those that were had been created prior to the general adoption of GIS technologies, and thus ill-suited for fine-resolution, large-domain modeling. In order to satisfy the four criteria above, a combination of existing and newly-development models were brought together into an integrated suite called the Integrated Landscape Hydrology Model (ILHM).

ILHM is similar to this new class of integrated hydrologic models in that it explicitly simulates the processes of the surface routing, canopy and root-zone, deep unsaturated-zone, and saturated zone domains. It was written in a modern, high-level programming language, MATLAB, that facilitates rapid development and testing. The code is modular, and developing and implementing new capabilities requires days, rather than weeks or months common to older codes written in compiled languages. Eventually, the core code itself may be re-written in a compiled language to enhance performance and interoperability with other models.

The development effort was focused along three pathways: 1) writing the core code

and process capabilities, 2) writing an interface to the code that was fast, powerful, flexible, and scalable to eventually enable much larger simulations, and 3) developing a series of methods and practices for building large-domain, fine-resolution hourly climate inputs over a three decade period. In total, the code is now 70,000 lines, not including pre-compiled libraries or external interface tools. From the beginning, my intention has been that the code may eventually be adopted by others, thus significant effort was invested in extensive comments within the code, documenting the modules themselves, and creating a support infrastructure to handle multiple contributors and developers.

ILHM is intended to be both a fully-explicit water- and energy-balance code, with run-time coupling between all domain modules (surface routing, canopy and root zone, unsaturated zone, and saturated groundwater). Currently, however, coupling is one way though full-coupling is under active development. Additionally, while the entire model preserves water-balance, only portions such as the snowpack module and lake temperature module, preserve energy-balance. A new set of physical process modules have been developed that will complete the energy-balance capabilities of ILHM, but are not yet fully complete and tested.

1.2 Description of Chapters

All of the chapters in this Thesis address the same common theme: how do climate and land use impact terrestrial hydrology? In the next Chapter, very simple models are used in conjunction with time- and frequency-domain analytical techniques. The following two Chapters describe the development of ILHM, and its application first to a small catchment in central Lower Michigan, and then finally to a regional watershed and surrounding domain nearly of 19,000 km². The Appendices present detailed equations and tables describing the structure of and processes simulated by

ILHM. The final Chapter focuses on interpreting the results from a 28-year ILHM simulation to understand how local and regional climate, soil texture, and land use impact hydrologic fluxes.

Chapters 2 and 3 were published in a slightly different form in 2007 (*Kendall and Hyndman, 2007; Hyndman et al., 2007*). They have been modified to fit the formatting of this document.

CHAPTER 2

Examining Watershed Processes Using Spectral Analysis Methods Including the Scaled-Windowed Fourier Transform

This chapter was published in 2007 in a slightly modified form in an AGU Geophysical monograph, *Subsurface Hydrology: Data Integration for Properties and Processes* (Kendall and Hyndman, 2007).

2.1 Abstract

Important characteristics of watershed processes can be extracted from hydrologic data using spectral methods. We extract quantitative information from precipitation, stream discharge, and groundwater head data from watersheds in northern-lower Michigan using Fourier Transform (FT) methods. By comparing the spectra of these data using similar units, we graphically illustrate the hydrologic processes that link precipitation to stream discharge and groundwater levels including evapotranspiration. We also demonstrate how unit hydrographs can be efficiently and non-parametrically derived using the FT in a manner that allows for a quantitative seasonal comparison of precipitation and the resulting stream discharge response. This

analysis clearly illustrates the reduction in summer discharge levels due to canopy interception and evapotranspiration. We also develop a systematic application of the FT we call the Scaled-Windowed Fourier Transform (SWFT), which extracts time-varying spectral content using a similar approach to the wavelet transform. While computationally less efficient than the wavelet transform, the SWFT allows for embedded detrending and tapering. Application of this method clearly illustrates the non-stationarities of spectral content within the three chosen data types, leading to a greater understanding of discharge-generating processes.

2.2 Introduction

Spectral analysis (SA) provides a powerful means of extracting information from hydrologic data. This type of analysis can reveal processes that may be obscured in direct time-series analysis by providing data not just on temporal fluctuations, but also on the spectrum of frequencies within those fluctuations. Furthermore, integrating multiple datatypes in a quantitative and meaningful way is relatively simple using spectral analysis. Here we apply both existing and novel SA techniques to hydrologic data from two watersheds in northern lower-Michigan. We use these methods to illuminate the processes that link precipitation to stream discharge and groundwater levels.

Spectral analysis has been applied within the hydrologic sciences for decades (e.g., *Bras and Rodriguez-Iturbe* (1985); *Hameed* (1984)), most commonly as a means of estimating coefficients for linear autoregressive or stochastic models (e.g., *Naff and Gutjahr* (1983); *Jukic and Denic-Jukic* (2004); *Zhang and Schilling* (2004)). During the last decade, SA has been applied to examine fractal and multi-fractal behavior of systems (e.g. *Tessier et al.* (1996); *Pelletier and Turcotte* (1997); *Kirchner et al.* (2000)), and to examine linkages between hydrologic and climatic processes (eg.

Tessier et al. (1996); Pelletier and Turcotte (1997); Coulibaly and Burn (2004)).

Here we apply SA techniques to explore linkages between hydrologic processes and to provide a deeper understanding of those processes. Previous SA process comparison studies have generally not assured the similarity of measurement units, nor have they (with the exception of more recent wavelet studies including, e.g. *Gauchere* (2002) and *Coulibaly and Burn (2004)*) considered the time-variant nature of process spectra. We demonstrate the quantitative utility of comparing data and spectra with similar physical units for inference of process relationships. Also, using an application of the common Fourier transform we call the Scaled-Windowed Fourier Transform (SWFT), we illustrate the non-stationary behavior of precipitation, stream discharge, and groundwater head spectra. Comparing Fourier spectra to the SWFT output, we illustrate how ignoring such non-stationarity can result in misinterpretations of spectra and thus of inferred process details. Finally, using a spectral derivation of seasonal unit hydrographs, we demonstrate how simple stream discharge models can be improved by considering the seasonality of watershed processes.

Recognizing that spectral analysis is not a standard tool in the hydrologic science toolbox, *Fleming et al. (2002)* published a practical introduction to SA that focused on applications of the Fourier Transform (FT) including direct frequency domain investigation, spectral filtering, and spectral simulation model validation. Here we apply SA in a similar manner, explaining our assumptions and the common complications, and demonstrating how these methods can be used across the hydrological sciences. Additionally, parts of the methods section are intended to contribute to a set of "best practices" of SA in the hydrologic sciences.

2.3 Datasets and Study Area

Six data types from two northern lower-Michigan watersheds were used in this study: daily temperature, precipitation, and snowfall, both 15-minute and hourly stream discharge, and bi-hourly groundwater heads. Figure 2.1 shows the locations of both Michigan watersheds, as well as the locations of our groundwater transducers. The 3711 km² Evert sub-basin of the Muskegon River Watershed (MRW) was selected for this study because there are no actively controlled flow structures on the Muskegon River or its tributaries upstream of this station.

Because the Evert sub-basin lacks a network of monitored groundwater wells, we used water level data from our pressure transducers in the nearly adjacent Grand Traverse Bay Watershed (GTBW), which has a similar hydrogeological setting. Unconsolidated sediments within the Evert sub-basin and the GTBW were deposited by the same set of glacial episodes. These sediments are characterized primarily by coarse to fine sands and gravels with a small percentage of fine-grained material (*Farand and Bell*, 1982). The similarity in depositional history and topographic variation between these two basins suggests that GTBW groundwater head fluctuations can be used as a proxy for the nearby upper MRW.

Preliminary 15- and 60-minute discharge data from the United States Geological Survey (USGS) stream gauge on the Muskegon River at Evert (Station ID: 4121500) were used in this study (*USGS*, 2005). Hourly data for this gauge were used from October 1989 until 15-minute discharge data became available in 1997. Combining the hourly data with resampled 15-minute data, the total data record extends from 10/1/1989 through 9/30/2004.

For finer temporal-scale resolution, we processed hourly 4-km NEXRAD data (*Andresen*, 2004) for 2004 and extracted the mean rainfall over the Evert sub-basin at each hourly interval. NEXRAD data show a high degree of correlation to the corresponding gauge values in this region (*Jayawickreme and Hyndman*, 2007). These

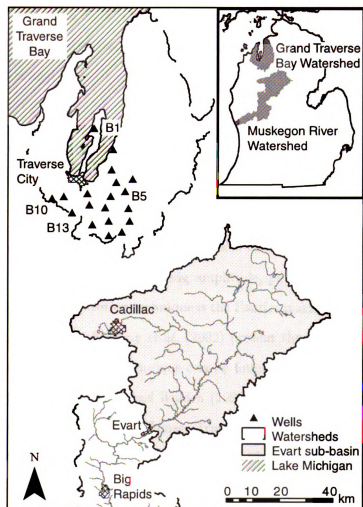


Figure 2.1: Map of the Muskegon River and Grand Traverse Bay watersheds, with an inset map of Michigan showing their locations. Groundwater wells with transducers are marked with triangles, and the Evart gauge sub-basin is shaded in grey.

radar-based precipitation data are only available for approximately 9 months out of the year in this region (from 4/1/2004 to 11/30/2004) due to errors in snowfall estimates, so they are only suitable for evaluating shorter-period system behavior.

We installed a network of 17 pressure transducers in USGS groundwater wells across the GTBW. Water table elevations were recorded every two hours beginning either in June 2003 (9 transducers) or June 2004 (8 transducers). Data through 7/1/2005 were used in this analysis.

2.4 Methods

2.4.1 Fourier Transform

Fourier's Theorem states that any complex periodic function can be decomposed into a set of periodic basis functions of varying amplitude, period, and phase shift. The most common such decomposition technique is the Fourier transform (FT), which uses sinusoidal basis functions. *Fleming et al.* (2002) present the basic theory, and a full mathematical treatment of this technique can be found in textbooks (e.g. *Bras and Rodriguez-Iturbe* (1985); *Percival and Walden* (1993)). A common implementation of the discrete FT is the discrete Fast Fourier Transform (FFT) popularized by *Cooley and Tukey* (1965). In this study, we use the FFTW libraries that are integrated into the MATLAB computing environment (*Frigo and Johnson*, 1998). This particular set of general-radix algorithms does not constrain the user to the requirement in *Cooley and Tukey* (1965) that the number of samples (N) be a power of 2.

The output of the FFT algorithm is a complex array representing the magnitude and phase shifts of the Fourier coefficients. For instance, the FFT of a sinusoid of unit period and amplitude is an array with a single non-zero value corresponding to a period of 1. The FFT of a complex summation of sinusoids would produce non-zero

peaks in the spectrum. If instead the time-series input to the FFT were a broad, single-peaked curve such as a gaussian, exponential, or gamma function, spectral amplitudes would be non-zero across a broad range of periods.

Fourier spectra are generally plotted as power vs. frequency, however we find that plots of amplitude vs. period provide a more natural basis for viewing spectra of interrelated physical phenomena. The spectral amplitude of a time-domain input corresponds directly to hydrologic flux quantities such as precipitation and stream discharge. Note that the log-log slope of the amplitude spectrum is $\frac{1}{2} \beta$, where β is the slope of spectral power in log-log coordinates. A system is typically considered fractal (or multi-fractal) if its power spectrum roughly follows behavior (*Avnir et al.*, 1998).

If one assures the similarity of the units of each time-series dataset, the amplitudes of multiple spectra can be compared in a physically meaningful manner. To accomplish this, a suitable unit for comparison must first be chosen. For this study, we chose length/time (L/T) units because precipitation is the most common hydrologic forcing mechanism. In order to change the volumetric discharge units (L^3/T) of stream discharge to L/T units, the discharge was divided by the drainage area of the watershed upstream of the gauge. Groundwater head data, measured in L units can be differentiated to yield L/T units. Using the differentiated data, an approximation of the rate of head relaxation at the well location can be obtained by selecting only the negative values. Note, this technique assumes constant lateral inflow. This was then multiplied by an estimate of drainable porosity for the sediments (0.2), to correct for water level differences in porous media vs. open water.

2.4.2 Assuring Periodicity

The FFT algorithm assumes that the data are periodic, namely the starting and ending points of the dataset are identical. Violating this condition results in spurious

features in the output spectrum (*Bach and Meigen, 1999*). However, time-series data from environmental systems rarely satisfy this criterion. There are four primary means of assuring periodicity: 1) filtering, not considered here, refer to *Fleming et al. (2002)* for a discussion, 2) taper function multiplication (tapering), 3) data subset selection (discussed in "Reducing Aliasing and Leakage" below), and 4) trend subtraction.

Tapering can be a valid means of forcing periodicity if one considers how it affects the resultant spectra. Tapering refers to the multiplication of a mean-removed signal by a "tapering function" (sometimes referred to as a windowing function) that smoothly tapers from a peak of 1 at its center to 0 at the edges. There are a variety of pre-defined tapering functions (*Blackman and Tukey, 1958; Harris, 1978*), and each affects the spectra differently. Tapering has the side effect of reducing the amount of information in the signal, thus limiting its applicability for short data records (*Fleming et al., 2002*). The shape of the tapering function, and how gradually it tapers near the edges, controls how much information is lost in the process. There is a tradeoff since the tapering functions that reduce information loss have more severe spectral "side lobes", which distort the transformed spectra by shifting power from primary frequencies to harmonics (integer multiples or factors) of those frequencies. To recover the amplitude of an isolated peak within a spectrum (but not the entire spectrum itself), spectra from each tapering function can be corrected to account for the effect of side lobes. The multiplicative correction factor for each tapering function is given by the ratio of unaltered- to tapered-amplitude (*Harris, 1978*). In this study, tapering is primarily applied within the SWFT method where distortion is minimal since only a single amplitude is selected from each FT.

Trend removal is often necessary for environmental data series to assure periodicity while minimizing loss or side-lobe distortion from tapering. The simplest method is to linearly remove the trend from the data, which can be effective if the non-

periodicity of the data is associated with a nearly linear trend. If there is some roughly sinusoidal long period fluctuation, a more valid means of trend removal may be to subtract a half-period sinusoid from the signal. In this case, the peak and trough of the subtracted function occur at each end of the original signal. A variety of additional trend-removal techniques have also been developed (*Mann*, 2004). We used the sinusoidal trend removal method in this study due to its simplicity and physical basis.

2.4.3 Reducing Aliasing and Leakage

If the FT is blindly applied to data without thought to dominant system processes, sampling rates, or sampling interval, aliasing and leakage may occur. Both aliasing and leakage act to shift spectral power (or amplitude) from “true” frequencies to harmonics of those frequencies, although each acts differently. Aliasing results from under sampling high-frequency fluctuations (*Bras and Rodriguez-Iturbe*, 1985), while leakage or overspill is caused by both the non-periodicity of the system as well as non-periodicity of the processes within that system (*Bach and Meigen*, 1999). It is important to note that leakage will occur if the endpoints do not match but the inverse is not always true. Even if the time-series endpoints match, leakage may occur due to variability of the processes that contribute to the sampled time-series.

In theory, aliasing can be avoided by merely increasing the sampling rate until the time series is fully resolved. Specifically, one cannot resolve spectral peaks with frequencies greater than half the sampling rate (the Nyquist frequency) (*Bras and Rodriguez-Iturbe*, 1985). If a time-series has significant power in frequencies above half the sampling rate, aliased power will be present in the empirical spectrum. In the case of many environmental datasets, aliased peaks may not be significant because as is shown below, these datasets are typically strongly damped at high frequencies. Our analysis indicates that some data do require sampling frequencies on the order

of once per hour, but most of those discussed here are sufficiently sampled with daily sampling rates.

Spectral leakage can be more persistent and troubling than aliasing because periodic processes within a system are rarely sampled over an integer number of cycles. Leakage is commonly reduced by applying a tapering function to the time-series, detrending the time-series, or both (*Fleming et al.*, 2002). However, when the entire FT spectrum is of interest rather than specific spectral peaks, a more effective means of decreasing leakage in environmental datasets may be to carefully select subsets of the data that correspond to natural breaks in processes, thus assuring near-integer sampling without distortion.

Data subset selection is also important because most environmental processes are non-stationary, thus each occurrence of a process may vary in both period and amplitude. If one includes multiple cycles of a periodic process in the FT, the true peak location, shape, and amplitude can be obscured by differences in system states across cycles. Additionally, selecting a subset of data in which some system processes are inactive can also greatly simplify spectral analyses and reveal the spectra of weaker processes in portions of datasets that would otherwise be obscured by dominant, but intermittent, processes. For example, the diurnal fluctuation of stream discharge is often clearly visible during baseflow conditions but can be obscured by runoff and near-stream groundwater discharge during late spring and early summer. Alternately, if one were only interested in the spectral behavior of runoff, for instance, then selecting data surrounding an isolated moderate-precipitation event during a dry season yields a stream discharge response primarily to direct precipitation and runoff.

2.4.4 Unit Response Functions

While there are a variety of techniques to derive unit hydrographs from discharge and precipitation data, most are either ill-posed or require assumptions about system

behavior (*Yang and Han, 2006*). But direct FT deconvolution can produce unit hydrographs quickly and deterministically. The total time-series response of a linear system to a forcing input can be derived by the convolution of the system unit response and the input time-series as follows (*Smith, 1997*):

$$q(t) = \int_0^{\infty} h(\tau)p(t - \tau)d\tau \quad (2.1)$$

where q is total time-series response (i.e. stream discharge), h is the unit response (for the case of stream discharge, this unit response has the special name of “unit hydrograph”), and p is the input precipitation time series. According to the convolution theorem,

$$Q(k) = H(k)P(k) \quad (2.2)$$

where $Q(k)$, $H(k)$, and $P(k)$ are the FTs of q , h , and p , respectively. The unit hydrograph time-series is then

$$h(t) = F^{-1} \left[\frac{Q(k)}{P(k)} \right] (t) \quad (2.3)$$

where $F^{-1}[\dots](t)$ denotes the inverse Fourier transform. Thus the unit-response hydrograph is given by the inverse FT of the ratio of the discharge and precipitation spectra.

Though not strictly necessary, the analysis is simplified if the sampling frequency and units of the precipitation and discharge time-series are identical. The data should be resampled so that the number of samples, n , and the sampling frequency, f , are equivalent. The unit-response function, $h(t)$, has length n , however, in this case the unit response function is only valid up to the point where it becomes negative, since precipitation can not directly produce a decrease in stream discharge (*Yang and Han, 2006*). The negative response is therefore the signature of some other watershed process. If baseflow separation is used, this issue can be avoided. Though some small uncertainties will remain due to the data themselves, perhaps producing

negative responses. However, we chose not to use baseflow separation as it produces a synthetic dataset not directly tied to watershed process, thus conflicting with the investigative nature of these spectral analysis techniques.

The resultant unit-response hydrograph is not an invariant property of the watershed, as it is sensitive to variations in runoff-generating processes. These processes can be studied by directly comparing different unit-response hydrographs. Differences in response curve timing, peak, and shape can all be used to infer the activity and relative influence of various watershed processes. Applying data subset selection with these process differences in mind can allow for a quantitative sensitivity analysis of system sub-processes.

In this study, we compare seasonally derived unit-response hydrographs for the Evart sub-basin averaged over 10 consecutive years. The derived unit hydrograph will be incorrect if stream discharge is still responding to precipitation inputs that occurred prior to the start of the data period (*Smith, 1997*). However, applied over entire seasons this error, as well as any error resulting from noisy data, is greatly reduced. Nevertheless, the derived unit-response function remains highly sensitive to edge conditions of the time-series inputs, thus the nominal time period (given by Table 2.1) of each season was adjusted to remove precipitation events or sudden increases in discharge near the edges of the time-series. The nominal time period for each season does not correspond to starting and ending dates of each season because they were chosen to assure similarity of hydrologic response within a season based on assumptions of process activity, also listed in Table 2.1. Also note that the fall and winter seasons overlap because the minimum length of the time-series subset must be longer than the watershed response time, which in this case was on the order of 60 days.

To assure periodicity for the FT of each season's data, a Tukey tapering function (*Blackman and Tukey, 1958*) was applied with relatively steep taper (coefficient of 0.1).

Table 2.1: List of the nominal time periods used in each seasonal analysis, and of the years omitted from the average unit responses. Data were from 1990-2001.

<i>Season</i>	<i>Nominal Period</i>	<i>Justification</i>	<i>Years Omitted</i>
Winter	12/1-3/15	Mainly frozen precipitation	1991, 1992
Spring	3/15-5/15	Little evapotranspiration	1990, 1993-99 ¹
Summer	6/15-10/1	Maximum evapotranspiration	1990-91, 1993, 1995-96
Fall	10/15-12/25	Mostly liquid precipitation	1991, 1994-95

Note: ¹Only four years were included in the spring average discharge response, 1991, 1992, 2000, 2001.

Tapering was chosen over trend-removal because the magnitude of the discharge response to precipitation was of primary importance. After making these adjustments, some seasons continued to produce non-physical results (characterized primarily by sinusoidal unit-response behavior) and were thus omitted. These omissions are justified on the grounds that the non-physical results reveal only the sensitivities and limitations of the method, and nothing about watershed process.

2.4.5 Scaled-Windowed Fourier Transform

A key difficulty in applying the FT to environmental datasets is that non-stationarities in the data introduce artifacts in the Fourier spectrum. Here there are two types of non-stationarity to consider. The first is non-stationarity of process period, where subsequent cycles of a process within the system have slightly different spectra due to changes in system properties. The second is best described as intermittency, where a process may be active during only a portion of the sampled data. The first results in spreading of spectral power about a central period (if the process is sinusoidal), while the second results in spurious spectral power at harmonics of the primary period.

To avoid these effects and more clearly illuminate important changes in the spectra, we developed a method that we call the Scaled-Windowed Fourier Transform (SWFT). The SWFT is very similar to a sinusoidal wavelet transform (WT), but it differs in a number of respects. The mathematical development of the SWFT pre-

sented here is fundamentally different from that commonly presented for the WT, in a way we feel increases the utility of this method of hydrologic scientists. In particular we feel that there is great value in demonstrating that the SWFT produces Fourier coefficients localized in both the time- and frequency-domains, as opposed to the similarly localized WT coefficients that are wavelet-dependent. Additionally, the SWFT is capable of embedded detrending rather than relying on tapering alone to reduce leakage, potentially producing better spectral estimates. Finally as developed, the periods and times queried by the SWFT are more flexible than typical WT schemes, with the tradeoff of decreased computational efficiency. The SWFT also differs from the traditional windowed FT that only transforms data within a specific subset of the overall time series called a data window (here the window is different from a tapering function). This data window is then slid along the time series to produce a map of spectral power varying in both period and time. Unfortunately, the windowed FT forces a choice between severe aliasing of low-frequency components of the signal or poor resolution of high-frequency non-stationary processes (*Torrence and Compo, 1997*).

By contrast to the windowed FT, the SWFT method scales the width of the window over successive passes along the time-series. At each window position, the data are detrended, multiplied by a tapering function, and Fourier transformed. A single amplitude corresponding to the Fourier coefficient of a single frequency is selected from the complete FT at each window position. The window is then slid along the data, producing a time-varying series of amplitudes for that frequency. When the end of the dataset is reached the window width is rescaled and the process is repeated for the next frequency. Note, hereafter we use the word "period" solely to indicate the spectral period, or the inverse of frequency. We feel that the period of spectral content is generally more applicable to the hydrologic sciences than the frequency.

The SWFT produces the same type of scalogram as would be generated with the

WT. These scalograms are well defined mathematically and physically (if proper units are used), and can be examined using standard statistical techniques. Aside from simple examination of the scalogram, comparisons of related scalograms such as the cross-scalogram and the coherence phase map (see *Torrence and Compo* (1997)) can be calculated as well.

Conceptually, the SWFT scalogram is produced via the following:

$$F_{pq} = C \cdot F([x(j_{start} : j_{end}) - D] \cdot T)_k \quad (2.4)$$

where p and q are indices defined mathematically below, which correspond to the periods and times at which the SWFT transform is applied. $F(\dots)_k$ indicates a single value from the discrete Fourier transform spectrum, x is the time-series dataset, j_{start} and j_{end} are the beginning and ending indices of the current data window. T is a tapering function (optional), and D is a detrending function (optional), each defined over the current data window. C is a multiplicative correction factor unique to each type of tapering function that is computed as $C = |F(\sin(-\infty : +\infty))_k| / |F(\sin(-\infty : +\infty) \cdot T)_k|$ where k is the index that corresponds to the period 2π . Note that a reasonable approximation of C can be obtained using just three or four cycles of the sine function. If a tapering function is not used, $C = 1$. For our analysis, we chose a Tukey window with a gradual taper (Tukey window coefficient of 0.5) as a tradeoff between frequency resolution and side lobe distortion, resulting in $C \approx 1.33$. Tukey coefficients closer to 1 produce greater side-lobe distortion, while those nearer to 0 reduce side-lobe distortion but increase leakage.

The complex definition of the discrete Fourier transform (DFT) modified from *Press et al.* (1992) is

$$F_k = \sum_{j=1}^n x_j e^{-\frac{2i\pi}{n}(k-1)(j-1)} \quad (2.5)$$

where x is the time-series dataset, and k and j are indices running from $k = [1, \dots, n]$ and $j = [1, \dots, n]$ and n is the total number of data points to be transformed. The

DFT spectrum F has n points of which the first is the "gain" term, and the next $n/2$ points correspond to the periods $n/f \cdot [1, 1/2, 1/3, \dots, 2/n]$.

To compute the SWFT, we would like to extract a single Fourier coefficient corresponding to a certain period, P_p , from the entire DFT spectrum. Also, we need the coefficient not for the entire dataset x , but for a windowed subset x_j with indices $j = [nstart, \dots, nend]$ (which relate to times $t = j/f$) where the total number of points in the window is N_p . To minimize leakage, the window width N_p should be chosen such that an integer number of sinusoidal cycles fit within it, which can be expressed as

$$N_p = P_p \cdot f \cdot w. \quad (2.6)$$

Here w is an integer multiple with possible values $[1, 2, \dots, \text{floor}(n/P_p \cdot f)]$ where "floor" is a function that rounds the argument toward zero. In general, higher values of w increase the frequency resolution of the SWFT while decreasing the time resolution. Note that the entire dataset x is used, and the standard DFT Fourier coefficient is produced when w is set equal to $n/P_p \cdot f$.

When Equation 2.5 is applied to the data window N_p , we note that the indices k in the DFT output then correspond to periods $N_p/f \cdot [1, 1/2, 1/3, \dots, 2/N_p]$, and the first point in the DFT output is the DC gain term. The SWFT requires only the value F_k corresponding to P_p . So $P_p = N_p/f \cdot 1/w$, and therefore (replacing k with p) we get

$$F_k = F_p = F_{w+1} \quad (2.7)$$

Substituting Equations 2.6 and 2.7 into 2.5, and recalling the definition of N_p , we get

$$F_p = \sum_{j=nstart}^{nend} x_j e^{\frac{-2i\pi w}{N_p}(j-1)}. \quad (2.8)$$

The window of width N_p is slid along the dataset producing an array F_{pq} where q corresponds to the center-window time τ_{pq} with indices p and $q = [1, 2, \dots, qmax_p]$

. The maximum resolution of P is such that the product $P_{max} \cdot f = [2, 3, \dots, n]$, since the shortest period allowed is given by the Nyquist critical frequency $(2/f)^{-1}$. However, for computational efficiency when P_{max} spans several orders of magnitude, the user can specify any subset P of P_{max} . For instance, one could choose $P \cdot f = [2, 5, 12, 14, \dots, n]$, or any other arbitrary subset of P_{max} . The index $p = [1, \dots, \text{length}(P)]$ then refers to the periods at which the SWFT will be calculated, where “length” is a function that calculates the size of the 1st dimension of the array P .

Substituting Equation 2.8 into Equation 2.4 yields the complete definition of the SWFT:

$$F_{pq} = C \sum_{j=n_{start}}^{n_{end}} (x_j - D_m) \cdot T_m e^{\frac{-2i\pi w}{N_p}(m-1)} \quad (2.9)$$

where m is an index $[1, 2, \dots, P_p \cdot f_s \cdot w]$, or $m = (j + 1) - n_{start}$ and n_{end} are functions of p and q :

$$n_{start} = (q - 1) \cdot P_p \cdot f \cdot w + 1 = (q - 1) \cdot N_p + 1, \text{ and} \quad (2.10)$$

$$n_{end} = n_{start} + N_p - 1. \quad (2.11)$$

Because the Fourier coefficients can only be time-localized to a window of width N_p , we include the capability for each data window to overlap the previous one in order to increase the temporal resolution of the transform. As an example, if $N_p = 10$ and no overlap is allowed, the minimum temporal resolution at this value of P_p would be 10. Instead overlap is allowed such that the minimum resolution can be as low as 1. This could either be done with a fixed overlap (i.e., each window overlaps half of the other across all periods), or the overlap can be scaled in any other fashion, such as linearly with P . For this study, we define the overlap value, o_p to be given by a simple linear scaling as

$$o_p = \text{floor} \left[o_{min} + \frac{P_p}{\max(P)} (o_{max} - o_{min}) \right] \quad (2.12)$$

where o_{max} and o_{min} are specified by the user, and $\max(P)$ is the maximum value within the array P . The minimum value of $o_{min} = 1$, and the maximum value of $o_{max} = N_p$. With this modification, Equation 2.10 becomes

$$n_{start} = (q - 1) \cdot N_p / o_p + 1 \quad (2.13)$$

and $q(max)_p$ is given by

$$q(max)_p = \text{floor} \left[(n - N_p) \cdot \frac{o_p}{N_p} + 1 \right]. \quad (2.14)$$

The amplitude or power spectra can be extracted from the full SWFT spectrum in the same way it would be for the standard FT. Here we use the amplitude spectrum that is calculated via

$$A_{pq} = 2 \cdot |F_{pq}| / N_p \quad (2.15)$$

where the vertical bars indicate the magnitude of the complex value F_{pq} and the factor of 2 arises because the DFT spectrum is symmetric about the vertical axis and thus distributes half of the spectral power at a period to the each of the positive and negative instances of that period. The center-window times τ associated with the arrays F and A are given by

$$\tau_{pq} = f \cdot \left[(q - 1) \frac{N_p}{o_p} + \frac{N_p}{2} \right] = f \cdot \left[N_p \left(\frac{q - 1}{o_p} + \frac{1}{2} \right) \right]. \quad (2.16)$$

Note that F , A , and τ are, in general non-rectangular arrays (the exception is when $w = N_p$). MATLAB's cell array capability was used to store these values. For convenience of both visualization and further processing, such as contouring or computing cross-scalograms and phase-coherence maps, these arrays can be interpolated to rectangular grids.

The SWFT array F can either be computed directly via Equation 2.9, or it can be computed via Equation 2.4 using the values of j_{start} and j_{end} from Equations 2.13 and 2.11. In either case, the tapering and detrending functions T and D are calculated using the data $x(j_{start} : j_{end})$.

The SWFT was developed primarily for flexibly visualizing the non-stationary spectral content of a time-domain signal. Including the ability for ω_p to scale with period greatly reduces the computational demand by calculating F_{pq} less frequently for longer periods. Allowing w to vary enables flexibility between time- and frequency-localization, as the needs of the user demand. Though not used here, w could also be a function of p allowing the frequency resolution to also scale with the period. Including explicit tapering and detrending further improves the ability of the SWFT to represent dynamic spectral content when P spans several orders of magnitude.

In order to illustrate the interpretation of the SWFT scalogram, we examine a simple test case using a summation of three separate sinusoids given by: $f(x) = f_1(x) + f_2(x) + f_3(x)$, where $f_1(x) = \sin(4x)$ over $0 \leq x \leq 12\pi$, and $f_2(x) = \sin(10x)$ over $0 \leq x \leq 12\pi$, and

$$f_3(x) = \begin{cases} 0 & 0 \leq x \leq 4\pi, 8\pi \leq x \leq 12\pi \\ \sin(x) & 4 < x < 8\pi \end{cases}.$$

Figure 2.2a illustrates the successive superposition of the three sinusoids, the darkest curve plots $f_3(x)$, the mid-tone curve plots $f_3(x) + f_2(x)$, and the lightest curve plots $f(x)$.

The SWFT scalogram (Figure 2.2b) of the function $f(x)$ reveals the periods, amplitudes, and ranges of activity of each of the simple sinusoids. The method extracts the periods of the three sinusoids and reconstructs the amplitudes accurately, except for the $f_3(x)$. This is simply because only two cycles of the sinusoid were used in $f_3(x)$, and the window width was chosen as twice the period, thus only the very center point should reach an amplitude of 1. The shorter-period sinusoids both have peak amplitudes very near 1, although the middle has amplitudes >1 in some locations because of the interaction between the two longer-period sinusoids. Importantly the longest period sinusoid, which is defined only for $4\pi < x < 8\pi$, only has large amplitudes in this range, thus revealing the time-varying spectral behavior of the input signal.

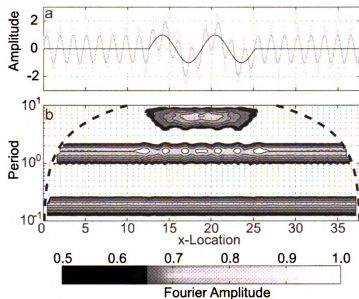


Figure 2.2: SWFT results for a test function, $f(x)$. a) Curves representing the successive superposition of the three sinusoids with different frequencies, and b) the SWFT scalogram. The shaded contours represent amplitudes between 0.5 and 1.0 as shown in the colorbar. The dashed line represents the boundary beyond which the SWFT scalogram is undefined (where half the window width is greater than the available number of points within the dataset).

In cases where the input is something other than a summation of sinusoids, the scalogram output will exhibit large amplitudes across a range of periods, as expected from Fourier theory. If the broad-spectrum behavior of the input data is relatively time-localized, such as the runoff response to a brief storm event, the scalogram will exhibit lineations corresponding to that event. Thus, temporally continuous but spectrally limited high amplitude regions indicate periodic processes within the input data while temporally limited but spectrally broad lineations indicate broad-spectrum processes.

2.4.6 Scalogram Averaging

Averaging the scalogram amplitudes across periods or time gives the period-averaged or time-averaged amplitude spectra, respectively, of the SWFT scalogram (*Torrence and Compo*, 1997). All of these three averages are computed using a rectangularly interpolated grid, A_{sr} , at user-selected periods s and times r , from A_{pq} . Since the period spacing in the SWFT scalogram is not necessarily uniform, we use the period-weighted mean amplitude Ap , calculated using

$$Ap_r = \frac{\sum_{s=1}^L P_s A_{sr}}{\sum_{s=1}^L P_s}, \quad (2.17)$$

where L is the dimension of the rectangular grid in the period direction. This provides aggregate information on the temporal variation of spectral amplitude.

The time-averaged amplitude At , given by

$$At_s = \frac{1}{M} \sum_{r=1}^M A_{sr} \quad (2.18)$$

where M is the dimension of A in the time direction, displays average amplitudes across time at a single period. Referred to as the global SWFT spectrum, this time-averaged spectrum is qualitatively similar to the FT spectrum, but differs in physical meaning.

13

14

15

21

22

23

24

25

26

27

28

29

30

31

32

33

34

Bulk changes in the relative influence of short- vs. long-period fluctuations across time can be visualized using the amplitude-weighted average period Pt ,

$$Pt_r = \frac{\sum_{s=1}^L P_s A_{sr}}{\sum_{s=1}^L A_{sr}}. \quad (2.19)$$

All three of these scalogram averages are demonstrated below.

2.4.7 Watershed Available Precipitation: Snowmelt Modeling

Although this study focuses on revealing and exploring watershed process using a purely data-driven SA, one model is required for the analyses. Because precipitation falls as snow during most of the winter months in northern lower-Michigan, a snow storage-and-release model is needed. The data required for a full energy/water balance model were either unavailable or unreliable for the entire data period, thus we implemented a simple heuristic snowmelt model. Because data are available for both fresh snow totals and observed snow depth, the model can simply identify when a decrease in snow pack thickness corresponds to melting event or densification of the snow pack. We then compared one year of the output from the heuristic model to the full energy-balance UEB snow model (*Tarboton and Luce, 1996*) in order to assure that it acceptably predicts both snowmelt timing and volume.

This heuristic model tracks snow water equivalent and releases snowmelt based on a three-part rule structure:

1. The density of newly-fallen snow is calculated from daily precipitation and snow fall totals. Since precipitation can be mixed frozen and liquid, a maximum new-snow density cutoff, new_{max} , was determined from the data. Any precipitation in excess of this cutoff is considered equivalent to rainfall and immediately released.
2. The total snowpack density is updated based on the new snow depth and the

accumulated water content of the pack.

3. If the average snowpack density exceeds the maximum pack density, $pack_{max}$, it is assumed that some of the snow has melted. Melt is then generated equivalent to the depth of the pack multiplied by the difference between the calculated pack density and the maximum density.

There are three important assumptions in this model. First, we assume that drifting does not affect the observed snow depths at the measurement location (typical of snow models). Second, we assume that the snow pack properties are uniform throughout, which is realistic in the MRW region as total pack thicknesses are typically less than a third of a meter. Finally, this model assumes that the density at which the snow pack releases water remains the same throughout the season.

The maximum densities of the new snowfall and the snowpack are physical quantities that are generally not equal. Maximum snowpack density, $pack_{max}$, is determined by both its composition and water holding capacity, which are functions of the thermal history of the pack and new snowfall conditions. Direct comparison of snow depth and stream discharge in our study area suggests that water is released from the snowpack when the combined snow/water density reaches approximately 0.35 g/cm³, according to this model, thus this value was chosen for $pack_{max}$. The maximum new snowfall, new_{max} density of 0.23 g/cm³ (determined as the ratio of new snow water content to new snow depth) was extracted from the data, as this was the relatively abrupt limit above which higher-density new snowfall events were obvious outliers.

The model was tested relative to the UEB snow model using data from the winter of 1999/2000. The solid line in Figure 2.3a is the heuristic snowmelt model, and the dashed line is the output from the UEB snow model. Both models output the combined snowmelt and liquid precipitation, referred to hereafter as watershed-available precipitation. Watershed-available precipitation is that which can be acted upon by physical or biological processes. Note that the models provide very similar

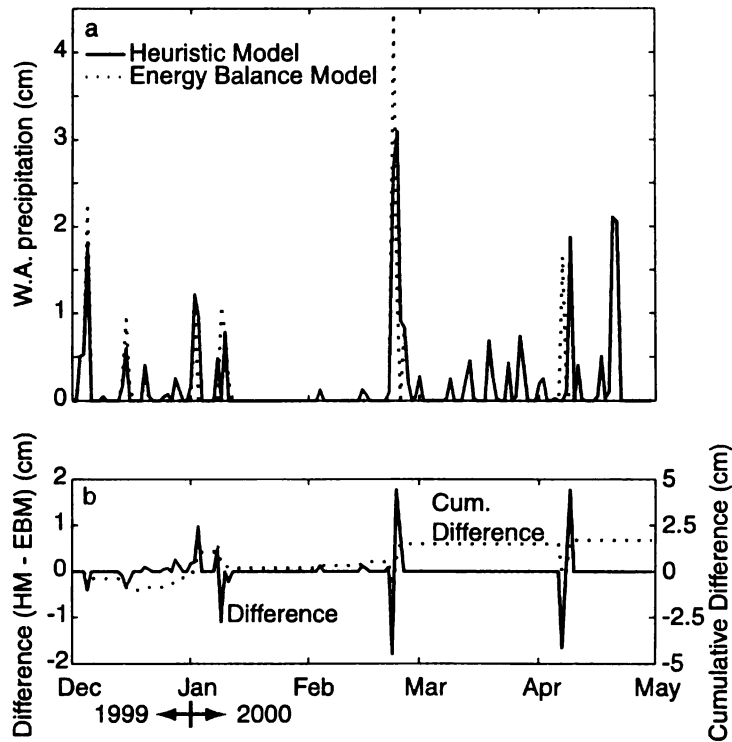


Figure 2.3: a) Comparison of the heuristic snowmelt model to the UEB model, b) the difference between the two (heuristic-UEB) and the cumulative difference.

results, with the UEB model predicting slightly more melt early in the season and the heuristic model predicting more late in the season. The heuristic model predicts approximately 1.5 cm more snowmelt during the season than the UEB model. However, the heuristic model does not allow for sublimation, which entirely accounts for the ~ 1.5 cm difference at the end of the modeled period. The heuristic model was chosen for our analysis since it performs acceptably, despite minimal data requirements and a simple structure.

2.5 Results and Discussion

2.5.1 Examining Watershed Processes: Spectral Comparison

The spectra of stream discharge, watershed-available precipitation, and relaxation of the water table elevation in well B13 can be directly compared to infer interaction timescales between the data types and to examine details of processes within each type. Well B13 (data not plotted) was chosen because the water table is deep enough (>30 meters) to preclude direct evapotranspiration effects. Four key features of the spectra will be compared to integrate the hydrologic datasets and reveal process details: 1) spectral peaks, 2) log-log linear slopes (i.e., fractal scaling), 3) locations of slope-breaks, and 4) relative spectral amplitudes. Since watershed-available precipitation is the primary forcing function for natural watershed processes in our study region, the spectra of well-head relaxation and stream discharge can be viewed as modifications, or fractal filters (*Kirchner et al.*, 2000), of the watershed-available precipitation spectrum. The same is true to some degree for the well-head relaxation and stream discharge spectra, as groundwater inputs to the Muskegon River account for a majority of its annual discharge (*Jayawickreme and Hyndman*, 2007).

The spectra of these three data types are plotted in Figure 2.4 along with log-log linear fits and 95% confidence intervals. Also plotted is fixed period-width binned spectrum to aid visualization. Slopes for selected linear portions of the spectra were calculated using least-squares regression between user selected bounds. These bounds were chosen to match portions of the spectrum that exhibited a linear slope. The slope breakpoints are then calculated at the intercepts of the separate linear fits.

The 95% confidence interval (dotted line) is determined by multiplying the χ -square value for a system with 1 degree of freedom by the average amplitude given by a noise (or scaling) model (*Torrence and Compo*, 1997). In this case, the noise model was assumed to be given by the linear fits. This enables the flexibility of applying

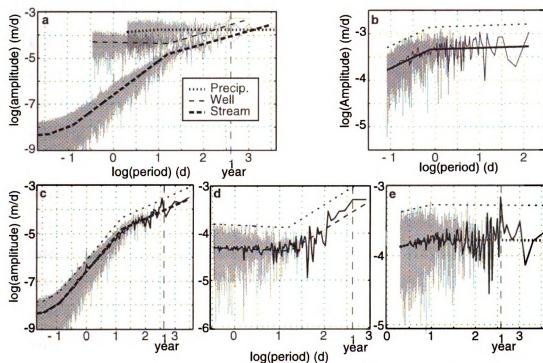


Figure 2.4: Composite plot of stream discharge, well-head relaxation, and precipitation spectra. Grey lines plot the raw FT spectra, solid lines are the binned-mean amplitudes, and dashed lines give the upper 95% confidence intervals. a) An overlay of the spectra in parts c-e and their linear fits, b) NEXRAD precipitation spectrum for 4/1/2004-11/30/2004, not plotted on part a, c) stream discharge spectra for 10/1/1997-9/30/2004, d) well-head relaxation for well B13 from 7/1/2003-7/1/2005, and e) watershed-available precipitation from 1/1/1990-1/1/2001. Note that, in part c the limited resolution makes it appear as if much of the stream discharge spectrum is above the confidence limit, however fewer than 5% of points exceed the limit in a spectrum with over 130,000 points.

statistical confidence intervals to datasets without assuming a priori a particular type of noise. This is useful when working with spectra that exhibit multi-scaling behavior (*Tessier et al.*, 1996; *Dahlstedt and Jensen*, 2004), where a single-scaling noise model, and therefore confidence test, would be inadequate.

All three process spectra have annual-cycle peaks at or near 365 days, although the peak in the head relaxation data is significantly below the 95% CI boundary. This is probably due to the relatively short data record available for the head relaxation. The discharge spectrum has a series of peaks at the harmonics of the 1 day peak that are artifacts, as later discussion will demonstrate. The head relaxation spectrum also has a peak near 117 days, along with weaker peaks near 70 and 50 days (Figure 2.4d). These are near the integer factors 3, 5, and 7 of the 365 day peak, again suggesting a spectral artifact rather than processes acting at these periods. The watershed-available precipitation spectrum has two additional peaks at ~ 174 days, and another at ~ 65 days. These may not indicate sinusoidal processes active at those periods, but may instead be the spectral signature of a non-sinusoidal process characterized primarily by a longer-period oscillation. In particular these two peaks are near the integer factors 2 and 6 of the primary 365 day peak.

The slopes of the spectra and slope breaks (Table 2.2) provide provocative evidence of linkages between watershed processes. If a quantity, such as stream discharge, is being forced directly by another, such as precipitation, then fractal scaling active in the forcing input should exhibit itself directly in the response variable (*Tessier et al.*, 1996). However, a non-linear system response behaves like a fractal filter, modifying the input scaling behavior (*Kirchner et al.*, 2000). There are two examples of this in Figure 2.4 and Table 2.2. The first is the segment of the discharge spectrum between one and three hours with a β of 1.5. This is roughly similar to the slope of the NEXRAD hourly precipitation spectrum (data provided by *Andresen* (2004)). The Slope break locations and β values for Figure 2.4. $\beta = 0.9$ slope in the

Table 2.2: Slope break locations and β values for Figure 2.4.

<i>Discharge</i>		<i>Watershed-Available Precip</i>		<i>Head Relaxation</i>		<i>Head Fluctuation</i>	
<i>Break</i>	β	<i>Break</i> ¹	β^1	<i>Break</i>	β	<i>Break</i> ²	β^2
1.0 hour	0.3						
3.0 hours	1.5		0.90				0.03
15.4 days	2.9	18.2 hours	0.30		-0.08	14.0 hours	1.9
	1.1	9.6 days	0.01	16.5 days	1.2	19.3 days	4.4

Note: ¹These bold italicized entries correspond to results from the NEXRAD dataset, normal entries refer to the watershed-available precipitation from gage data.

²These italicized entries correspond to slopes from Well B13, whose binned-mean spectrum is unplotted.

NEXRAD spectrum is an underestimate of the true spectral slope, given that the hourly NEXRAD spectrum represents just a single year of data that undersamples precipitation, and thus suffers from high-frequency aliasing and slope-flattening as described in *Kirchner* (2005). If, as indicated in these empirical spectra, the “true” spectral slopes of streamflow and precipitation match in this range, we interpret the similarities of slope up to a period of three hours to indicate that direct precipitation is the dominant flow-generation process. Beyond three hours, processes with a different scaling relationship dominate flow.

Process linkages are also apparent between long-period stream discharge and head relaxation spectra. Both spectra have a slope break at approximately 16 days and follow a $\beta = 1.1 - 1.2$ scaling relationship to longer periods. This suggests that variations in groundwater discharge control stream discharge variability at periods longer than approximately 16 days. The exact value of this slope break is approximate because of the gradual transition between linear segments in the discharge spectrum between approximately 10 and 30 days. *Zhang and Schilling* (2004) observed slope breaks in Iowa streams at approximately 30 days. The similarity in slopes between head-relaxation and discharge in Figure 2.4 suggests that groundwater inputs dominate streamflow in these Midwestern streams for periods longer than 10-30 days while in-stream and near-surface watershed processes appear to dominate at shorter

periods.

The spectral slope values also provide useful information. Particularly interesting is the $\beta = 2.9$ slope seen at periods between about 3 hours and 16 days in the discharge spectrum. The uniformity of scaling in this portion of the spectrum means that any watershed processes active in this period range also exhibits similar scaling. Fundamentally, this is because linear processes that are additive in the time-domain also add in the spectral domain (*Smith*, 1997). If the scaling of any hydrologically significant watershed process differed from the others, it would preclude the observed uniform scaling. This uniformity is interesting considering the variety of watershed processes active in this period range, including bank storage and release, precipitation runoff, near-surface “interflow”, near-stream saturated groundwater response, and evapotranspiration. Further investigation of this uniformity could be undertaken with a more concentrated set of data designed to explore these processes, or using a detailed process-based hydrological model.

Analysis of the relative amplitudes of the three spectra in Figure 2.4a provides a hydrologic response time to precipitation. The amplitudes of watershed-available precipitation and discharge converge at a period of approximately 2.5 years. Thus any long-period fluctuation in precipitation will be followed by an equal magnitude fluctuation in stream discharge. Therefore this watershed and its unsaturated and saturated groundwater processes do not appear to control hydrologic fluxes with periodicities greater than 2.5 years. This observation could be used in autoregressive models of basins with short discharge records but more extensive precipitation data. Another interpretation of the similarity in magnitude between the precipitation and discharge spectra is that the combined response time of both surface and groundwater systems is approximately 2.5 years under current conditions, although extended drought periods could certainly affect this value. Therefore, in order to assure insensitivity to initial conditions, a model of this watershed must be “spun up” with

realistic meteorological inputs for at least 2.5 years.

Beyond 10-30 days, the well-head relaxation amplitudes are approximately 2-3 times greater than those of stream discharge. This disparity is likely a combination of both physical processes as well as our assumptions. First, the well-head relaxation amplitudes would exceed those of stream discharge because evapotranspirative losses. Also, the head in an individual well may fluctuate more than the average of all wells in the watershed, particularly if the water table at that well is deep. Thus a more representative comparison would be between stream discharge and the average of spectra from wells distributed across the watershed. However, in this case many of our wells showed signs of periodic anthropogenic disturbance that violates the assumptions of our simple differencing technique. Other factors that contribute to the disparity between head relaxation and stream discharge amplitudes may include overestimation of porosity or average recharge rate because of the short data record in the well, as well as differences in annual recharge between the Evart watershed and the location of the well in the Grand Traverse Bay region.

The head relaxation results presented in Table 2.2 are taken from the spectrum of smoothed-differentiated head fluctuations shown in Figure 2.4d while the head fluctuation data in Table 2.21 were taken from the spectrum of actual head fluctuations (not plotted). Unlike head fluctuations, head relaxation (and therefore a decrease in storage) is directly related to groundwater discharge, allowing direct comparison of their spectra. Other studies have reported head fluctuation spectra, (*Zhang and Schilling, 2004; Lee and Lee, 2000; Naff and Gutjahr, 1983*), but because the basic units of head fluctuation and stream discharge differed in these studies, their reported groundwater head and baseflow scaling laws can not directly be related.

2.5.2 Examining Intra-Annual Spectral Variability Using the SWFT

Physical interpretations of Fourier transform spectra assume that system processes are both non-intermittent and stationary. However, many watershed processes are either inactive for parts of the year (intermittent) or possess different spectral characteristics over successive cycles (non-stationary). Thus, interpreting spectra from many occurrences of a given process is problematic. To overcome this, we use the Scaled-Windowed Fourier Transform (SWFT), which does not require either stationarity or non-intermittency. As Figures 2.5-2.8 demonstrate, all three watershed processes examined in this study exhibit both non-stationary and intermittent processes.

The SWFT spectrum of stream discharge (Figure 2.5) displays its time-varying Fourier spectral content, revealing details about how discharge responds to hydrologic inputs and watershed processes. As described previously, the vertical lineations in the SWFT scalogram are caused by the broad spectrum of time-series stream discharge peaks with each such lineation corresponding to a pulse increase in stream discharge. Clearly separated lineations, which primarily occur during the summer months, extend from very short periods and tend to reach maximum amplitudes at periods between 20-40 days. This range of periods corresponds to the width of the time-series discharge peak, and thus to the time-scale of surface and near-surface watershed response to precipitation inputs. As will be shown in the next section, this 20-40 day time scale of surface and near-surface hydrologic response is also similar to the width of the unit hydrographs developed for this watershed.

Spring months typically exhibit increased Fourier amplitudes relative to other seasons across all periods shorter than several hundred days. The spring of 1998 has a very different spectrum than that of 1996 or 1997. Those years had several spring discharge peaks followed by relatively high summer baseflow levels, whereas the single discharge peak of 1998 is followed by low baseflow and weakening of the 365 day am-

Chen, Y. and J. A. Roberts. 2002. *Journal of Great Lakes Research* 28: 1-12.

Figure 6

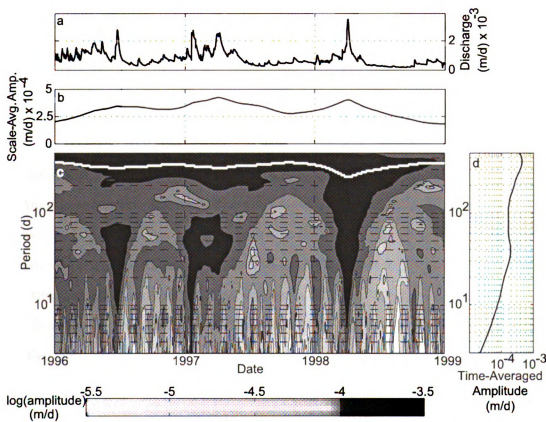


Figure 2.5: SWFT of stream discharge at the USGS gage in Evart, MI. a) Stream discharge time-series divided by the area of the watershed above this gage, b) Scale-averaged amplitudes, A_p , c) Filled contour scalogram of the amplitude (A) vs. both period (P) and time (τ), along with the location of the amplitude-weighted mean period, P_t (white line), d) Time-averaged amplitude spectrum, A_t .

plitude. This summer baseflow portion of the time-series discharge is accompanied by decreased amplitudes across nearly all periods, except near 365 days. Thus, stream discharge during this portion of the year is dominated by long-period fluctuation, most likely from groundwater inputs, as indicated by the amplitude-weighted mean period curve (Pt).

Seasonal differences in spectral character between the summer and spring/fall are not evident in the watershed-available precipitation SWFT scalogram (Figure 2.6c), but show up very prominently in discharge (Figure 2.5c). This suggests that although there is spectral power in that period range in watershed-available precipitation, summer evapotranspiration and canopy interception decrease the magnitude of discharge response and thus the spectral amplitudes.

Another important difference between Figures 2.5c and 2.6c is that the dominant power in precipitation spectrum occurs at shorter periods while the reverse is true for stream discharge. This corresponds to the behavior seen in Figure 2.4c, however the time-averaged amplitude (At) spectrum of watershed-available precipitation differs from the Fourier spectrum in Figure 2.4e. Such differences are artifacts produced by the violations of the assumptions of stationarity and non-intermittency inherent in the standard FT.

The SWFT discharge scalogram for the water years 10/1/1991-9/30/2004 (Figure 2.7) displays the spectral content of stream discharge at the Evart gage over periods of 0.3-1000 days. Although the longer-period spectrum is not plotted, 2-3 year and 6-8 year cycles are evident in periods between 100-400 days, perhaps related to climate cycles as seen in *Coulibaly and Burn* (2004). For 1991-93, long-period fluctuations are most active near a period of 180 days, which then cease during 1994-95 before resuming from 1996-98 at 365 days. Again, this long-period activity switches off for two years and then resumes centered about 180 days. Another interesting set of features in this scalogram are the summer-fall periods of 1998, 2000, and 2002 in

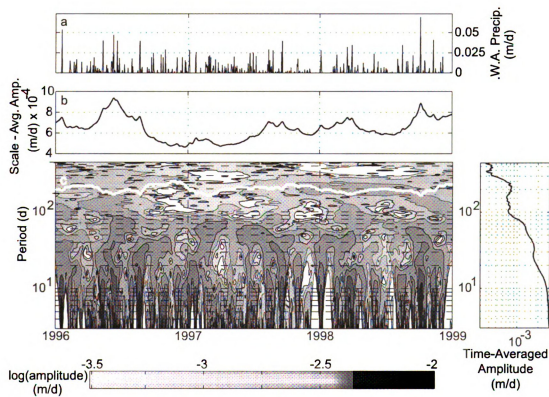


Figure 2.6: SWFT of watershed-available precipitation at Big Rapids, MI. a) watershed-available precipitation time-series, b) Scale-averaged amplitudes, A_p , c) Filled contour scalogram of the amplitudes (A) vs. both period (P) and time (τ), along with the location of the amplitude-weighted mean period, P_t (white line), Note, white areas have amplitudes $< \log(\text{amplitude}) = -3.5$, d) Time-averaged amplitude spectrum, A_t .

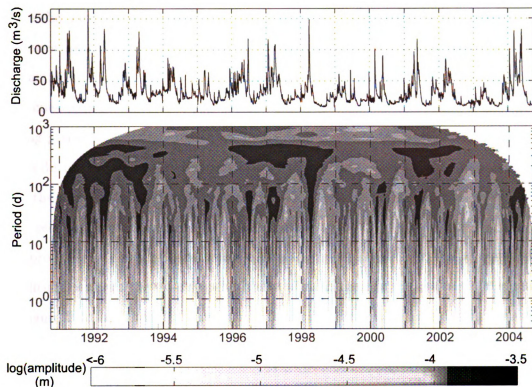


Figure 2.7: SWFT of stream discharge at Evert from 9/1/1991 to 8/31/2004. a) Stream discharge time-series, b) Filled contour scalogram of the amplitude (A) vs. both period (P) and time (τ). The year labels are centered at January 1.

which amplitudes are drastically reduced up to periods on the order of 100-200 days.

These observations from Figure 2.7 enable a deeper understanding of time-averaged spectra such as the FT spectrum or the SWFT time-averaged amplitude (At). The FT spectrum of discharge (Figure 2.4c) contains a weak spectral peak at 180 days along with a peak at 365 days. Prior to examining the spectrum as a scalogram, we were unable to distinguish between dominant spectral peak harmonics and peaks from independent processes at those periods. The scalogram shows that there are processes generating true peaks at both the 365-day and 180-day periods.

Additionally, we can use the information from the scalogram to indicate when processes that generate particular spectral peaks are most active. A prominent example of this is the 1 day peak in the FT spectrum of discharge. Plausible interpretations of

this 1 day peak include diurnal fluctuation in evapotranspiration or streambed conductance during the summer months. However, Figure 2.7 shows that the dominant 1 day amplitudes occur in the winter and early spring months. A close examination of the time-series reveals two important details: 1) the diurnal fluctuation is strongest during periods where daily maximum temperatures are subfreezing, and 2) discharge peaks during the coldest mid-morning hours of each day. These observations suggest that the diurnal signal may be related to icing effects at the instrument. The 1-day peak seen in Figure 2.4 is a “true” spectral peak, but it does not indicate cyclic system processes so much as inaccuracies in measurements. Also, there are no significant amplitude peaks at the 0.5-day period, thus confirming that the corresponding peak in Figure 2.4c was indeed a harmonic. This information could then be used to filter the discharge time-series and remove this apparently erroneous periodic discharge behavior.

The SWFT of head fluctuations from Well B10 (Figure 2.8) provides another example of the importance of viewing spectra as a function of both period and time. Well B10 was chosen because it exhibited the greatest amplitude of fluctuation in the period 1-30 days of the 17 wells in our study. Figure 2.8 displays both the head fluctuation data (mean removed) as well as the SWFT scalogram for periods between 0.3 and 100 days. The largest time-series amplitude head fluctuations occur from late October to early April. The scalogram reveals that most of the time series fluctuation is caused by larger Fourier amplitudes of the 1-30 day periods, which are greatest during the winter. Note that the dominant period in this range is not constant throughout the year. During the summer, early fall, and spring, the dominant period is near the 7-10 day range, that then shifts to the shorter 2-7 day range in mid-November. The SWFT scalogram clearly reveals this information that would be difficult to directly extract from the time-series data.

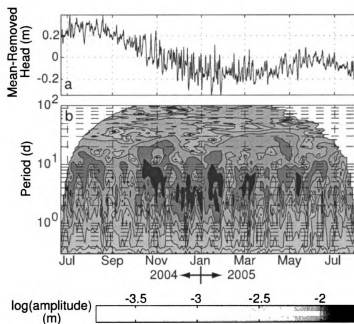


Figure 2.8: SWFT of groundwater-head fluctuations in Well B10 in the Grand Traverse Bay Watershed. a) Well-head fluctuation time-series, b) Filled contour scalogram of the amplitudes (A) vs. both period (P) and time (τ).

2.5.3 Spectrally-Derived Watershed Precipitation Response Functions

The watershed annual unit response functions for the portion of the Muskegon River Watershed above Ewart gauge were calculated using discharge and Big Rapids precipitation data between 9/1/1999 and 8/31/2000 (chosen arbitrarily). Two different unit response functions are shown in Figure 2.9a, that of the discharge response to watershed-available precipitation as well as to raw precipitation. Including the snow storage-and-release model creates higher peak discharge responses with a more physically realistic long tail due to groundwater discharge. The higher peak is expected because without a snow model this method treats the precipitation the same in January as it would in July, even though the January precipitation fell as snow and was stored until later, resulting in no significant short-term discharge response.

A convolution of the solid-line unit response function in 2.9a with watershed-

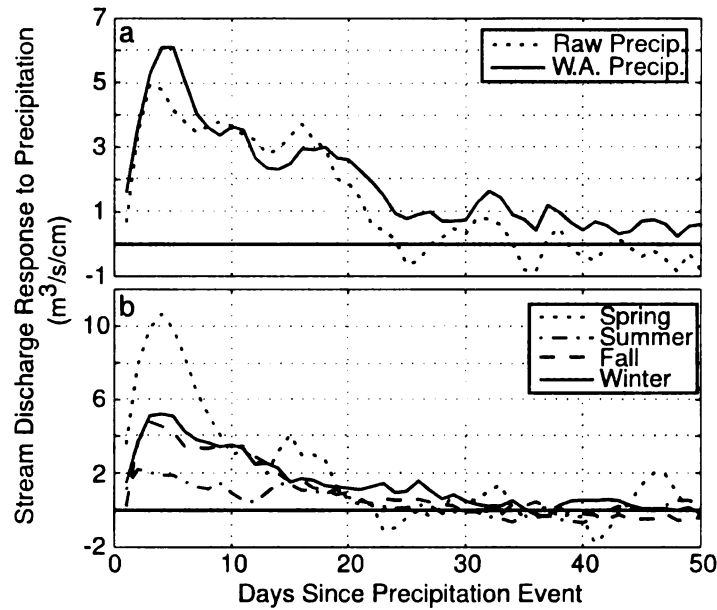


Figure 2.9: Stream unit response hydrographs. a) Unit response functions obtained using data from 9/1/1999 to 8/31/2000, and b) Seasonal unit response functions using selected years (see Table 2.1) from 1990-1999.

available precipitation according to Equation 2.1, produces the dashed modeled discharge and residual curves in Figures 2.10a and b. Because the unit-response was truncated as described in the methods, the convolution is not a perfect reconstruction. The resultant discharge is an overestimate in the summer and fall months, but an underestimate during the spring. This is to be expected as the annually-calculated response curve effectively averages the system behavior throughout the year.

If the unit response is calculated seasonally rather than annually, the resultant set of unit response curves reveals the seasonal differences among runoff responses in this watershed (Figure 2.9b). Discharge response during the spring is much higher than either the annual curve or those of the other three seasons, perhaps due to a combination of frozen soils and higher average soil saturation prior to watershed-available precipitation events (which can be either rainfall or snowmelt). Summer discharge response, on the other hand, is highly damped due to canopy interception, lower average soil moisture, and evapotranspiration. The fall and winter responses

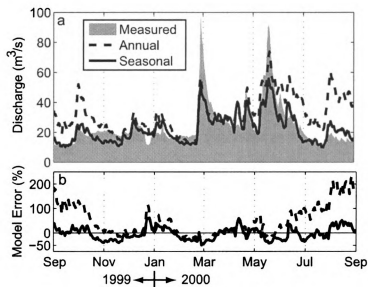


Figure 2.10: Plots of convolved stream discharge from 09/01/1999 to 8/31/2000 using the annually- and seasonally-derived unit hydrographs (initial modeled discharged matched to stream discharge) on top of the measured stream discharge (a), and on top of the measured stream discharge. b) Plot of the percentage difference between each model and the measured discharge $(\text{modeled}-\text{measured})/(\text{measured}) \times 100$.

appear to be very similar, suggesting that there are not large differences in discharge response between these two seasons. This result was somewhat unexpected since frozen soils are generally expected during the winter months due to long periods of sub-freezing temperatures. Significant areas of frozen soils would tend to increase discharge response, as infiltration capacity is greatly reduced. The lack of a response difference suggests that the soils are not homogeneously frozen throughout the winter season, or that this freezing is not important for runoff generation in this watershed during this time period.

Analysis of data from the wells in the GTBW indicates that the delays between peak spring recharge and peak saturated water table response scales approximately as 2-4 days/meter of unsaturated zone depth. Thus, for depths on the order of 30 meters, the delay between full groundwater response to a precipitation (or snowmelt) event can be as much as 120 days. As the data lengths included in the seasonal response

calculations are only on the order of 60-90 days the seasonal unit response curves can not properly represent the groundwater response. The tail in the annually calculated unit response curve of Figure 2.9a is likely a more reasonable representation of the average groundwater response.

In addition to providing more insight into watershed processes than the annually-derived unit response curve, the seasonally-derived curves produce a much more accurate estimate of stream discharge when convolved with watershed-available precipitation (Figure 2.10a). The residuals between the two convolutions and stream discharge (Figure 2.10b) quantitatively demonstrate the improvement in discharge estimation gained by seasonal convolution and consideration of non-stationary system behaviors. Figure 2.10a illustrates the utility of the seasonally-derived watershed unit response curves for providing a very simple means of forecasting discharge response to precipitation events. Because the seasonal curves average watershed responses across significant variability in watershed state properties (such as soil moisture), the seasonally-derived convolution underestimates the largest peaks in the discharge data by less than 50%, while predicting measured flows to an accuracy of $\pm 25\%$ in most other cases. Much of the remaining residual is because the 50-day unit response curves fail to capture much of the groundwater response to spring and fall precipitation. An analysis that accounts for the different seasonal responses between wet and dry years, and explicitly incorporates the full groundwater response, might further improve the accuracy of forecasting with this technique.

2.6 Conclusions

We present the application of three spectral analysis techniques, direct spectral comparison, the Scaled Windowed Fourier Transform (SWFT), and the derivation of the unit hydrograph via FT deconvolution. We have discussed and demonstrated how

each technique requires consideration of limitations and possible pitfalls in order to be applied successfully, and we elucidated a set of best practices in their application. Most importantly, we have demonstrated that these spectral analysis methods can be used to integrate hydrologic data in order to evaluate watershed processes.

The spectra of related data types were directly compared to infer process linkages between hydrologic inputs, watershed processes, and stream discharge. Similarities in amplitude peaks, log-log linear fractal scaling behaviors, and breaks in scaling slopes among datasets indicate the nature of linkages. For example, fractal scaling in precipitation may be matched in the stream spectrum at periods shorter than approximately 3 hours for the Evert, MI sub-basin of the Muskegon River Watershed. From periods of 3 hours to approximately 10-30 days, stream scaling follows a $\beta = 2.9$ slope. This single scaling relationship is notable considering the variety of processes active in this period range, suggesting mathematical similarities among these processes. Beyond 30 days, the scaling apparent in the stream spectrum appears to be controlled by groundwater inputs. But, past a period of approximately 2.5 years, fluctuations in the precipitation spectrum control the stream discharge spectrum. This overall watershed response time should be considered when developing transient predictive simulations of watershed behavior.

We introduced the Scaled Windowed Fourier Transform (SWFT) technique to examine the time-varying content of fundamentally non-stationary hydrologic datasets. The SWFT scalograms revealed both the non-stationarity and intermittency of stream discharge, precipitation, and groundwater head fluctuation spectra. The effect of evapotranspiration and canopy interception is evident in a comparison of the SWFT scalogram of summer precipitation events to the highly damped discharge scalogram for those seasons. Also, the 1-day peak evident in the FT spectrum of stream discharge was shown to be due largely to measurement error rather than diurnal hydrologic processes. Importantly, these are a subset of many possible observations from the rich

set of information contained within the SWFT scalogram.

Using direct FT deconvolution, spectral analysis can be also be used to estimate stream unit hydrographs. Because of the simplicity of this method, temporally- and seasonally-varying hydrographs can be quickly derived to better understand non-stationary watershed processes. For the Muskegon River above Evart, MI, the groundwater dominance of the stream discharge spectrum beyond approximately 15-20 days is confirmed by visual inspection of the main unit response peaks. These peaks show a 15-20 day primary stream response period followed by a long-tailed groundwater response that continues out to at least 50 days. Also, the seasonally-derived unit hydrographs quantitatively reveal decreased discharge responses due to evapotranspiration during the summer months and augmented responses during spring snowmelt.

Acknowledgements

We are grateful for the financial support from the National Science Foundation grant (EAR-0233648), and secondary support from the Great Lakes Fisheries Trust. We thank Dr. Jeffrey Andresen and Dushmantha Jayawickreme for providing and processing the NEXRAD precipitation data used in this research. We also thank the hydrogeology research group at Michigan State University for their suggestions and contributions. Finally, we thank Dr. James W. Kirchner and an anonymous reviewer for their comments.

CHAPTER 3

Evaluating Temporal and Spatial Variations in Recharge and Streamflow Using the Integrated Landscape Hydrology Model (ILHM)

This chapter was published in 2007 in a slightly modified form in an AGU Geophysical monograph, *Subsurface Hydrology: Data Integration for Properties and Processes* (Hyndman, Kendall, and Welty, 2007)

3.1 Abstract

Projections of climate and land use changes suggest significant alterations to the hydrology of the Upper Midwest. Forecasting those changes at regional scales requires new modeling tools that take advantage of increased computational power and the latest GIS and remote-sensing datasets. Because of the need to resolve fine-scale processes, fully-coupled numerical simulations of regional watersheds are still prohibitive. Although semi-distributed lumped-parameter models are an alternative, they are often not able to accurately forecast across a broad range of hydrologic conditions such as those associated with climate and land use changes.

We have developed a loosely-coupled suite of hydrologic codes called the Integrated

Landscape Hydrology Model (ILHM), which combines readily-available numerical and energy- and mass-balance modeling codes with novel routines. This code is used to predict hydrologic fluxes through a 130 km² portion of the Muskegon River Watershed in northern-lower Michigan. We combine GIS maps of the land cover, soils, and sediments with a variety of gaged and remotely-sensed data for this watershed to simulate evapotranspiration, groundwater recharge, and stream discharge from 1990-2004. These estimates are compared to measured stream discharge data to demonstrate the capability of the ILHM to provide reasonable predictions of groundwater recharge with minimal calibration. The results begin to illustrate critical differences in hydrologic processes due to land cover and climate variability, including a demonstration that approximately 75% of precipitation becomes recharge during leaf-off periods while almost no recharge occurs during the growing season.

3.2 Introduction

Land use and climate changes are expected to alter the spatial and temporal distribution of groundwater recharge over the next century (*Bouraoui et al.*, 1999; *IPCC*, 2001). These changes could have far reaching consequences since recharge maintains groundwater supplies that are used as primary drinking water sources, and is critical to stream ecosystem health since groundwater is the main source of streamflow during dry periods. Despite the clear importance of groundwater recharge, its spatial and temporal distribution is generally poorly understood in humid regions. Many hydrologic modeling studies ignore both spatial and temporal variations in recharge rates, either because limited measurements of critical parameters are available, or because existing modeling methods are not adequate to accurately evaluate these variations at the scales of interest. Integration of available hydrologic and landscape data can help improve estimates of historic recharge rates, and can then provide the basis for

evaluating the range of impacts of anthropogenic alterations of the landscape and climate on future hydrological and ecological conditions.

A range of approaches have been developed to estimate recharge rates based on relatively simple analysis of flows and levels in surface water, the unsaturated zone, and the saturated zone, as reviewed by *Scanlon et al.* (2002). These methods include analyzing baseflows or tracer concentrations, developing estimates based on changes in groundwater levels (reviewed by *Healy and Cook* (2002)), and evaluating recharge through the unsaturated zone with lysimeters or well-instrumented field sites. A variety of empirical models have also been developed to estimate recharge across a range of scales (e.g., *Bogena et al.* (2005)), which can provide estimates with varying degrees of reliability and spatial extent depending on the types, quality, and density of the input data (*Scanlon et al.*, 2002).

Numerical models provide a powerful framework to integrate different data types for recharge estimation. Such models can be categorized as lumped parameter models or process-based models. Lumped-parameter semi-distributed models, such as SWAT (*Arnold et al.*, 1993) and TOPMODEL (*Beven and Kirkby*, 1979) have parameters that can be adjusted to fit measurements but can not be independently measured. As a result, such models tend to have difficulty either predicting flow in a new system without independent calibration, or projecting likely changes in a currently modeled system due to changes in factors including climate and land cover.

Process-based codes such as MODFLOW for groundwater flow are based on fully distributed parameters such as hydraulic conductivity, which can be independently measured based on laboratory analyses or using field evaluations such as pump or slug tests. Unfortunately most groundwater codes are not designed to estimate recharge rates because they do not incorporate important landscape and unsaturated zone processes that are critical to redistribution of precipitation from the soil surface and the vegetation canopy.

To address this limitation, several codes have previously been developed to link MODFLOW or other groundwater codes to landscape or watershed codes that incorporate aspects of the hydrologic cycle beyond groundwater flow. For example, MODFLOW has been linked with SWAT (*Sophocleous and Perkins, 2000*) and HSPF (*Said et al., 2005*), which are both lumped-parameter codes. A new Variably Saturated Flow (VSF) package was developed as a MODFLOW module (*Thoms et al., 2006*) to add unsaturated zone and overland flow processes to groundwater flow simulations, but the data and computational requirements appear to be too great for large watershed simulations based on our analysis of this code for the watershed presented in this paper.

A variety of process-based models have also been developed to simulate fully coupled surface water, and variably saturated subsurface flow. Such codes, which include SUTRA3D (*Voss and Provost, 2002*), Mike-SHE (*Refsgaard and Storm, 1995*), WASH123D (*Yeh and Huang, 2003*), MODHMS (*HydroGeoLogic, 2002; Panday and Huyakorn, 2004*), and InHM (*VanderKwaak and Loague, 2001; VanderKwaak, 1999*), provide powerful tools to examine complex interactions between flow and transport across the range of natural conditions observed in the surface and subsurface. Unfortunately, the data requirements and significant computational demands have generally limited the use of these codes to simulate flows through fairly small domains.

In this research, we develop a new Integrated Landscape Hydrology Model (ILHM) to integrate widely-available hydrologic and landscape data in a synergistic and computationally efficient manner to assess temporal and spatial changes in important hydrologic processes. Since the focus of this monograph is data integration in hydrology, we begin by describing the watershed that we chose for testing and development of the code along with the available hydrologic and landscape data used in this simulation. This is followed by a detailed description of the model development and results.

3.3 Methods

3.3.1 Study Region: Cedar Creek Watershed

The Cedar Creek Watershed, in southwestern Michigan (Figure 3.1), was chosen as a site to test the ILHM since it is one of our main field sites in an ongoing eco-hydrological monitoring and modeling study. Cedar Creek flows through the lower half of the Muskegon River watershed (7,052 km²), where urbanization of previously agricultural and forested landscapes is projected to increase runoff volumes and the associated solute transport over the next 35 years based on an empirical model (*Tang et al.*, 2005). The spatial distribution of land uses within the Cedar Creek Watershed facilitates evaluation of differences in recharge associated with land cover types since the upstream portion of this area is dominated by agriculture while the downstream portion is predominantly forested (Figure 2a). The quaternary geology ranges from medium and coarse-textured glacial tills that drape the northern watershed, to glacial outwash and lacustrine sand and gravel in the central and southern watershed (Figure 3.2B).

The groundwater source area, which we call a groundwatershed, of Cedar Creek was delineated using a two-layer groundwater model of the region encompassing the Muskegon River Watershed (Figure 3.1). The groundwatershed (~ 130 km²) was used in addition to the surface watershed (~ 100 km²) for this study because regional modeling of the Grand Traverse Bay Watershed in Michigan by (*Boutt et al.*, 2001) indicated that surface-and groundwatersheds can differ significantly. The regional Muskegon River groundwater model was developed by expanding the watershed boundaries to significant hydrologic features (i.e., the next large stream or lake beyond the surface watershed) to avoid this issue at regional scales (Figure 3.1). The groundwatershed boundary does fluctuate somewhat with both seasonal and long term climatic variations, but for simplicity in this study we have defined the groundwatershed using the

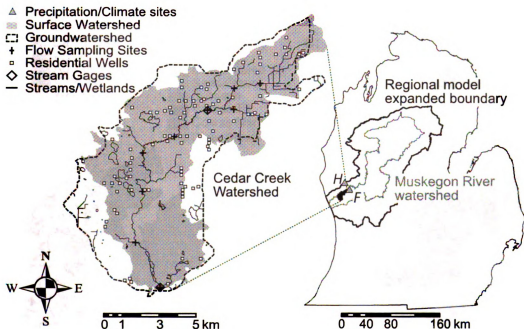


Figure 3.1: On the left is a map for the Cedar Creek watershed (shaded) along with the groundwater contributing area to Cedar Creek (dashed outline). Also displayed on this map are locations of two stream gages, nine discharge measurement cross sections, and residential drinking water wells located within the watershed. On the right, a map of the lower peninsula of the state of Michigan shows the Cedar Creek watershed within the greater Muskegon River watershed. The boundary of a regional groundwater model of the Muskegon River is shown in bold on the state map. The three precipitation and climate gage locations are Hesperia (H), and Fremont (F).

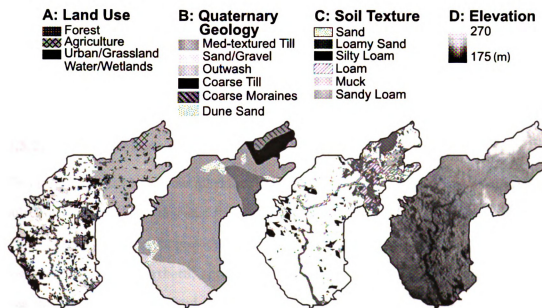


Figure 3.2: Static GIS datasets that were integrated into the ILHM framework. From left to right, the datasets are land use/cover from IFMAP, quaternary geology from (Farrand and Bell, 1982), SSURGO soil textures, and land surface elevation from the NED 26.5 m DEM.

steady-state model.

3.3.2 Data Collection and Analysis

Before constructing the groundwater model for the expanded MRW region and the ILHM to define the Cedar Creek groundwatershed, we assembled the available landscape, hydrology, and climate data for the region into a geodatabase. In many parts of the world, the types of data used for this analysis are commonly available as a free download from internet sites. However, supplementary data such as flows and water levels beyond those available from the US Geological Survey will often need to be collected for model calibration or optimization.

3.3.2.1 Hydrologic Data

Two pressure transducers installed in Cedar Creek recorded stream stage at hourly to sub-hourly intervals (*Wiley and Richards*, 2006) from mid to late 2002 through 2004. These surface water levels provide critical information for this study. One transducer was installed in the northern, agricultural portion of the watershed, while the other was installed in forested land near the watershed outlet (See Figure 3.1). Stream discharge measurements (*Wiley and Richards*, 2006) were used to construct rating curves between stage and discharge. The stage discharge relationships were developed using pairs of measured streamflow paired with water levels from the transducers. For the upper watershed site 19 measurement pairs were used, while 23 measurement pairs were used for the lower watershed site.

Groundwater levels for this region were collected from the Michigan Department of Environmental Quality (MDEQ) residential well database, as no monitoring well data are available in this region except at our surface water-groundwater interaction site adjacent to Cedar Creek. Unfortunately, the wells at this site are too close to the stream to provide useful information about groundwater levels for this watershed-

scale model testing. Observations were available from 99 wells installed across the watershed during the simulation period in the MDEQ database. For each well, one static water level measurement was taken by the well driller at the time of installation. The static water level measurements were used in a preliminary calibration of an early version of the Cedar Creek groundwater model, but there is a significant amount of error associated with these water level measurements since a variety of methods were used by well drillers to identify the location of the well, the elevation of the ground surface, and the depth to water. Residential wells were not included in the model as extraction wells since most of the extracted water is assumed to return via septic systems and the remainder is assumed to be a very small component of the water budget for this region. There are no known irrigated agricultural areas within the Cedar Creek watershed.

3.3.2.2 Geologic, Landscape, and Remote Sensing Data

We established a GIS database for the Cedar Creek region with topography, land use, hydrography, and hydrogeology characteristics. These GIS datasets were compiled from the Michigan Geographic Data Library, established by the Michigan Department of Environmental Quality. These datasets are assumed to be static for the purposes of this analysis.

Land surface elevations (see Figure 3.2D), which were defined based on the National Elevation Dataset 26.5 m digital elevation model (DEM), were used for a range of model inputs. Flow direction and gradients for the overland flow and near surface soil moisture redistribution modules were calculated from the full-resolution DEM and then upscaled. This 4x upscaled DEM was used to set the land surface elevation for the soil water balance model. The drainage network and lake boundaries were defined based on the Michigan Framework GIS dataset, with stream crossing elevations manually extracted from a digitized version of the USGS 1:24000 topography quad-

ranges. The watershed does not contain any large lakes that are connected to the stream system, thus lakes are not separately considered in the groundwater portion of the integrated model.

The land cover distribution across the Cedar Creek model area (Figure 3.2A) is taken from the Integrated Forest Monitoring Assessment and Prescription (IFMAP) coverage, which is a statewide digital land cover map with 30 m resolution derived from 1997-2000 LANDSAT data (MDNR, 2001). This watershed is dominated by forested/openland/wetland land covers (60%), while 36% of the area is agricultural and the remaining 4% is urban. The IFMAP coverage provides inputs that are used in calculating evapotranspiration and overland flow. Land cover types are associated with transpiration estimates through a variety of terms including stomatal conductance, canopy height, and root depth, and with evaporation through changes in canopy interception, wind speed and interception of incoming radiation. Overland flow is also associated with land cover through Manning's roughness coefficients. The details of these connections are included in the modeling sub-sections below.

Hydrogeologic zones were parameterized according to a Quaternary Geology coverage of *Farrand and Bell* (1982) for the groundwater model along with the SSURGO soils database (see Figure 3.2C), which was then mapped into saturated and unsaturated zone parameters according to lookup tables based on literature values. The geometry of the aquifer base was interpolated between measured bedrock elevations by de-clustered and polynomial-detrended simple kriging of the elevations of the drift/bedrock contact from oil and gas wells across the entire expanded Muskegon River model domain. Initial hydraulic conductivity values were assigned to the geologic zones based on an optimization of these parameters for the nearby Grand Traverse Bay watershed that has the same geologic zones (*Boutt et al.*, 2001).

GIS grids of leaf area index (LAI), the ratio of one-sided green leaf area to ground area (*Myeni et al.*, 2002), were also used in calculations of potential evapotranspira-

tion (PET), canopy interception, and solar radiation interception. For this study, we used remotely-sensed LAI measurements from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) eight-day averaged product. Spatially-averaged LAI for both forest and agricultural land-use types are plotted for 2003 and 2004 in Figure 3.3A. In cases where these data were not available (i.e., prior to 2000), we average all LAI grids for each Julian day and apply these multi-year averages to the earlier periods. This will have little effect on our results because we are only comparing simulated and observed flows for mid 2001 through 2004 when all datasets are available. However, it is important to spin up the model using realistic data inputs because we found that it takes between two and three years before the model results are independent of the starting conditions.

3.3.2.3 Climate Data

Precipitation data was obtained from the NOAA gage at Hesperia, MI approximately 20 km NNW from the center of the Cedar Creek watershed (see locator map in Figure 3.1). This gage was chosen because lake effect precipitation is an important meteorological phenomenon in this area, and this gage lies at relatively the same distance from the Lake Michigan shoreline as the Cedar Creek watershed. NOAA data (shown in Figure 3.3B) included hourly precipitation totals, as well as daily measurements of new snowfall and snow pack depth.

Other climate data, including hourly temperature, relative humidity, wind speed, and incoming solar radiation (Figures 3.3C-F), were extracted from the Fremont, MI station of the Michigan Automated Weather Network (MAWN) (see Figure 3.1 for locations). This climate network is operated by the Michigan State University Extension, the Michigan Agricultural Experiment Station and the Michigan Department of Agriculture. Since the MAWN data did not exist prior to 1996, from 1990-1995 we used the Julian-day average of the available data.

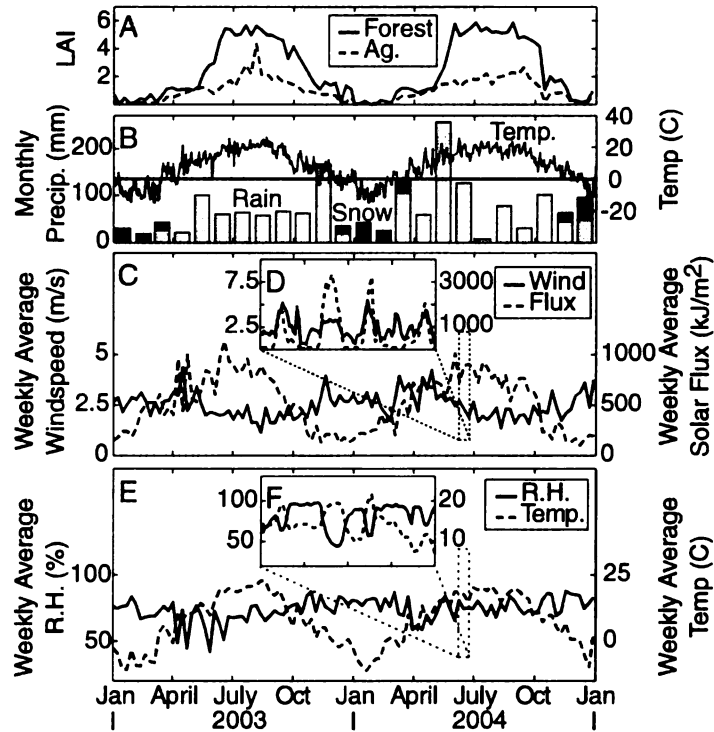


Figure 3.3: Time series data inputs to ILHM for 2003 and 2004: A) LAI of forest and agricultural land covers; B) monthly rain and snow along with daily average temperature, a horizontal line at 0°C is included as a visual aid; C) weekly averaged windspeed and solar flux; D) inset of hourly windspeed and solar flux data from 6/21/04 through 6/24/04; E) weekly averaged relative humidity (R.H) and temperature; F) inset of hourly R.H. and temperature from 6/21/04 through 6/24/04.

3.3.3 Components of the Integrated Landscape Hydrology Model (ILHM)

Figure 3.4 illustrates our conceptual model of the most important hydrologic processes in the Cedar Creek watershed, and diagrams the linkages between input datasets, ILHM modules, and model outputs. As mentioned earlier, this version of ILHM was developed by linking a novel landscape water balance model with a simple linear - delay unsaturated zone model and MODFLOW-2000 (Harbaugh *et al.*, 2000), the most commonly used groundwater flow code. The landscape and near-surface portion of the ILHM combines several existing codes with a set of new modules, in order to speed development and to incorporate the full range of hydrologic processes. The canopy water balance model is based on equations published in Chen *et al.* (2005a). The surface hydrology model, including infiltration and runoff routing, are modified from the Distributed model for Runoff, Evapotranspiration, and Antecedent soil Moisture (DREAM) model by Manfreda *et al.* (2005). The snow pack is simulated using the UEB Snow Model by Tarboton and Luce (1996). Soil moisture accounting along with near-surface flows are handled by a set of codes we developed based on common unsaturated zone flow modeling methods.

The ILHM suite calculates each term in the full water balance equation:

$$\Delta S = P - T - E - Pc + Tr - Ex - R \quad (3.1)$$

where ΔS is the change in surface soil moisture storage, P is watershed available precipitation, T is transpiration, E is evaporation, R is precipitation excess runoff, and Pc is deep percolation beneath the root zone, Tr is lateral near-surface unsaturated flow called throughflow, and Ex is the exfiltration from each cell. For Equation 3.1 and the detailed equations presented in Appendix A, terms are in units of meters per unit time, unless specified otherwise.

The landscape portion of our model sequentially calculates the water balance along the paths water takes as it is redistributed from precipitation to various subsurface

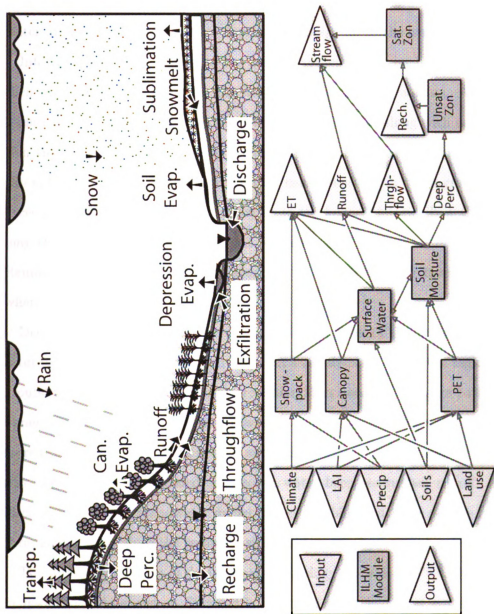


Figure 3.4: Conceptual model and simplified box diagram of the ILHM.

and surface pathways. Incoming rainfall is first subjected to canopy interception, while snow is routed directly to the snow pack model. Next, canopy throughfall and snowmelt are applied to the soil surface. These new inputs are then combined with any water stored in surface depressions and allowed to infiltrate into the soil. Any excess water at this point enters surface depression storage up to the available capacity.

Infiltrated moisture is added to the existing surficial soil layer budget, where it can then percolate downward under the influence of hydraulic gradients. Any moisture within the first soil layer is then available for evaporation, along with any transpiration that may occur in any of the biologically- active soil layers. Subsurface lateral throughflow is then calculated, which may cause moisture in down-gradient cells to exceed saturation. At this point, moisture in the lowest biologically-active soil layer may then percolate into the sediments beneath, where it becomes deep percolation. Remaining moisture in excess of saturation is exfiltrated back toward the surface where it also enters depression storage.

Deep percolation is then delayed as a linear function of the thickness of the unsaturated zone, which is estimated based on a steady-state run of the regional MRW groundwater model. The delayed percolation then becomes recharge to the three-dimensional transient groundwater flow model when it crosses the water table. Water stored in surface depressions is then subjected to direct evaporation. If depression storage capacity is exceeded, the excess water becomes surface runoff. Baseflow discharge from the groundwater model is then combined with the surface runoff and throughflow to produce the complete simulated stream hydrograph.

The surface component of the landscape hydrology model is a 177×153 grid at 106.3 m resolution. As shown in Figure 3.1, while the watersheds and groundwatersheds overlap for most of the modeling domain, some locations contribute only surface water or groundwater to Cedar Creek. To account for this, the landscape hydrology model is run for the entire domain while the unsaturated zone model only allows groundwater

Table 3.1: List of adjustable parameters in the UEB Snowmelt Model and their assumed values

<i>Parameter</i>	<i>Units</i>	<i>Value</i>
Snow Density	kg m ⁻³	200
Liquid Holding Capacity	-	0.15
Thermally Active Soil Depth	m	1.0
Snow Thermal Conductance	m hr ⁻¹	0.2

recharge in active cells of the saturated groundwater model, and the stream routing module only includes areas within the surface-watershed of Cedar Creek.

3.3.3.1 Precipitation and Snowmelt

Watershed available precipitation (P) is the sum of liquid rainfall and snowmelt. From late December through mid March, precipitation falls predominantly as snow in the Cedar Creek watershed. To model the storage and release of snow we used the UEB Snowmelt Model by *Tarboton and Luce* (1996), which is an explicit energy and water balance model designed to track three state variables: snow water equivalent, energy deficit (i.e., how much energy would be required to return the snow pack and soil layer to the 0 degree C reference condition), and the snow surface age. The model is computationally efficient because it assumes no temperature gradient within the snow pack and the layer of soil with which it interacts. For ease of integration the rest of the ILHM model suite, we ported the FORTRAN version of this code into MATLAB. The full UEB model requires air temperature, wind speed, relative humidity, and solar insolation. In addition it has several adjustable parameters, including the density of the snowpack, the thermal conductance of the snow, liquid water holding capacity of the snowpack, and the depth of soil with which the snow interacts. Preliminary calibration to our snow depth data provided the parameter values given in Table 3.1.

The ILHM suite only runs the UEB model if either the air temperature during a precipitation event is below freezing or the snow water equivalent of the snowpack is greater than 0.01 mm. Any water remaining in the snowpack at this point is then

applied to the surface as additional snowmelt. The watershed available precipitation (P) is defined, then, as the sum of all available snowmelt and liquid precipitation as determined by the UEB model.

3.3.3.2 Evaporation and Transpiration

To calculate the E and T terms in the water balance equation, potential evaporation and transpiration were first calculated. All evaporation and transpiration potentials, (canopy, PE_c ; depression, PE_d ; soil, PE_s ; and transpiration, PT) are calculated using the modified Penman-Monteith equation (*Monteith*, 1965) presented by *Chen et al.* (2005a) and shown as Equation A.1 in Appendix A. Evapotranspiration and transpiration rates vary according to land cover types through a stomatal conductance map which is assumed to be temporally invariant as discussed in *Chen et al.* (2005a) and with 8-day LAI scenes.

Incoming rainfall is first subjected to interception up to the water holding capacity of the canopy, which is related to LAI in the cell. Evaporation (E) terms are calculated for the canopy, soil, and any surface depressions within each cell. The total transpiration in each cell T is calculated to be the sum of the root water uptake from each biologically active soil layer.

The evaporation of moisture in depression storage is calculated after any infiltration and exfiltration (described below) in a given time step have occurred. Total depression storage capacity is determined by land use, soil texture, and slope class; a tabular reference of storage capacities can be found in *Manfreda et al.* (2005). Depression evaporation occurs at the potential rate until depression storage is depleted. Evaporation from depressions and directly from soil is allowed only from the proportion of soil that is exposed to solar radiation, thus assuming no soil evaporation from the portion shaded by canopy or covered with snow or ice.

Direct soil evaporation is allowed only from the first soil layer, which is a reasonable

assumption in this relatively humid region. Soil evaporation occurs at the lesser of calculated potential rate or the soil exfiltration depth (discussed further in Appendix A following Equation A16). We are exploring alternative strategies for calculating evaporation using the model of *Ritchie* (1972). The total transpiration in each cell T is calculated to be the sum of the root water uptake from each biologically active soil layer. We assume that transpiration only occurs above a dormant threshold temperature, which was chosen to be 40 degrees F for this study. Stomatal conductance is also assumed to be constant for this case, although we plan to incorporate variations due to changes in temperature, carbon dioxide concentrations, and soil moisture as described in *Chen et al.* (2005a).

3.3.3.3 Infiltration, Percolation, Throughflow, and Exfiltration

For this study we assume the biologically active soil can be described by two layers, with a total thickness calculated according to Equation A14 as the depth above which 90% of the root mass lies. The first soil layer, from which evaporation occurs and that controls infiltration capacity, is on the order of several centimeters thick. Infiltration capacity is calculated as the greater of either the soil-texture dependent saturated infiltration capacity, i_{sat} , or the first layer moisture deficit from saturation. We chose the maximum infiltration rate to be $(2 \times i_{sat})$ which determines the choice of the first soil layer thickness. This formulation produces similar results to empirical infiltration rate descriptions, and has the advantage that it does not require storm event tracking or single storm event modeling.

Here we assume that i_{sat} can take either its nominal soil-texture dependent value taken from literature values (see Appendix A), or $i_{sat} = 0$ if the soil is frozen, which we only allow to occur in agricultural soils based on *Schaetzl and Tomczak* (2001). The soil is assumed to be frozen if its temperature is below -0.25°C , measured by the MAWN station in Fremont at a depth of 10 cm.

In addition, we assume that no infiltration occurs in cells classified as permanent water features. Although the water features in this watershed often cover only a small portion of each land-use cell, our assumption may nevertheless be fairly realistic. The true physical system process is more accurately described as rapid percolation to a shallow water table followed by equally rapid rise and subsequent relaxation of the water table. The effect is a temporary increase in groundwater discharge that generally does not modify what is more traditionally called stream baseflow. Our stream gage data indicate that the characteristic time of this response may be on the order of twice the surface runoff response, or perhaps 5–7 days. This does not allow for significant losses due to evapotranspiration, thus the combined increase in stream discharge due to percolation to the near-stream water table and direct overland flow would be nearly equal to that expected from an assumption of zero infiltration. Streamflows in our model would thus be expected to peak higher and return to baseflow levels more rapidly than observed.

Throughflow, defined here as lateral subsurface flow within the biologically active soil zone is calculated using a simplified Richards equation model. For a full development of the Richards equation, see (*Hillel*, 1980). For purposes of computational efficiency, and in order to assure that the subsurface redistribution of moisture occurs in only one dimension, we assume that flow only occurs parallel to the dip of the slope and thus cannot flow uphill on the ~100 meter scale of our model cells. For environments where these assumptions may be invalid, alternate two-dimensional formulations could be substituted for this ILHM module given adequate computational resources. The van Genuchten model was used to calculate all soil-moisture dependent properties (*van Genuchten*, 1980) with parameter values given in Appendix A. The downgradient cell is determined via the D8 flow direction function (*ESRI*, 2003). However, each cell can have more than one upgradient cell, thus throughflow is the summation of the shallow subsurface flow out of all upgradient cells.

3.3.3.4 Unsaturated Zone Delay

Deep percolation beneath the biologically-active soil layer is delayed prior to becoming recharged as a linear function of the depth to the water table. The slope of this delay in units of days/meter, was determined from wells installed in the nearby Grand Traverse Bay Watershed, and was fixed at 2.5 for this study. It is important to note that this would not be the same as a solute transport time through the unsaturated zone. Delayed deep percolation then becomes recharge once it reaches the water table. The depth of the water table is assumed temporally invariant for the current version of the unsaturated zone delay module for ease of implementation. By fixing the depth of the water table for this purpose, we do not account for seasonal or trending differences in water travel times through the unsaturated zone. The seasonal differences in the depth of the water table are small (typically <1 meter) relative to average depths to water over most of the model domain. Locations very close to surface water features are an exception, but these areas comprise a small fraction of the total watershed area and thus will not significantly affect the dynamics of modeled recharge.

3.3.3.5 Groundwater Model

The groundwater model for the expanded MRW was developed using a suite of MATLAB utilities that we developed to create input files from GIS layers for MODFLOW. A regular grid of 1798×1865 cells (106.3 meters on a side) was used so that each cell directly overlies 16 Digital Elevation Model (DEM) cells. The vertical domain of the saturated zone model was then subdivided into two layers with approximately equal saturated thicknesses based on the simulated water table in a single layer model. Automated parameter estimation routines were applied to an early version of the Cedar Creek groundwater and soil balance model to estimate hydraulic conductivity values for aquifer sediments in geologic zones parameterized using a digital map created from *Farrand and Bell* (1982).

The Cedar Creek model is a single layer with dimensions of 184×162 and cell size of 100 meters. In this case, a single vertical layer provided an adequate description of flows through the Cedar Creek watershed, because it does not have significant vertical relief, extensive low-permeability subsurface layers, high-capacity pumping wells, or other features that tend to induce significant vertical head gradients. Groundwater discharge to streams is calculated with the Stream Flow Routing (SFR) package in MODFLOW that routes water via the kinematic wave equation (*Prudic et al.*, 2004).

3.3.3.6 Runoff and Stream Routing

Surface runoff is routed to the streams using an approach modified from that presented by *Manfreda et al.* (2005). In this version of the code, we assume that runoff cannot infiltrate once it is generated. It is routed overland and through streams according to the D8 flowdirection algorithm in ARC (*ESRI*, 2003) with runoff times given by the velocities in each cell along the flowpath. Runoff is assumed to travel overland at a velocity given by the Kerby time of concentration equation (*Kerby*, 1959). Once the runoff enters the stream channel its velocity is calculated using Manning's Equation.

For this study, the hydraulic radius, r , for Manning's Equation was determined as a function of discharge using low-flow channel geometry measurements in and around the Cedar Creek watershed along with geometries reported by the USGS for their stream gages. Wetted perimeter was assumed equal to $2 \times \text{depth} + \text{width}$, while area was simply $\text{depth} \times \text{width}$. A power law fit to these data produced the empirical relationship for this watershed:

$$r = 0.9046 \cdot Q^{0.283} \quad (3.2)$$

with a correlation coefficient R^2 of 0.77 (see Figure 3.5). Q in the above equation is the measured stream discharge in m^3/s .

While streamflow velocity can be dynamically calculated, for simplicity we have assumed a temporally constant v_{stream} for each stream cell. To calculate v_{stream} ,

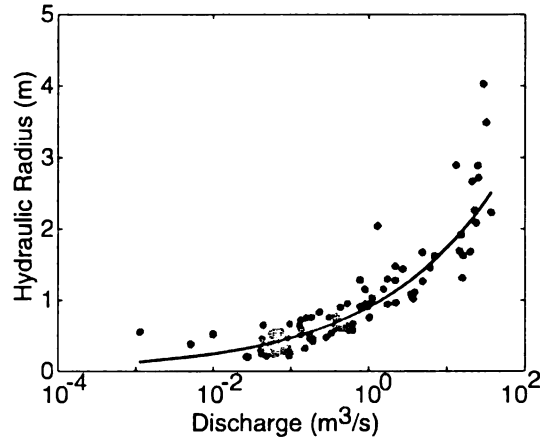


Figure 3.5: Plot of measured discharge versus hydraulic radius in streams in the greater Muskegon River Watershed and adjacent watersheds. The hydraulic radius calculation assumes rectangular cross-section geometry. The power-law fit to the data produced a correlation coefficient, R^2 , of 0.77.

we assume that each cell in the Cedar Creek watershed contributes a unit of runoff which is then routed to the gages using ARC's *flowaccumulation* function. To rescale the output to match a typical discharge event in a stream cell (Q_{ij}), we multiply the measured Q at the outlet by the ratio of the *flowaccumulation* value in each cell (i, j) to that of the outlet:

$$Q_{ij} = Q_{outlet} \left(\frac{flowaccumulation_{ij}}{flowaccumulation_{outlet}} \right) \quad (3.3)$$

from each cell to the outlet. Here the weighting is the inverse of the velocity in seconds/meter. ARC then multiplies this “cost” by the distance traveled through each cell along the entire flowpath and outputs to the travel time to the outlet. The travel time grid is then used to transform the precipitation excess grid into a runoff hydrograph.

3.4 Results

The results of the ILHM simulation for the Cedar Creek watershed are discussed in the context of the broader goals of the code, which are the prediction of temporal and

12
13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

Table 3.2: List of calibrated parameters for the unsaturated and saturated ground-water models.

<i>Parameter</i>	<i>Units</i>	<i>Value</i>
Outwash conductivity	m d ⁻¹	11.0
Till conductivity	m d ⁻¹	4.4
Unsaturated zone delay	d m ⁻¹	2.5

spatial variations in recharge with very little direct calibration of model parameters using readily available remote-sensed and ground-based data sources. All parameters for this prediction were based on literature values (Tables 3.2 and A1–A3), except for the UEB model parameters (Table 3.1) as well as two hydraulic conductivity values and one unsaturated zone delay parameter that were calibrated using an early version of the model (Table 3.2).

A plot of simulated versus observed heads (Figure 3.6) across this region shows a reasonable degree of agreement given the measurement uncertainty. Figure 3.6 shows no trending bias between simulated and observed heads, though a slight high-side bias is present at observed heads lower than approximately 200 meters. We would expect a higher degree of correlation between simulated and observed water levels in regions where pressure transducer data are available from wells, or if a parameter optimization were to be performed for this ILHM simulation.

Detailed evaluation of modeled flows is hampered by the lack of long duration stream gages with stable channels in the basin. All observed stream discharges were collected via established methods, however the rating curves for the two stream gages are currently inadequate to account for temporal adjustments of the channel geometry after flood events. As is commonly the case, we have few high flow measurements, which limits the accuracy of our flows calculated from the rating curve during large floods. In addition, flows in the Lower Cedar gage from January through March appear to suffer from ice-induced over-pressurization not observed at the Upper Cedar gage.

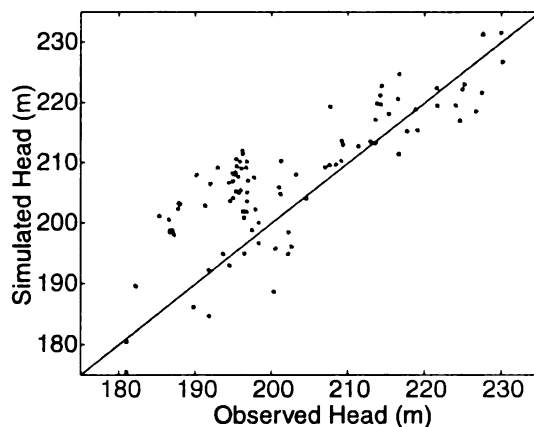


Figure 3.6: Uncalibrated simulated vs. observed groundwater heads plotted on top of a 1:1 line. Observations are from 96 residential wells installed in the Cedar Creek Watershed between 1990 and 2004 (see locations in Figure 3.1).

Despite almost no calibration of parameters in the near-surface components of the ILHM code, the model provided a reasonable prediction of observed flows (Figure 3.7) for the two available gage sites in this 100 square kilometer watershed (Figure 3.1) during the fall and winter months. The ILHM also provided reasonable predictions of baseflow for this watershed system during the entire year. Because this prediction is based almost entirely on a set of widely available meteorological inputs and GIS datasets combined with literature parameter values, the code appears suitable for directly simulating streamflows in ungaged basins.

The close agreement between observed and simulated baseflow levels also suggests that the model is providing reasonable predictions of recharge, which provides baseflow in these streams during low flow periods. During May and October 2003, and January–March 2004, the ILHM-simulated total flows typically agreed with gaged values within 10% (when the lower Cedar gage was not affected by ice cover). Baseflow levels in the smaller upper Cedar catchment proved highly sensitive to hydraulic conductivity in the outwash sand/gravel zone (Figure 3.2A). The conductivity values presented in Table 3.2 should not be viewed as a fully calibrated parameter set, as optimization of the total stream-flow simulation is the subject of ongoing evaluation.

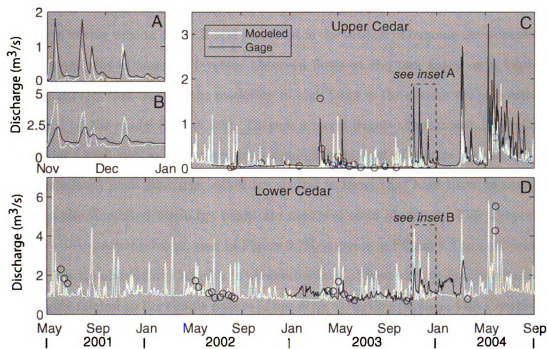


Figure 3.7: A & C) Upper and B & D) Lower Cedar Creek stream discharges with simulated values shown in white and gage values in black calculated based on stage discharge relationships. Manually measured discharge values are shown as black circles.

Dej

de M

des

Des

de

des

des

de

A

de

de

per

de

de

de

de

de

de

de

de

de

de

de

de

de

de

de

Despite the limitations imposed by some parts the stream gage data, the values from April through December of 2003 are reasonable for quantitative flow comparisons. During lower ET periods from April through early June and October through December, the ILHM-simulated total f luxes are approximately 20% higher than calculated from the flow gages installed at both the Upper and Lower Cedar gage sites. ILHM-simulated total fluxes during higher ET periods from June through October are much greater than observed, also likely due to an incomplete description of ET processes in this version of the ILHM code that is the subject of ongoing development.

A scatter plot of simulated versus observed flows at the two gages on a log-log scale illustrates that most of the moderate to high flows in the system are reasonably described by the model (Figure 3.8). There is a larger degree of mismatch in the Upper Cedar Creek site, as illustrated by a significant amount of scatter about the 1:1 line. Simulated peak discharge values are similar to those that have been measured, however the simulated discharge peaks are narrower than observed. This temporal offset at near-peak discharge, seen in Figure 3.7B, appears in Figure 3.8 as a tendency toward low simulated flows relative to observed values. These narrow simulated peaks are largely related to the simple nature of our stream routing package, and the unidirectional linkage between groundwater and surface water processes in this version of the ILHM code. The assumption of a temporally constant v_{stream} results in average flow velocities lower than those that would be expected, with the effect that simulated peaks would be even sharper if v_{stream} were a function of discharge. However, there are several known flow damping mechanisms, including f low through wetlands and bank exchange, which are not yet represented in this version of the code.

A map of the simulated average annual recharge across the watershed implies that agricultural areas may have higher recharge than forested areas according to this simulation (Figure 3.9). Several inter-related factors combine to account for the simulated differences. Agricultural areas experience less canopy interception than

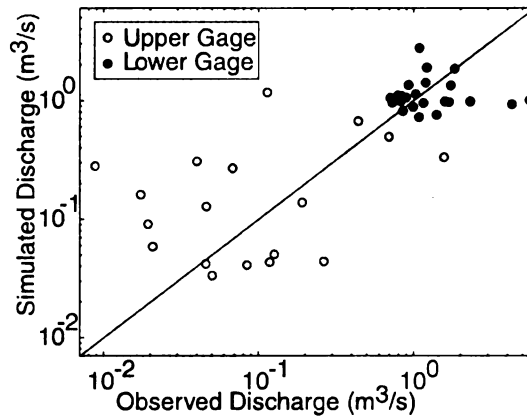


Figure 3.8: Uncalibrated simulated vs. observed flows for the Upper and Lower Cedar Creek gages plotted on top of a 1:1 line from approximately 20 manual discharge measurements taken at each of the upper and lower Cedar Creek gage locations.

forested areas due to lower LAI values. Although this increases the infiltration into agricultural soils, a resulting increase in transpiration tends to narrow the difference in recharge between these land use types. Our model may also under-represent soil evaporation in agricultural soils because it does not incorporate solar heating of the shallow soil layer. This simulated difference is the subject of future evaluation across the much larger MRW where more flow data are available.

The blocky nature of the simulated recharge in this map is mainly due to the large (1 km^2) LAI cells. The effect of these coarse cells is to decrease forest LAI and increase agricultural LAI in regions with mixed land-uses. We plan to resolve this issue through downscaling the LAI information by assuming that the measured LAI value is a linear combination of the LAI of forest and agricultural land-uses represented by the much higher resolution IFMAP dataset. Thus unique “agricultural LAI” and “forested LAI” values can be approximated at the resolution of the IFMAP data constrained by the total measured LAI from MODIS.

Areas with low hydraulic conductivity soils experience reduced recharge (Figure 3.9), such as portions of the upper watershed to the south side of the stream where loams and silty loams are common (compare with the soils map in Figure 3.2C).

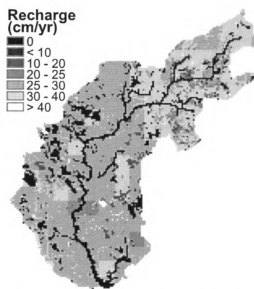


Figure 3.9: ILHM-simulated average annual recharge for the Cedar Creek watershed. Annual total precipitation averaged approximately 83 cm during the period of simulation.

In contrast, recharge can be greatly enhanced in internally drained regions. In this simulation we deactivated the runoff mechanism in internally drained areas, thus potentially increasing infiltration and shallow subsurface flow. This may be very important in areas with moderate to low-conductivity soils, but the internally drained areas in this watershed tended to also be sandy so the effect is only localized.

The Cedar Creek region experiences very little runoff from upland areas, which is consistent with the simulated map of precipitation excess (Figure 3.10). Nearly all the simulated precipitation excess in this watershed occurs in cells that are classified as “water” because there is no transpiration or percolation from those cells. The only cells with any significant precipitation excess that are not classified as water are in a region of lower conductivity sediments in the upper watershed. Despite a simple description of runoff processes, we do not expect significant runoff in most of the sediments across this watershed due to high infiltration capacities and saturated hydraulic conductivities. There is also very little simulated subsurface redistribution,

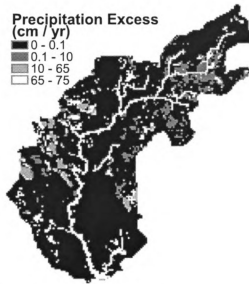


Figure 3.10: ILHM-simulated average annual precipitation excess runoff for the Cedar Creek watershed. Annual total precipitation averaged approximately 83 cm during the period of simulation.

or throughflow, throughout most of the watershed due to the highly conductive soils and relatively gentle topography. This process is most active in areas with steep slopes or water tables very near the surface.

The simulation provides evidence for a strong seasonality in recharge rates for the Cedar Creek watershed. The temporal variations in simulated deep percolation are shown in Figure 3.11. From September through March in the four illustrated years, the model predicts that approximately 70–80% of watershed available precipitation will percolate into the deep aquifer sediments where it eventually recharges groundwater. In contrast, the simulations show virtually no deep percolation over the growing season from May through September for the same years, which is consistent with the statistical findings of *Jayawickreme and Hyndman* (2007). This simulation indicates that agricultural areas have more recharge in the fall months than forested areas, while the opposite occurs with higher relative forest recharge in the spring months. This is reasonable as forests tend to have less extensive frozen soils during snowmelt

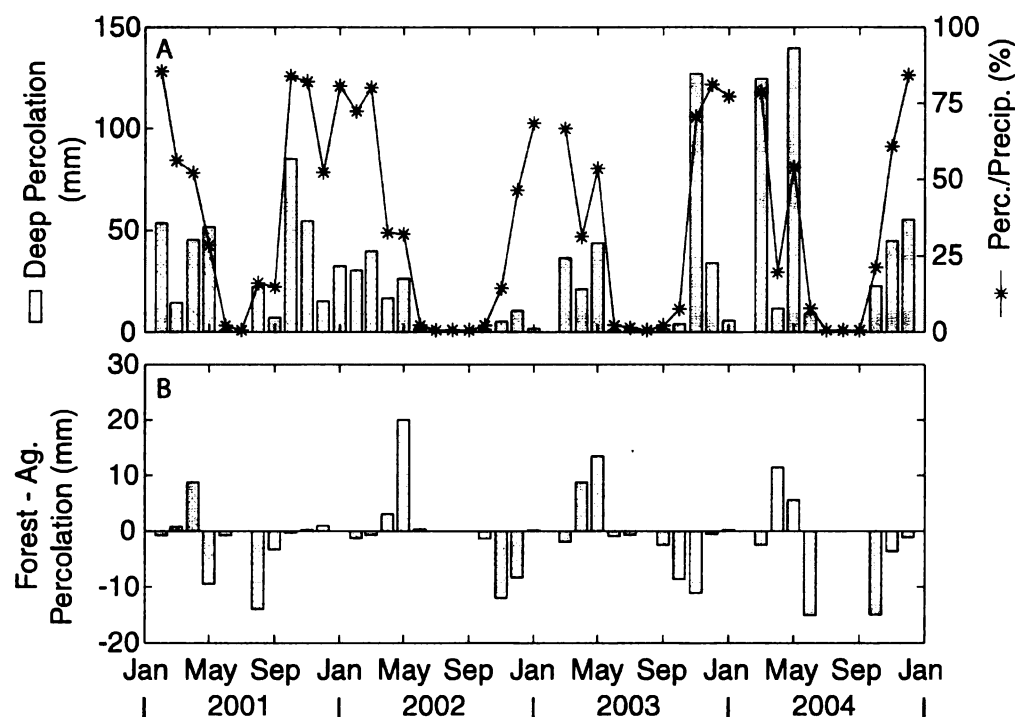


Figure 3.11: A) Monthly average deep percolation (bars) in the Cedar Creek watershed plotted with the ratio of percolation to total precipitation for each month (stars). B) Monthly difference in deep percolation between forested and agricultural land covers.

periods *Schaetzl and Tomczak* (2001), and agricultural LAI often begins to decline earlier in the year than in forested areas. Although coniferous forests may transpire year round, they represent only a small percentage of the forested areas in this study region.

Temporal variation in evaporation and transpiration are clearly the causes of most of the simulated variations in deep percolation because these are the primary loss processes. As Figure 3.12 illustrates, evaporation is generally a much smaller component of water loss in this watershed than transpiration, and this component is larger in forested land relative to agricultural land due to much higher forest canopy interception. Transpiration shows a stronger seasonal trend than evaporation, as it depends more strongly on LAI. Total agricultural transpiration is greater than that of forested

areas despite much greater potential transpiration in forested areas. Agricultural areas experience less canopy interception than forests, and thus greater infiltration and higher average soil moisture. As a result, agricultural areas tend to transpire closer to their potential rate than forested areas. Unexpectedly, agricultural transpiration also rises in the spring more quickly than that of forested areas according to these model results due to the similarity in LAI values during early spring and higher stomatal conductance values for agricultural areas relative to forests. As the LAI of forests increases in the late spring, the transpiration in these areas becomes larger than that of agricultural areas, until they reach approximate equality in late June that continues through the rest of the summer. Also during the summer, soil moisture levels reach their lowest point and often approach the permanent wilting point. As deep percolation cannot occur until the field capacity of the soils is reached, most of the water that does infiltrate the soil is transpired. Thus deep percolation is almost non-existent during summer months according to these simulations.

3.5 Discussion and Conclusions

We present the development and testing of a new suite of loosely coupled process-based codes that we call the Integrated Landscape Hydrology Model (ILHM). This modeling framework has several advantages over existing coupled hydrology codes. It can simulate much larger domains than fully coupled process-based codes, with fewer data requirements. In addition, the ILHM accounts for the processes and mass balance in a more rigorous manner than semi-distributed codes, which tend to lump or oversimplify important watershed processes and use parameters that cannot be independently measured. The ILHM also facilitates model development via direct input of readily available GIS data, in contrast to the impractical level of manual data input required for large domains from some existing process-based models such

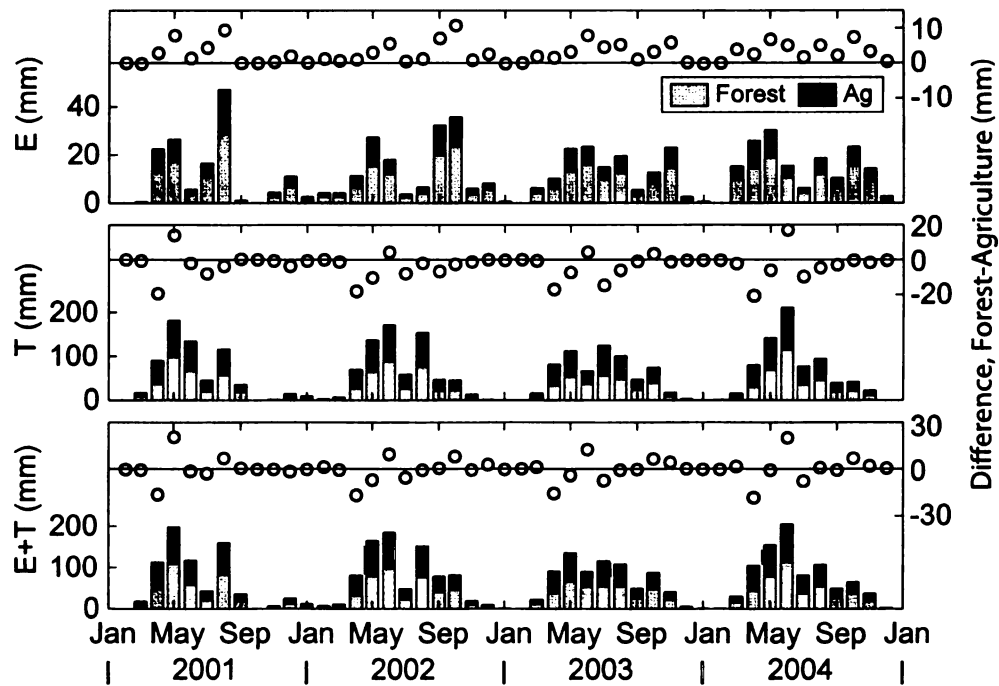


Figure 3.12: Average monthly evaporation (E), transpiration (T), and evapotranspiration (E+T) plotted for both forested and agricultural landscape as stacked bars along with the difference (Forest-Agriculture) between the two land-use types plotted as circles.

as Mike-SHE. Finally, ILHM is well-suited for forecasting purposes because it allows forcing data and component process models to be interchangeable; thus a model developed and calibrated with current data can be rapidly converted to a forecast simulation by adding the appropriate component process code.

This new modeling framework was designed to make development of models for large domains as simple as possible, while maintaining a rigorous fluid mass balance based on the primary processes that drive water movement over the landscape and through the subsurface. The approach is computationally efficient because it allows some processes to be simulated based on full numerical models while others can be described by simpler and thus faster water- and energy-balance approaches. Due to the loose-coupling framework, individual components can also be simulated at a variety of spatial and temporal scales appropriate to the individual processes. This framework also allows more rigorous simulation modules to be used in place of a simpler routine in cases where the additional computational burden provides necessary improvement in the model predictions. Alternatively, in cases where enough data exist to adequately describe a particular process, the data can be used in lieu of that process simulation module.

As currently configured, ILHM is designed to simulate flows through regions with connected surface water and groundwater regimes such as Cedar Creek. This test watershed has a sub-humid and temperate climate, with flow through a glacio-fluvial aquifer, largely covered with deciduous forests, agricultural land, and small percentage of urban cover. Thus the code is expected to provide reasonable predictions for similar environments in the sub-humid Midwest. The general processes are the same in arid and montane regions; however alternate modules would likely provide more accurate simulations in such cases. In particular, some high-relief environments may require a full two-dimensional representation of overland flow, including depth-dependent velocities for sheet or rill flow. Areas with large proportions of urban land uses will

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

require additional modifications, especially when engineered storm water systems have a significant effect on the hydrograph shape after a storm event.

In the Cedar Creek watershed, precipitation excess runoff routing and subsurface moisture redistribution are both largely inactive over most of the modeled domain and timeframe. As a result, the simulation results are similar even if these modules are not active for upland areas. Therefore, further testing in domains where these processes are responsible for a significantly larger percentage of the flow in a river system will be needed for these modules. The current unsaturated zone module is a very simple representation of hydrologic processes, thus we will explore the use of direct solution methods ranging from the Green-Ampt model through the full Richards equations. Additionally, using MODFLOW or any finite difference scheme has the disadvantage of requiring somewhat cumbersome rectangular grids that limit cell refinement at regional scales. However, ILHM can easily be altered to interface with a finite element code capable of representing and accounting for groundwater discharges to streams.

The ILHM was tested in the Cedar Creek watershed because of the need for high-resolution flow simulations that provide the interface between land use change models and ecohydrology models in the near future. This first evaluation of the ILHM modeling framework demonstrated that these codes can reasonably predict groundwater recharge and streamflow through a 130 square kilometer watershed with very little calibration using readily available data. The simulation represented overall basin recharge accurately, but it appears to have slightly overestimated recharge in agricultural areas. This is likely due to an inadequate representation of soil evaporation that will be addressed in a future version of the ILHM. The simulated hydrograph peaks are too narrow and decline more rapidly than is observed because the current surface water/groundwater linkage cannot represent bank storage and release processes, nor can the unidirectional coupling between surface water and groundwater fully repre-

sent near-stream processes at our chosen spatial scales. Nevertheless, the recharge and streamflow predictions provide reasonable descriptions of system behavior and will be further refined in future versions of the ILHM applied to much larger domains.

Acknowledgements

We would like to thank Mike Wiley and Paul Richards for supplying us with surface water flow measurements and stream levels from transducers, and Dushmantha Jayawickreme and Cheryl Kendall for their help with data processing and analysis. We are also grateful for the financial support of National Science Foundation grant (EAR-0233648), with supplemental support from both the Great Lakes Fisheries Trust's Muskegon River Initiative and Environmental Protection Agency grants (R830884). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the Environmental Protection Agency.

CHAPTER 4

ILHM Development, Validation, and Simulation of Regional-Scale Stream Discharge

4.1 Introduction

Climate and land use change are altering the terrestrial hydrologic cycle at all scales from hectare to continent . Over the last century, runoff from local to continental scales has increased in many areas of the world(*Probst and Tardy*, 1987; *Labat et al.*, 2004), but the reasons for this are poorly understood (*Gedney et al.*, 2006). Changes at the regional basin scale emerge from smaller-scale fluctuations within both meteorological inputs and the catchments of higher-order tributaries that respond heterogeneously to underlying drivers. Those catchments themselves are collections of hectare-scale plots, each of which have a unique history of land use and management. Even at this scale, more localized alterations such as a re-planted hillslope or riparian buffer strip on the order of several meters in width may significantly impact the quantity and quality of overland flow reaching a stream. Ultimately, predictive understanding of the impacts of climate and land use change on the terrestrial hydrologic cycle requires models capable of explicitly simulating this entire range of scales.

Historically, most hydrologic models have been developed for application across a

limited range of scales. Codes suited for regional to continental scales typically omit or simplify groundwater and include lumped parameters to explain much of the smaller-scale variability within model discretization units. The predictive capability of these models is limited outside of their calibrated set of conditions, as climate and land use change modify shallow groundwater systems in subtle and complex ways, and alter the values of lumped model parameters. Fully-coupled 3D numerical models are capable of predicting small catchment-scale fluxes but are generally either too computationally intensive, or too logistically cumbersome, to apply to large domains (*Markstrom et al.*, 2008). In addition to the models discussed below, a number of models have appeared once or at most a few times in the literature (*Manfreda et al.*, 2005; *Querner*, 1997; *Wigmosta et al.*, 1994; *Chiew et al.*, 1992).

Originating out of a need to describe mass, energy, and momentum fluxes at the land surface within general circulation models, land surface models have since evolved in complexity to describe much of the terrestrial hydrologic cycle. These models include, notably, BATS (*Dickinson et al.*, 1993), LSM (*Bonan*, 1996), CLM (*Yongjiu et al.*, 2001), LEAF (*Walko et al.*, 2000), and IBIS (*Foley et al.*, 1996). All share a common design goal: to accurately describe land surface fluxes within climate model cells on the order of $1^\circ \times 1^\circ$, far too coarse for most hydrologic applications (*Yu et al.*, 2006). Furthermore, the models either ignored or vastly simplified most hydrologic processes including runoff generation, routing, and groundwater flow. Some incorporate an approach similar to TOPMODEL (*Beven and Kirkby*, 1979), though at scales that preclude direct hydrograph simulation in all but the largest basins. To correct some of these shortcomings, VIC (*Liang et al.*, 1994) was written primarily as a surface hydrologic model, though it, also lacks a comprehensive treatment of groundwater. A more recent efforts at coupling a detailed hydrologic models within a global climate model is detailed in *Yu et al.* (2006).

Watershed hydrologists have developed a separate class of models that describe the

mass and energy fluxes within catchments at a variety of scales. These watershed models, including the Stanford model (*Crawford and Linsley*, 1966), SWAT (*Neitsch et al.*, 2002), and HEC-HMS (*USACE-HEC*, 1998), have evolved to suit the needs of soil conservationists and agricultural scientists studying water quality and quantity within catchments. These are semi-distributed, or lumped parameter models, whose parameter values are distributed within sub-aggregation units, sometimes referred to as hydrologic response units, or HRUs. Rainfall-runoff models, including HSPF (*Johanson et al.*, 1984) and PRMS (*Leavesley et al.*), are similar to watershed models though are typically more finely discretized at the surface. These have arisen from the need to forecast floods within gaged basins. Both watershed and rainfall-runoff models all share a similar treatment of subsurface flows, which are routed through typically two (or more) subsurface pathways that are equated to near-surface throughflow and deeper saturated groundwater flow. These pathways are simple linear reservoirs, and must be calibrated for most applications.

Agronomists and soil scientists required the capability to simulate the plot-scale, and evaluate the impacts local impacts such as tile drainage or tillage practices. They have developed models that simulate the full mass and energy balance at the surface. Examples of plot-scale models include HYDRUS-1D (*Simunek et al.*, 2005), SHAW (*Flerchinger*, 2000), and RZWQM (*Ahuja et al.*, 1999), which are fully-distributed, process-based (typically) 1D numerical models. These also include capabilities to assess nutrient fluxes through various storage reservoirs in the subsurface. Fundamentally, however, these are limited for application to broader domains because of their construction as 1D models, and their lack of groundwater components.

There are a number of models that include fully coupled surface and subsurface flow equations, with landscape process models of varying degrees of complexity. These include InHM (*VanderKwaak and Loague*, 2001), PIHM (*Qu and Duffy*, 2007), HydroGeoSphere (*Therrien et al.*, 2004), MODHMS (*HydroGeoLogic*, 2002), GSSHA

(*Downer and Ogden, 2003*), and CLM-PARFLOW (*Maxwell and Miller, 2005*). Applications of these models to small to mid-sized ($1 - 100 \text{ km}^2$) catchments demonstrate the value and efficiency of simultaneously solving fully-coupled systems of equations. However, upscaling these models can be prohibitive, or require coarsening of the discretization such that the value of the coupled system of equations approach is unclear.

A more recent approach has been to couple 1D land surface models with full 3D subsurface models. This method scales efficiently to regional domains while still representing the dominant hydrologic processes at those scales. The recently-released GSFLOW is a coupling of the rainfall-runoff model PRMS with MODFLOW (*Markstrom et al., 2008*), a 3D saturated groundwater model. MODFLOW has also been the subject of other integrated modeling efforts including HYDRUS-1D (*Seo et al., 2007*), the VSF package for MODFLOW (*Thoms et al., 2006*), and HSPF (*Said et al., 2005*). These models incorporate land surface components of varying complexity and coupling schemes, though all are fully-coupled. MIKE-SHE is another fully integrated model, though its cost is considerable (*Refsgaard and Storm, 1995*).

We have developed the Integrated Landscape Hydrology Model (ILHM) to simulate fine-resolution hydrologic fluxes across regional to continental model domains. ILHM explicitly incorporates process modules that describe nearly the entire terrestrial hydrologic cycle, including groundwater. Its parameters are both physical and measurable, though generally available from literature. These attributes make it well suited for highly-calibrated applications within a single catchment or collection of catchments, modeling ungaged basins, and hindcast or forecast simulations. Here we present an application to a regional (7400 km^2) catchment in Michigan, USA. This study demonstrates that ILHM, is capable of predicting hydrologic fluxes over a large domain at unprecedented resolution with minimal calibration.

4.2 The Integrated Landscape Hydrology Model

ILHM is composed of four loosely coupled “domain” modules: 1) surface water routing, 2) canopy and root zone moisture, 3) deep unsaturated zone, and 4) saturated groundwater (see Figure 4.1). The individual modules were selected based on the criteria that each includes: 1) fully-distributed physical representations of the most important hydrologic fluxes, 2) measurable parameters whose values are generally accessible through literature, 3) computationally efficient, highly scalable algorithms, from desktops to high performance computing (HPC) environments. Also, while not a formal criteria, existing open-source or published algorithms and solvers were chosen where available to reduce development time.

Evaluating possible modules to include within the context of the study area that would be simulated resulted in the following four domain modules: 1) a simple, de-coupled kinematic wave surface routing model, 2) a comprehensive 1D/2D water-balance canopy, root-zone, and wetland hydrology module, 3) a simple 1-D unsaturated zone pulse delay module, and 4) MODFLOW, a widely-used 3D saturated zone groundwater model (*Harbaugh et al.*, 2000). To facilitate rapid development while assuring platform independence, ILHM was written in the MATLAB computational environment (*Mathworks*, 2008). Also, a high-performance disk I/O back end, HDF5 (*HDF*, 2008), was chosen for its speed, scalability, and portability.

An alternate study area or set of questions to evaluate would have likely produced a different set of domain modules. Therefore, the code was written to incorporate alternate modules without modifying any existing code, and entirely new processes by modifying only a few lines. The choice of MATLAB greatly facilitated development due to its fully interactive development environment, however this imposes certain performance penalties. The structure of the code itself is highly modular (though not

explicitly object-oriented), allowing C/C++/FORTRAN/CUDA routines to replace computationally-intensive components. Furthermore, while the simulations presented in this study were all run in single-threaded desktop environments, the code may be readily extended and modified for HPC environments. The surface routing, canopy and root zone, and unsaturated zone modules are all “embarrassingly parallel”, and will scale to many nodes with little optimization.

Figure 4.1 illustrates the components, dimensionality, and spatial discretization schemes employed by the four domain modules in ILHM. The dashed arrows represent flux directions or dimensions that are partially supported. The root zone soil moisture module does allow for lateral cell to cell fluxes in the down slope direction. Additionally, vertical fluxes in the root zone are calculated if using the full 1D Richard’s Equation module, though that was not chosen for this study. Note that the root zone and saturated zone can consist of multiple layers, while the deep unsaturated zone currently a single layer (though tracks multiple wetting fronts). Within ILHM, each domain can be independently discretized, though currently only structured grids are supported. The root zone moisture module allows for a second level of discretization, allowing each cell to be subdivided into upland, wetland, and stream fractions.

There are three general types of model input requirements: 1) static inputs, 2) dynamic inputs, and 3) observations data. Static inputs describe the model domain boundaries, thicknesses, layering, model stresses, and parameters. Dynamic inputs include climate and landscape data such as precipitation, solar radiation, relative humidity (or dew point), wind speed, air temperature, and leaf area index (LAI). Other dynamic inputs might be parameters whose values change throughout the simulation period, land use type, or anthropogenic stresses such as pumping wells or reservoirs. ILHM also accepts observation data for any quantity that it simulates, which may be used to implement sensitivity analysis or parameter estimation schemes, though

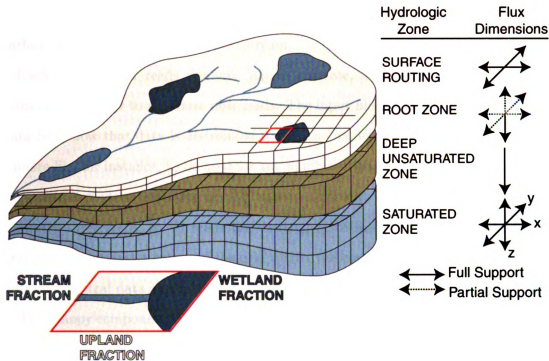


Figure 4.1: Graphic depiction of the spatial discretization scheme used in ILHM.

these are not included at this time.

4.2.1 Model Process Flow and Coupling

For this study, the canopy, root zone, and wetland hydrology modules run independently of the unsaturated, saturated, and surface routing modules, thus they are coupled serially rather than at run-time. Within each separate component, independent spatial and temporal discretization schemes may be employed. Currently, the surface processes run with hourly timesteps, while the unsaturated and saturated zone modules run on daily timesteps with weekly input stress periods. Each module writes outputs to disk, which are read by subsequent modules upon initialization.

Figure 4.2 illustrates the flow of processes within ILHM. In general the processes follow the path of water, interacting with the canopy, infiltrating or running off, being redistributed in the shallow root zone or transpired, percolating into the deep

unsaturated zone, recharging the saturated groundwater, and finally discharging to surface water and being routed downstream.

Each hour, ILHM reads dynamic inputs (climate, LAI, parameters) from disk, using a data buffer to minimize disk reads. The input file format separates the input data from how that data is distributed within the model. This allows a single input climate file, for instance, to be used for any number of parallel threads, or for smaller sub-domain models useful for testing and development. Thus the input controller performs three steps, it: 1) reads the data, 2) distributes those data to the model grid, or possibly just a subset of the grid within the current thread, and 3) stores the data in a central data controller that process modules may then request data from.

The canopy components are run next. First, the model updates albedo values across the landscape and determines the shortwave radiation intercepted by the canopy, and upland or wetland portions of each cell. Next, solid and liquid precipitation are intercepted and stored as canopy water. Finally, canopy evaporation and transpiration are calculated.

For the upland fraction of each cell, evaporation terms are then calculated for water stored in surface depressions and from the soil. Any canopy throughfall precipitation is first sent to the snowpack module. This module maintains both an explicit water and energy balance, accumulating snow in a single layer as well as calculating snow melt and sublimation. Snow melt or other liquid precipitation is then added to surface depression storage where it can either infiltrate into the soil or become overland flow runoff. Infiltration is then added to root zone soil moisture storage, transpiration is removed, and water is redistributed vertically and horizontally. Finally, if active, ex-filtrated water is passed to surface depressions where it may again runoff if depression storage capacity is exceeded.

Potential evaporation from shallow and deep wetlands within the wetland fraction of each cell is calculated next. Then, for deep wetlands, the temperature is calculated

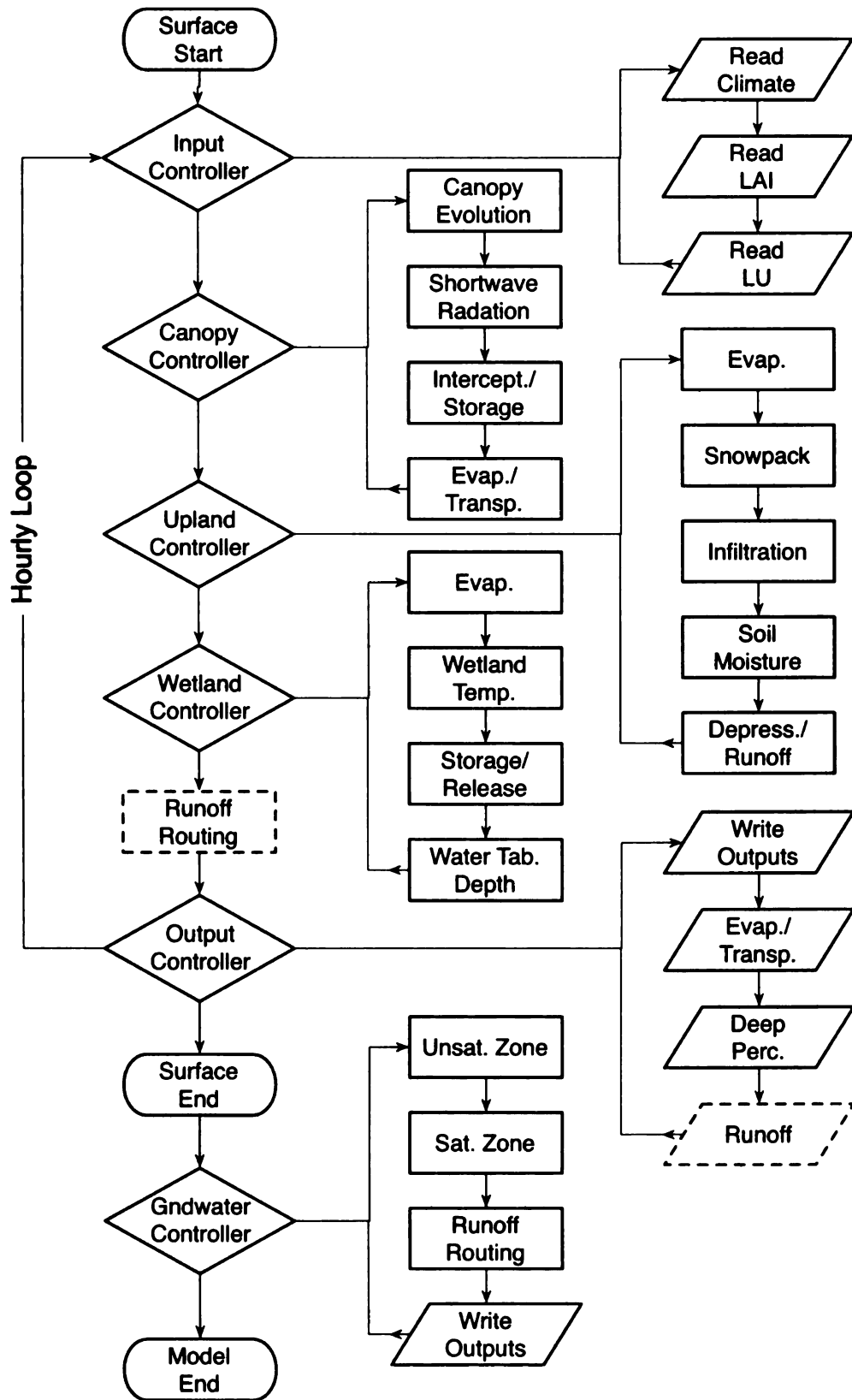


Figure 4.2: ILHM generalized process flow diagram, focusing on surface processes. Dashed outlines indicate optional items.

and an ice-pack accumulation and melt module is run. In addition to mass balance, the temperature and ice pack modules also maintain energy balance. The wetland water storage and release module calculates actual evaporation, transpiration, deep percolation, and runoff. Finally, the water table depth is updated.

Following the physical process modules, outputs are written to disk for coupling with other domain modules or for comparison to observations and visualization. The model controller then loops through the remaining hours within the simulation. When the surface process modules are complete, the unsaturated zone, saturated zone, and runoff routing modules are then run in sequence.

4.2.2 Simulating Hydrologic Fluxes

Hydrologic fluxes are calculated according to the equations in the Appendix A, however note that for consistency, the symbols have been updated for this Chapter. There are numerous improvements and modifications to the code, which are described in detail in the sections below and Appendix B.

4.2.2.1 Canopy

The canopy water balance is

$$\Delta S_{CAN} = W_{PCP} - W_{THR} - W_{EC} \quad (4.1)$$

where ΔS_{CAN} [m/s] is the change in water stored in the canopy S_{CAN} [m], W_{PCP} [m/s] is the total incoming precipitation, W_{THR} [m/s] is canopy throughfall ($W_{PCP} - W_{THR}$ represents water intercepted by the canopy), and W_{EC} [m/s] is canopy water evaporation. The canopy intercepts precipitation up to its storage capacity, which is a function of the Leaf Area Index (LAI) of each grid cell. Canopy evaporation and canopy transpiration, W_{TU} [m/s], are calculated via a modified Penman-Monteith equation. ILHM treats the canopy as continuous across the upland, stream, and wetland fractions within each cell.

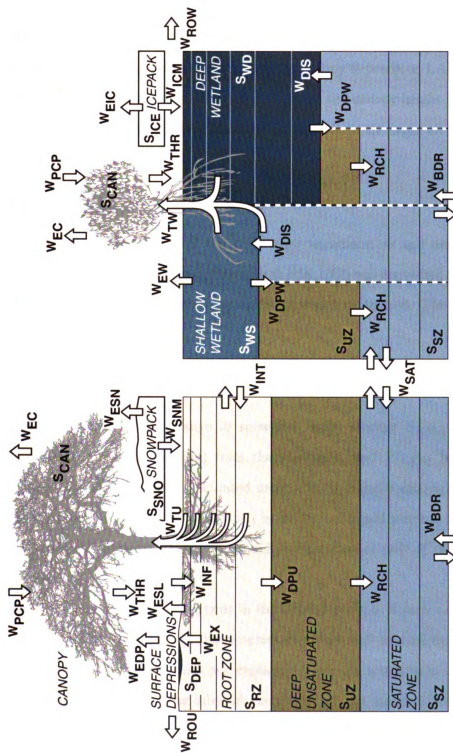


Figure 4.3: Detail of hydrologic fluxes calculated by ILHM for upland (left) and wetland (right) portions of cells.

Total available energy for evaporation is a function of intercepted solar radiation, governed by the LAI, Stem Area Index (SAI), incident angle of solar radiation, relative proportions of beam and diffuse solar radiation, and albedo of the ground beneath the canopy. Resistance to vapor transport within the canopy depends on LAI. Resistance to momentum transport depends on input wind speed and canopy height (which varies as a function of LAI). Atmospheric vapor deficit depends on the input air temperature and vapor pressure.

4.2.2.2 Snowpack

ILHM uses a modified version of the UEB snow accumulation and melt model by *Tarboton and Luce* (1996). For coupling with ILHM, UEB was converted to MATLAB code from FORTRAN, and made fully distributed and vectorized. The snow model maintains the water balance

$$\Delta S_{SNO} = W_{THR} - W_{ESN} - W_{SNM}. \quad (4.2)$$

Here, ΔS_{SNO} [m/s] is the change in snowpack water storage S_{SNO} [m], W_{ESN} [m/s] is evaporation (sublimation) from the snowpack, and W_{SNM} [m/s] is snow melt. Snowpack sublimation is calculated using a mass-transfer approach, assuming neutral atmospheric conditions. Snow melt is the sum of liquid water in excess of the snowpack storage capacity, either from liquid precipitation or melt of the solid phase of the snowpack.

Currently, the snow model only runs in the upland portion of each cell. The wetland and stream portions have an ice accumulation and melt routine that is roughly equivalent. This snow model does not explicitly track the temperature of the snow pack, but rather accumulates an internal energy deficit as a state variable.

4.2.2.3 Infiltration and Depression Storage

The upland portion of each cell has a surface depression storage capacity that must be overcome prior to overland flow runoff. The capacity of this storage is a property of the land cover, soil texture, and average slope of each cell. The depression storage water balance is

$$\Delta S_{DEP} = W_{THR} + W_{SNM} + W_{EX} - W_{INF} - W_{EDP} - W_{ROU}, \quad (4.3)$$

where ΔS_{DEP} [m/s] is the change in surface depression storage S_{DEP} [m], W_{INF} [m/s] is the rate of water infiltration into the soil, W_{EX} [m/s] is the rate of ex-filtration from the soil due to near-surface interflow over saturation, W_{EDP} [m/s] is depression evaporation, and W_{ROU} [m/s] is the overland flow runoff rate from uplands. Depression evaporation occurs at the potential rate, as calculated by the modified Penman-Monteith equation, using a resistance to vapor transport of 0, and a resistance to momentum transport determined by soil surface roughness.

Soil infiltration depends on the average saturated hydraulic conductivity of the first soil layer, the extent of impermeable surfaces within each cell, and whether the first soil layer is frozen. The maximum infiltration rate is set equal to the average saturated conductivity, which underestimates total infiltration by not incorporating matric and ponding depth potentials. For this study, the soils are relatively coarse with high conductivity, thus direct overland flow runoff rarely occurs over much of the watershed.

4.2.2.4 Root Zone

The overall root zone water balance in the upland portion of each cell is

$$\Delta S_{RZ} = W_{INF} - W_{EX} - W_{ESL} - W_{TU} - W_{INT} - W_{DPU}. \quad (4.4)$$

Here, ΔS_{RZ} [m/s] is the change in root zone water storage S_{RZ} [m], W_{ESL} [m/s] is direct evaporation from the soil, W_{INT} [m/s] is near-surface interflow, or hori-

zontal redistribution between adjacent cells, and W_{DPU} [m/s] is deep percolation to the unsaturated zone. Both interflow and over saturation due to variable vertical conductivity can generate W_{EX} . For this study, W_{INT} is a very minor term in the overall water balance, and was excluded for computational efficiency. The transpiration rate, W_{TU} , is controlled by the total shortwave radiation intercepted by the canopy, canopy stomatal conductance, and soil moisture. Transpiration is partitioned among soil layers according the fraction of total root mass within each layer.

Moisture is redistributed vertically between layers using either the 1-D Richard's equation, or free-drainage at the rate of unsaturated hydraulic conductivity. For this study, the free-drainage method was chosen because of its numerical stability and computational performance. Water cannot enter the root zone from below, due to the one-way coupling between ILHM's surface components and its deep unsaturated and saturated groundwater modules. Efforts are underway to couple these modules at run-time and incorporate capillary rise, saturated exfiltration, and other important fully-coupled processes.

4.2.2.5 Wetland Storage

Shallow wetlands, deep wetlands, and streams, maintain the water balance

$$\Delta S_{WD \text{ or } WS} = W_{THR} + W_{ICM} - W_{DPW} - W_{EW,S} - W_{TW,S} - W_{ROW}, \quad (4.5)$$

where ΔS_{WD} [m/s] and ΔS_{WS} [m/s] are the changes in storage S_{WD} [m] of deep wetlands, and S_{WS} [m] of shallow wetlands, W_{ICM} [m/s] is stored precipitation released by ice pack melt, $W_{EW,S}$ [m/s] is the rate of evaporation from surface storage in inundated wetlands, $W_{TW,S}$ [m/s] is the rate of transpiration from surface storage, and W_{ROW} [m/s] is the runoff or release rate.

Inundated area within a watershed can vary dramatically seasonally, particularly in regions with shallow water tables. With a run-time coupling between the saturated groundwater and surface components of ILHM, direct calculation of inundated area is

possible. However, with the one-way coupling currently implemented in ILHM, a less direct means of calculating inundated area is required. ILHM currently implements a user-specifiable seasonal water table fluctuation function, which combined with input average depths of wetland complexes allows inundated area to vary.

Runoff in wetlands is calculated differently depending on the inundation state. Inundated wetlands, i.e. those whose storage is positive or with a water table at or above the surface, release a portion of their storage each timestep. W_{ROW} for inundated wetlands is calculated as the product of a user-specified recession constant and the storage within each cell. For dry wetland cells, W_{ROW} is calculated as for upland runoff, if the rate of throughfall exceeds infiltration capacity in the wetland soils, runoff is generated.

Wetland storage will be negative when the water table is beneath the surface. In this case, the minimum storage equals the depth of the water table, multiplied by the average porosity of the soil in between the surface and water table. When storage reaches this minimum, deep percolation no longer occurs. Deep percolation, W_{DPW} from wetlands is calculated as the product of a specified wetland conductance and the storage head in each wetland cell. This head is the total head above the water table, including both sub-surface saturated water and ponded water at the surface.

Potential transpiration is calculated with the Penman-Monteith equation. If the wetland is perched, transpiration $W_{TW,S}$ is removed from wetland storage provided water is available. Transpiration in connected wetlands is assumed to primarily occur within the phreatic zone, and thus is passed to the saturated groundwater module as $W_{TW,P}$.

The rate of evaporation is calculated via the Penman-Monteith equation for shallow inundated wetlands, and via the energy balance equation for deep wetlands whose temperature has been explicitly calculated. If water storage is positive, evaporation is removed from storage as $W_{EW,S}$, otherwise this evaporative demand is passed to

2000

2001

2002

2003

2004

2005

2006

2007

2008

2009

2010

2011

2012

2013

2014

2015

2016

2017

2018

2019

2020

2021

2022

2023

2024

2025

2026

2027

2028

2029

2030

the runoff routing module where it will be removed (if available) from groundwater discharge to surface water or directly from the saturated groundwater module. If the wetland cell lies within an internally-drained zone (i.e. does not drain to the host watershed), it is passed to the groundwater module as $W_{EW,I}$ [m/s], otherwise, evaporation $W_{EW,E}$ [m/s] is passed to the runoff routing module. Evaporation in wetlands with negative storage is calculated as soil evaporation with the Penman-Monteith equation. The difference between inundated and dry wetland evaporation is the surface resistance term, which is 0 for inundated wetlands and dependent on depth-to-water in dry wetlands.

4.2.2.6 Wetland Temperature and Ice

The energy balance of a wetland can be expressed as

$$D_w \rho c_p \Delta T_W = Q_R + Q_A - Q_H - Q_E - Q_C, \quad (4.6)$$

where the terms on the left hand side are, in order, the depth of the wetland, density of the water, specific heat capacity of water, and change in temperature of the wetland. The energy flux terms are, in order, net radiative flux, advective flux, sensible heat flux, latent heat flux, and conductive heat flux.

For computational efficiency, terms that are less important in shallow wetlands, including temperature change (i.e. heat storage), advective flux, and conductive flux may be ignored in ILHM. For deep wetlands, heat storage within the wetland becomes important. For those cells, a full multi-layer energy balance module that implements the 1D eddy diffusion model by (*Hostetler and Bartlein, 1990*). This module incorporates wind-generated diffusive eddies, as well as full convective mixing.

A simple ice accumulation module has also been implemented that builds ice when the temperature of the 1st wetland layer drops below zero. From there, the energy balance of the ice is solved via Equation 4.6, ignoring conductive fluxes to the water below. Because of this simplification, modeled ice thicknesses are greater observations,

however timing of melt is reasonably simulated which is important for calculating evaporation flux. The ice pack stores incoming precipitation separately from ice accumulated from below. This water is then released to the system proportionately as melt occurs.

4.2.2.7 Deep Unsaturated Zone

The deep unsaturated zone module maintains the water balance

$$\Delta S_{UZ} = W_{DP} - W_{RCH}. \quad (4.7)$$

Here, ΔS_{UZ} [m/s] is the change in unsaturated zone storage S_{UZ} [m]. Total deep percolation leaving the root zone and wetlands, W_{DP} [m/s] is given by

$$W_{DP} = F_U W_{DPU} + (F_W + F_S) W_{DPW}, \quad (4.8)$$

where F_U , F_W , and F_S are the fraction of upland, wetland, and streams in each cell respectively. Individual wetting fronts representing the flux of water W_{DP} travel vertically downward through the deep unsaturated zone at a rate controlled by the hydraulic conductivity of the first saturated groundwater layer. The parameters for this were determined by fitting a 1-D Richard's Equation model to water table fluctuation data collected in a nearby watershed.

This simple routing greatly enhances the stability and speed of the model, but does not allow vertical fluxes through gradients in matric potential, nor wetting front attenuation and broadening. An adequate description with a 1-D Richard's Equation model was deemed too computationally extensive for a domain of this size, though it could be easily implemented for smaller model areas. Methods of intermediate complexity that may offer a more desirable trade off between speed and accuracy are under investigation, including transfer functions or wetting-front tracking as implemented in the UZF package in MODFLOW 2005 (Niswonger *et al.*, 2006).

After traveling through the unsaturated zone, water enters the saturated zone as groundwater recharge, W_{RCH} [m/s]. The total travel time depends on rate of wetting

front movement and the thickness of the deep unsaturated zone. For this study, that thickness is assumed constant in time, and was set by extracting the average depth to the water table from a preliminary model run. Full run-time coupling of the surface, unsaturated, and saturated zone codes already under development will allow the thickness of this zone to vary.

4.2.2.8 Saturated Zone

Saturated zone groundwater flow is calculated using MODFLOW-2000 (*Harbaugh et al.*, 2000), and maintains the water balance

$$\Delta S_{SZ} = W_{RCH} - W_{SAT} - W_{DIS} - W_{TW,P} - W_{EW,I} - W_{WEL} - W_{BDR}, \quad (4.9)$$

where ΔS_{SZ} [m/s] is the change in saturated zone storage S_{SZ} [m], W_{DIS} [m/s] is saturated discharge to water bodies, W_{WEL} [m/s] are anthropogenic withdrawals, and W_{BDRY} [m/s] is the sum of a variety boundary condition terms. ILHM prepares all input files for MODFLOW, and passes groundwater recharge and evapotranspirative fluxes to it after running the surface components. Groundwater discharge to drains, rivers, streams, or general head boundaries is then passed to a post-processing routine that routes these through the drainage network, and calculates remaining open-water evaporation.

A virtue of the one-way linkage to MODFLOW within ILHM is that any groundwater code may be used without modification of the source. Also, legacy models within those codes may be used directly. This allows users to leverage the thousands of existing groundwater models to rapidly build a more capable integrated hydrologic model within ILHM. As already discussed, this one-way linkage does limit the ability to fully describe certain processes within ILHM. The importance of the full coupling can vary greatly depending on the degree of interaction between surface- and ground-water. Thus even after completing the development of full-coupling, some applications (for instance, parameter estimation) may still benefit from simpler one-way coupling.

4.2.2.9 Runoff Routing

Total stream discharge generated in each cell is

$$W_{RUN} = A_S [F_U W_{ROU} + (F_W + F_S) W_{ROW}] + A_G (F_W + F_S) W_{DIS} - W_{EW,E}. \quad (4.10)$$

Here, W_{RUN} [m^3/s] is total runoff from each cell, W_{UP} [m^3/s] is the up gradient discharge, A_S [m^2] is the area of each cell in the surface components, and A_G [m^2] is the area of each cell in the groundwater module. Note that currently, ILHM does not allow loss along the flowpath. In relatively humid regions with shallow water tables, soil saturation increases downslope, and groundwater heads generally exceed heads in the stream. Thus streamflow routing can be decoupled from the surface and groundwater processes, and handled as a post-processing step.

The total runoff generated in each cell is then routed along a flowpath generated by a D8 flow algorithm (*O'Callaghan and Mark, 1984*). The velocity along the flowpath depends on the kinematic wave velocity, Manning's roughness of the cell and the average slope of the cell. Runoff is aggregated only at a user-specified set of points. Flow times between points are also specified. This simple routing approach is very fast, but tends to under-represent stream water velocities at high flow levels, thus flood peaks tend to be too narrow. In environments that produce more overland flow, enhanced flow routing algorithms could be implemented with little performance penalty, including $D\infty$ flow routing (*Tarboton, 1997*), and stage-dependent channel velocity.

4.3 Model Domain and Input Data Preparation

4.3.1 Study Area, Simulation Period, Model Domain, and Validation Data

The Muskegon River Watershed (MRW) is located in central Lower Michigan (Figure 4.4), and covers approximately 7,400 km². The Muskegon River ultimately drains into Lake Michigan, the second largest of the Laurentian Great Lakes. Its primary drainage, the Muskegon River, begins as an outflow from Higgins Lake, and flows through another large inland water body, Houghton Lake. It discharges to Lake Muskegon, which is connected to Lake Michigan via a narrow canal. There are several other sizable lakes along the Muskegon River, as well as a significant number of wetlands, particularly in the upper portion of the watershed. According to other studies, the majority of the total flow in the watershed comes from groundwater discharge to streams (*Jayawickreme and Hyndman, 2007*). Several large control structures modify the watershed, including Hardy, Rogers, and Croton dams in the central portion of the watershed.

The model boundary domain for this study has been extended well past the MRW to a total area of approximately 18,900 km². The boundary was chosen to coincide with major hydrologic divides in neighboring basins. Integrated surface- and ground-water models must incorporate boundaries appropriate to both domains. In this region, particularly in low-relief areas, the watershed divide does a poor job of describing the “groundwatershed” of the Muskegon River. The model bottom is given by the contact between the surficial unconsolidated aquifer and the bedrock beneath, which is assumed to be impermeable. A significant portion of the bedrock contact is comprised of permeable aquifer materials, however for this study we have assumed that the fluxes between the surficial and bedrock aquifers are small relative to the surface to shallow-

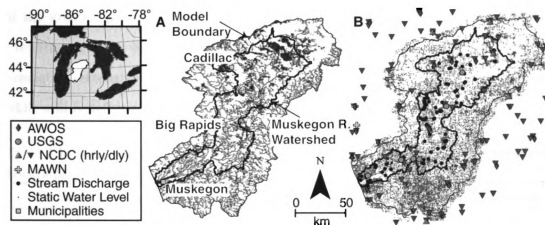


Figure 4.4: Map of the Muskegon River Watershed and surrounding model boundary. In part A, major municipalities along the drainage network are highlighted. Part B shows the locations of input gages and observation data locations.

groundwater fluxes of interest here. This assumption is supported by two lines of evidence: 1) the freshwater-saline contact in bedrock aquifer materials beneath much of the MRW is relatively shallow, suggesting a small downward flux from the glacial sediments to the bedrock, and 2) the waters within the three primary bedrock aquifers are saline or brine (*Westjohn and Weaver, 1998*).

A total of 28 years of hydrologic fluxes are simulated, from January 1st, 1980, to December 31st, 2007. During this period, new input data sources became available, greatly increasing the quality of model input data (see Table 4.1). A new gage network came on line in 1996, the Michigan Automated Weather Network (MAWN) (*MAWN, 2008*), providing the first reliable measurements of total solar radiation in the region. Also beginning in 1996, NEXRAD precipitation data at 4 km, hourly resolution became available (*Andresen, 2008*). In 2000, Leaf Area Index (LAI) data became available from the Moderate Resolution Imaging Spectrometer (MODIS) on NASA's TERRA and AQUA satellites (*Knyazikhin et al., 1999*). Prior to this, LAI had to be inferred from the Normalized Difference Vegetation Index (NDVI) products produced using data from instruments such as the Advanced Very High Resolution Radiometer (AVHRR). These heterogeneities of data sources and quality present some

of the greatest challenges to long-term change modeling. Below, (in Climate Data Preparation, Solar Radiation Preparation, and Precipitation Data Preparation) we present the extensive efforts taken to minimize systematic differences between datasets while still benefiting from newer, higher-quality data sources.

Table 4.1: Input climate and LAI data sources. Start and end dates are generalized. Locations of gage data are shown in Figure 4.4.

<i>Source</i>	<i>Type</i>	<i>Units</i>	<i>Start</i>	<i>End</i>	<i>Resolution</i>	<i>Frequency</i>
AWOS ¹	dew point 2 m air temp. 10 m wind speed sky conditions pressure	F F knots oktas mbar	1980	2006	15 stations	hourly
MAWN ²	relative humidity 1.5 m air temp. 3 m wind speed total solar rad. 2 in soil temp. 4 in soil temp.	% C m/s kJ/m ² C C	1996	2007	4 stations	hourly
NCDC ³	daily precip. hourly precip.	in·100 in·100	1980 1980	2007 2007	67 stations 10 stations	daily hourly
NEXRAD ⁴	hourly precip.	mm·10	1996	2007	4 km	hourly
GIMMS ⁵	NDVI	-	1981	2006	8 km	15 days
AVHRR ⁶			1989	2003	1 km	7 days
MODIS ⁷	LAI (v5)	-	2000	2007	0.93 km	8 days

¹(Andresen, 2007) ²(MAWN, 2008) ³(Center, 2001; NCDC Hourly, 2008)

⁴(Andresen, 2008; Fulton et al., 1998) ⁵(Tucker et al., 2005) ⁶(Eidenshink, 1992)

⁷(Knyazikhin et al., 1999)

In this study, two types of data will be used to validate the integrated model, groundwater heads and stream discharge. Static water levels were measured in nearly 50,000 wells during the 1980 - 2007 time period by drillers installing drinking water wells (see Figure 4.4). Though the quality of each individual measurement may be low, the dataset in aggregate presents an dataset for evaluating model-predicted groundwater heads. The USGS has maintained stream discharge gages at 6 locations, though no more than 5 were operating at any given time. These will be used to eval-

uate total simulated stream flow, along with errors in model fluxes. In addition to these, we have collected 105 stream discharge measurements during with a combination of wading and bridge-board cross-section velocity methods. These measurements were collected at or near baseflow conditions in the system and will be compared to simulated stream baseflow.

Table 4.2: Table of observations and validation data. Location abbreviations “MR” refers to “Muskegon River”. The “Area” column refers to the surface watershed area.

<i>Source/Type</i>	<i>Location</i>	<i>Area (km²)</i>	<i>Units</i>	<i>Start</i>	<i>End</i>
USGS ¹ daily gage	MR @ Newaygo	6,086	ft ³ /s	1980	1993
	MR @ Croton	5,991		1995	2007
	MR @ Evart	3,711		1980	2007
	Little MR @ Oak Grove	894		1995	2007
	Clam @ Vogel Center	629		1980	2007
	Bear @ Muskegon	43		1980	2007
This study synoptic flow	42 locations	4.2 - 928	m ³ /s	2003/07/23-25	
	28 locations	0.2 - 3,047		2004/08/21-24	
	35 locations	1.2 - 530		2004/10/09-10	
MDEQ ² static water level	~50,000 locations	-	ft	1980	2007

¹(USGS, 2008) ²(MDEQ, 2008)

4.3.2 Static Inputs

There are four primary static GIS inputs, shown in Figure 4.5. Many of the ILHM parameters depend on either soil texture, land cover type, or both. Here we use land cover from a 1998 update of the 1978 MIRIS aerial-photo classified land use map of the MRW region (Torbick *et al.*, 2006). The two dominant land use classes in the MRW and surrounding region are forest and agricultural lands. During the 20th century, the extent of agriculture declined sharply, while forest, urban, and open/grassland land covers moved into abandoned marginal agricultural lands.

Three-dimensional soil textures to a depth of 3 m are taken from the digitized soil survey SSURGO data (USDA-NRCS, 2008). In the survey data, each map unit has a unique layering specified, based primarily on the soil horizons. To simplify

extracting the texture data, the texture values were mapped to four layers from 0-20, 20-50, 50-100, and 100-300 cm. To select the appropriate texture for each layer in the SSURGO map units, the majority texture was chosen. Soil hydraulic parameters were then mapped to these textures. Note, the layering within ILHM will not necessarily match the four layers extracted from the survey data as the ILHM total root zone thickness is variable, and the global number of layers adjustable. Thus on import, the soil parameters are mapped to ILHM layers based on the weighted average of the input layers within in each ILHM layer.

The hydraulic conductivities of the deeper unsaturated and saturated zone sediments are assumed to be reasonably parameterized by the zones in the surficial, Quaternary geology map (*Farrand and Bell*, 1982). As illustrated in Figure 4.5B, the dominant soil textures are sands, sandy loams, and loamy sands. These relatively coarse textured materials have high infiltration capacities and hydraulic conductivities. As a result, most of the rainfall and snow melt within the watershed percolates into the soil, producing very little direct overland flow runoff. The important exception to this is in flooded areas and low-lying uplands in close proximity to streams and lakes. These coarse-textured soils are produced from similarly coarse Quaternary sediments (Figure 4.5C).

This map has three relatively distinct regions: 1) a low-relief area in the upper watershed comprised of alternating glacial outwash sandurs and finer-textured lacustrine sediments from pro-glacial lakes, 2) an inter-lobate moraine region in the center of the watershed made up of chaotic assemblages of mostly coarse glacial till and subsequent incised alluvium, and 3) a lake shore region in the lower watershed of relatively high-energy near-shore deposits and coarse alluvium. The groundwater model consists of three layers, with conductivity parameters directly related to this map in the first layer, and a more generalized three-region conductivity map for the deeper two layers. A preliminary steady-state groundwater model was used to calibrated

hydraulic conductivity values, as discussed below.

Surface flow routing directions and velocities are determined in part by the digital elevation map (DEM) (*Gesch et al., 2002*). The expanded model boundary encompasses the highest elevation point in Lower Michigan, at roughly 500 meters. Lake Michigan defines base level of the watershed, at an average of approximately 176 meters during the simulation period. The highest relief portion of the watershed is located around Big Rapids, MI (see Figure 4.4A).

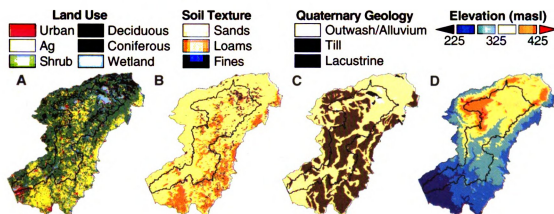


Figure 4.5: Static GIS inputs for ILHM expanded model boundary

Channel geometries were measured at each stream discharge measurement location. For measurements taken during low-flow conditions, discharge was correlated to stream width, average depth, and hydraulic radius with equations of the form $Y = AX^B$, where Y is either width, depth, or radius, and X is low-flow discharge. Next, values from flow accumulation routine in ArcGIS (which counts the number of cells upstream of every point in the watershed) were correlated to low-flow discharge at each discharge measurement location, using a power-law equation. This equation was then applied to the flow accumulation map to create an average low-flow discharge map. This map, along with the channel geometry correlations, then calculated width, average depth, and hydraulic radius across the model domain.

Lake and wetland depths are also required by the model for two purposes: 1) shallow

wetland depth information allows for a variable water table to affect inundation, and 2) deep wetlands and lakes require depths for temperature calculation. The Michigan Department of Natural Resources provides maps of the bathymetry of most of the large inland lakes in Michigan (*MDNR*, 2009). The Michigan Department of Environmental Quality has digitized some of these maps (*MDEQ Bathymetry*, 2008). For those that had not been digitized, the maps were imported into ArcGIS, georegistered, and digitized. Only wetlands larger than 0.25 km² area were digitized. For smaller wetlands with bathymetry information, average depths were estimated from the maps. Large wetlands or lakes without bathymetry were assigned a single average depth across the model area. Smaller wetlands were assigned depths based on interpretation of the National Wetlands Inventory classification scheme (Table 4.3). Note here that negative average depths indicate wetlands that are inundated for only part of the year.

Table 4.3: Wetland depths assigned based on NWI classifications. The depths are based on a sinusoidal water table fluctuation of ~2.5 m.

<i>NWI Water Class</i>	<i>Average Depth (m)</i>	<i>Typical Inundation</i>
Saturated	-1.5	Rarely
Temporarily Flooded	-1.0	Less than two weeks
Seasonally Flooded	-0.85	Dry by end of growing season
Semi-permanently Flooded	0.25	Wet during growing season
Artificially Flooded	1.0	Varies, controlled artificially
Intermittently Exposed	1.1	Exposed during drought
Permanently Flooded	2.0	Remains flooded in all years

4.3.3 Climate Data Preparation

The first step in preparing climate input data from a heterogeneous set of data sources is to create a set of utilities to interpret the native format of those data, including all data and quality control flags, into a standard format for ILHM. As Table 4.1 illustrates, some data are delivered in English units only, while others included options for output in SI units. ILHM uses the full SI standard set of units (meters, seconds, kilograms), so unit conversions were necessary in some cases. The AWOS

data included 10 m wind speed and dew point, where MAWN provides 2 m wind speed and relative humidity. While similar, these data were incompatible. 2 m wind speed values were approximated from the 10 m values using a power law wind profile (*Peterson and Hennessey*, 1978), and dew point was transformed to relative humidity using the equation from (*Lawrence*, 2005).

The data were then subjected to a further level of quality control to identify three problems: 1) individual erroneous values, perhaps due to instrument error or subsequent processing and data transfer, 2) erroneous periods within otherwise acceptable stations, 3) “bad” stations, whose measurements were consistent outliers. Individual bad measurements could be identified by specifying acceptable ranges of values; automated outlier detection methods did not fare as well because of the relatively small number of stations. These ranges were determined by visual inspection of the time series, and by plotting the frequency histograms. Erroneous periods in data were identified primarily via visual inspection, and could only be done for the MAWN dataset because of the small number of stations and relatively short time period. Entire stations were omitted whose values consistently differed from their neighbors past a specified threshold.

After these steps were complete, there were hours within the model time period during which no stations reported a particular required input value. Any gaps in data, except for solar radiation and precipitation, were filled with the day and hour average of that data type from other years. Gaps were filled only for hours with no station data present; ILHM does not assume that any given station must have data defined for all time steps, as discussed in the next section. This method was chosen, as opposed to interpolating temporally between present values, because gaps where all stations were missing tended to be wider than a few hours, thus interpolation would not preserve diurnal cycles.

4.3.4 Spatial Distribution of Climate Data

Data input to ILHM must specify three things: 1) the time of the measurement, 2) the value, and 3) how to distribute those data over the model domain. Spatially-distributed datasets, such as NEXRAD radar or satellite-sensed LAI inherently contain all three types of information. Gage data, however, must be interpolated to the ILHM domain. As each interpolation must be done every hour for over a dozen input data types, the scheme must be very fast, or be run prior to model execution and amortized over multiple model runs. Also, the scheme must handle stations dropping off- and coming on-line.

An early version of ILHM applied to small sub-watershed of the MRW, published in *Hyndman et al.* (2007), accomplished this through a fast inverse-distance-weight method. Such a method, however, is fundamentally unsuitable for hourly interpolation at larger spatial scales. As weather systems move across the model domain, they cause abrupt changes in data values, for instance temperature and precipitation. If the gage spacing is greater than the distance across a weather front, as is often true for the MRW, any interpolation method that includes more than one station value and assumes stationarity will fail to accurately represent the true field. This is particularly true with precipitation data, as the intensity of precipitation is critical to canopy interception and runoff generation. A more appropriate method is to distribute a single gage data values within Thiessen polygons. The precipitation data preparation section below describes how this method can be modified to better reproduce temporally-aggregated fields.

4.3.5 Solar Radiation Preparation

Solar shortwave radiation drives the hydrologic cycle, and is a critical input to ILHM. However, for much of the model period only a single gage recorded solar radiation in the watershed, near Muskegon (see Figure 4.4A). Comparison of this gage with the

MAWN data beginning in 1996, shows that this gage does a poor job capturing total solar radiation, therefore it was not used in this study. To provide solar radiation prior to the availability of MAWN data, more fully represent its spatial variability, and to split total solar radiation into beam and diffuse components, a real-sky radiation model was used. This model requires air temperature, relative humidity, and opaque cloud fraction.

The only cloud fraction data available came as sky condition observations from airport weather stations. These observations, (clear, overcast, etc.), crudely represent cloud cover conditions in oktas (1/8 increments). However, in 1996 the stations apparently changed their reporting, causing an abrupt shift in the distributions of cloud cover observations (*Dai et al.*, 2005). This shift had to be corrected to force consistency across the 28 model years. The distributions of cloud cover fractions from 1980 - 1996, and 1997-2007 were calculated. The distribution of data from 1980-1996 was shifted to match the later period by adjusting the mean and range of values within each quartile of the distribution.

Cloud cover observations are used to calculate real-sky radiation from clear-sky modeled radiation. Even with the available observations, the model cannot represent the full range of values present in the gage data, even though the annual mean values were well simulated. To partially account for the monthly variability in cloud cover that is not fully captured by the sky condition observations, the gaged data were compared to simulated values, and a monthly-average weighting was calculated. For the period 1997-2007 individual monthly weightings were used, while for 1980-1996, an average monthly correction values from 1997-2007 were applied as no gage data were available for that period. After these corrections, the model captures much of the variability present in the gaged data, and provides additional spatial information not available from the more limited MAWN locations. Figure 4.6 plots annual model-averaged total solar radiation. The period of 1980-1996 includes only mod-

eled radiation, while 1997-2007 includes modeled and gaged values. Notice that the mean radiation is reasonably well simulated, but inter-annual variability is not fully captured prior to the availability of radiation data.

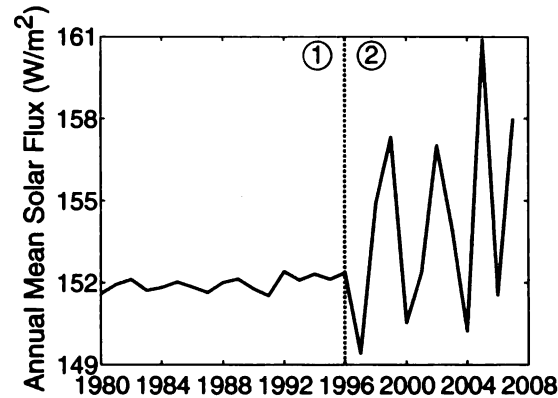


Figure 4.6: Plot of annual model-averaged solar radiation, from 1980 through 2007. There are two primary periods, 1) only cloudiness observations are available for short-wave radiation modeling, 2) cloudiness observations along with limited solar radiation data are available.

4.3.6 Precipitation Preparation

Within and around the model domain, many more stations report daily precipitation than hourly (see Table 4.1). To incorporate the spatial variability present in the daily data into the hourly precipitation dataset, a four-step procedure was used.

1. For each model hour, hourly precipitation data are interpolated to a grid using Thiessen polygons (this is identical to a nearest-neighbor interpolation). This method was used rather than a scheme that uses multiple stations to predict unknown locations because of the issues with weather fronts mentioned previously.
2. Daily precipitation totals are not recorded at identical times at each station. Most of the stations recorded near midnight GMT, while others recorded near

noon GMT. Because of this, the daily data could not be directly interpolated to create a single daily precipitation grid. The daily data at each station were summed weekly, and interpolated to the grid with a inverse distance weighting scheme.

3. The hourly grids were summed weekly, a weight value for each grid cell was calculated as the ratio of the summed daily and summed weekly grids.
4. For each hour in the week, the weight grid was multiplied by the hourly precipitation grid.

The net effect of this procedure is that total precipitation values in each cell are fixed by inverse distance weighted interpolation of daily precipitation stations, while the hourly distribution is governed by the nearest hourly station.

NEXRAD precipitation data is capable of providing relatively high resolution distributed precipitation data at hourly intervals. For the MRW, a 4 km Level III calibrated, gage-adjusted dataset became available in 1996 (*Seo, 1998*). To determine whether NEXRAD data could be used in place of interpolated station values, monthly grids of each were compared. A time-series from 1996 - 2007 of the model-average monthly precipitation estimates are plotted in Figure 4.7. From visual inspection, there are three distinct periods of NEXRAD accuracy. From 1996 - 2000, model-average NEXRAD precipitation matches interpolated station values quite poorly. Beginning in April of 2000, growing-season NEXRAD values seem to provide reasonable monthly precipitation estimates. After March of 2003, NEXRAD and station values compare favorably in all months.

This study assumes that if monthly NEXRAD precipitation compares favorably with interpolated station values, the NEXRAD dataset then likely provides more accurate hourly distributed estimates of precipitation over the watershed. Based on the comparison in Figure 4.7, no NEXRAD data prior to 2000 were used. Beginning

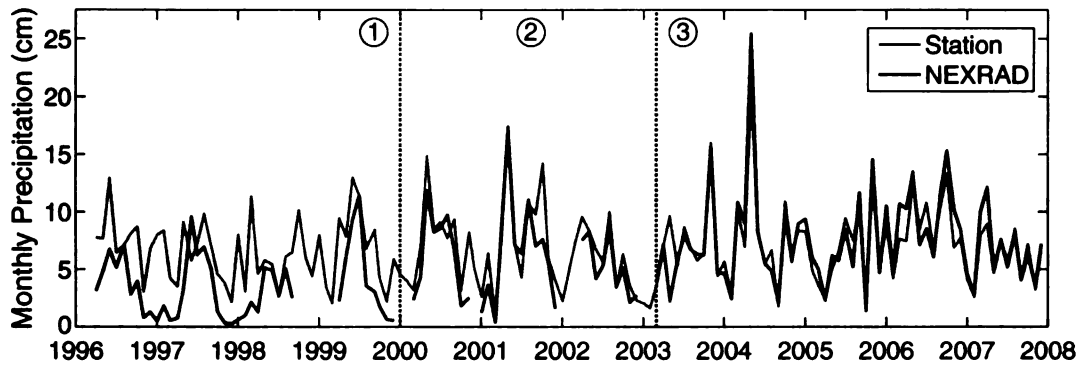


Figure 4.7: Plot of monthly model-averaged precipitation from daily interpolated NCDC precipitation gages, and NEXRAD radar from 1996 - 2008. The quality of the NEXRAD data can be differentiated into three periods, 1) all monthly observations are a poor fit to gage data, 2) NEXRAD provides a good match to months with liquid precipitation only, and 3) NEXRAD data and gage data match closely in almost all months.

in April of 2000, NEXRAD data were used in place of station values for April - October. After April 2000, NEXRAD data were used exclusively.

4.3.7 Leaf Area Index (LAI) Preparation

Leaf area index (LAI), a measure of the one-sided leaf area per unit area of horizontally-projected ground surface, drives seasonal vegetation change in ILHM. LAI controls vegetation height, canopy moisture and radiation interception, canopy water capacity, albedo, and canopy resistance to vapor and momentum transport. In 2000, a 1 km, 8-day LAI product (MOD15A2) was released, calculated from data captured by the Moderate Resolution Imaging Spectrometer (MODIS) instrument carried on-board NASA's TERRA and AQUA satellites (*Knyazikhin et al.*, 1999). Prior to this, however, there are no readily-available distributed LAI products.

To extend the LAI record as far backward as possible, a multi-step procedure for deriving LAI from normalized difference vegetation index (NDVI) values was developed. NDVI is calculated using two spectral bands, the red ($0.6 \mu\text{m}$) and near infrared

(0.9 μm), and is calculated as

$$NDVI = \frac{NIR - RED}{NIR + RED}. \quad (4.11)$$

Prior to MODIS, instances of the Advanced Very High Resolution Radiometer (AVHRR) instrument aboard several NOAA satellites have been in orbit since 1981. Two NDVI products derived from AVHRR data are readily available, the 15-day 8 km GIMMS global dataset, and the 7-day 1 km continental US dataset (see Table 4.1) (*Eidenshink, 1992; Tucker et al., 2005*).

The GIMMS dataset was specifically produced to correct certain errors present in the AVHRR data, including sensor and orbital drift (*Tucker et al., 2005*). Its coarse resolution, both temporally and spatially, are less than ideal for this application. The GIMMS grids were used to correct the 1 km NDVI. The mean value of the 1 km cells were calculated for each 8 km GIMMS grid cell. A weight grid was then calculated as the ratio of the 8 km values to the upscaled 1 km data. This weight grid was then multiplied by the 1 km data. This re-weighting allows the corrections present in the GIMMS dataset to be applied to the higher resolution 1 km, 7 day data.

As discussed in *Baret and Guyot (1991)*, the theoretical relationship between NDVI and LAI is an exponential of the form

$$LAI = C_A \cdot \exp(C_B \cdot NDVI) \quad (4.12)$$

where C_A and C_B are constants that vary by biome type. Values of these constants for each land use class were optimized by comparing, AVHRR-derived NDVI to MODIS LAI values. The 1 km NDVI dataset most closely matches the MODIS resolution, and four years of overlap between the two datasets were available for the following procedure:

1. Select all cells whose fraction of a given land use exceeds 90%.

2. Extract LAI and NDVI values for these cells. This requires re-interpolating one of the datasets to match the other spatially. Here, the LAI data were resampled to 1 km resolution.
3. Temporally interpolate one of the datasets to match the other for each model cell. Here, the LAI data were linearly interpolated to match the dates of the NDVI grids.
4. Optionally, randomly select a subset of these points for computational efficiency. 150 points from each 7 day grid were chosen for this analysis.
5. Fit an equation of the form shown in Equation 4.12 to the LAI and NDVI data, optimizing the values of C_A and C_B . Here, a bound optimization was conducted in MATLAB, using a mean-absolute residual objective function.
6. Calculate LAI from NDVI using Equation 4.12 and the constants optimized from step 5. Constrain the LAI values produced not to exceed the maximum and minimum values in the MODIS dataset.

Optimized values of C_A and C_B are shown in Table 4.4, along with their R^2 values. For all land use classes except urban, the empirical function provides reasonable estimates of LAI from NDVI grids.

Table 4.4: Table of optimized constants for the NDVI $\rightarrow$ LAI mapping function 4.12, and coefficient of determination R^2 . Open water and barren are assumed to have $LAI = 0$ in ILHM.

<i>Land Use</i>	C_A	C_B	R^2
Urban	0.103	4.13	0.33
Agriculture	0.082	4.03	0.67
Open/Shrub	0.103	4.99	0.70
Deciduous	0.091	4.79	0.74
Coniferous	0.152	4.21	0.62
Wetland	0.096	4.39	0.67
Open Water	-	-	-
Barren	-	-	-

After calculating LAI grids from the 8 km and 1 km NDVI grids, the result is a heterogeneous resolution dataset from 1981 through 2007. To resample these grids to model resolution (425 m), the following downscaling procedure was applied:

1. Identify cells for each land use class that exceed 75% of that type, and extract LAI values for each of these cells.
2. Interpolate these cells to a grid at model resolution using the nearest neighbor. This produces eight grids for each input LAI grid, one for each land use class, representing spatial distribution the LAI of that land use at a point in time.
3. Add these eight grids together based on the fraction of each land use class within each model cell. For example

$$LAI = f_{urb}LAI_{urb} + f_{ag}LAI_{ag} + f_{open}LAI_{open} + f_{con}LAI_{con} + f_{dec}LAI_{dec} + f_{wet}LAI_{wet}$$

where the sum of the land use fractions f is 1.

4. The result from Step 3 is a grid of LAI values at model resolution. Now, correct the cell-by-cell values to match the coarse input dataset by resampling this grid to the original resolution (8 km, 1 km, or 0.92 km) and calculating the mean of the model grid values. Calculate a weight grid as the upsampled mean divided by the original LAI value.
5. Resample this weight grid to model resolution, and multiply it by the model resolution LAI grid, resulting in a grid whose values match the original LAI data (or the NDVI-derived LAI values) at the coarse scale, but incorporate finer resolution information from land cover.

The result of these procedures is a 27 year continuous time series of LAI values from two different instruments and three different sources.

Consistent spatial biases remain between sources, primarily as a result of factors that control LAI other than NDVI. This can be easily visualized by calculating the

average annual maximum LAI from each dataset (not shown). In particular, areas of mixed land use exhibit significant variations between the downscaled NDVI-derived and MODIS-derived LAI. However, the spatial bias varies consistently throughout the year, thus continuous spatial bias correction is needed. The following procedure was applied:

1. For the period of overlap between datasets (7 years for 8 km NDVI and MODIS, 4 years for 1 km NDVI and MODIS), temporally resample one of the datasets linearly (they have already been downscaled to the same spatial resolution) to match the other. For this study, both datasets were resampled to the same 7 day intervals.
2. Calculate the ratio, LAI/NDVI for each grid. Then, calculate the Julian-day average ratio across all overlap years. The result is a set of ratio grids at 7 day intervals.
3. For the temporally-resampled NDVI-derived LAI grids, multiply each grid by the ratio grid for the corresponding Julian day. For instance, day 21 of 1992 is multiplied by the ratio grid for day 21 from step 2. Constrain the cell-by-cell maximum and minimum values to those observed in the MODIS LAI grids.

The result from this last procedure is that the long-term spatial averages of the NDVI-derived LAI grids match the MODIS LAI data. This was deemed essential for maintaining consistency within the model time period.

Two more steps are then applied. The three data sources are composited, GIMMS data are used prior to 1989, following which the 1 km NDVI-derived LAI are used until MODIS data become available in 2000. Periods during which values were missing for over 1 month are filled with the Julian-day average from all model years. This includes the entire year of 1980, and a several-month period in late 1989. Plotted in Figure 4.8 are the monthly model-average LAI values after all steps are complete. In

general, the three datasets are in good agreement, though peak LAI values derived from NDVI still exceed those from MODIS by 10-20% in some years.

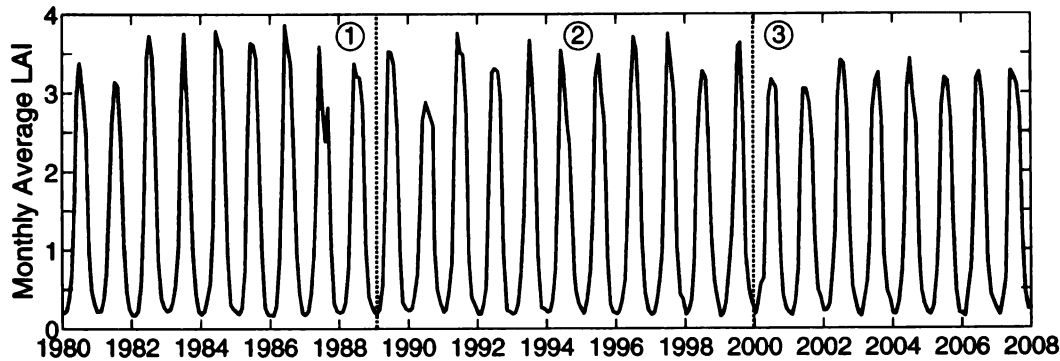


Figure 4.8: Plot of monthly model-averaged LAI from three sources: 1) 8 km GIMMS AVHRR data, converted from NDVI to LAI, 2) 1 km AVHRR data, uncorrected, converted from NDVI to LAI, and 3) MODIS LAI data.

4.3.8 Model Parameters

Model parameters in ILHM fall into four general classes: 1) physical constants with widely-agreed upon values (Table B.4), 2) land use or soil texture dependent values with readily-available literature values that generally agree but vary from source to source, 3) descriptive physical parameters that are constant across the model, but are expected to be domain-independent, and 4) site-specific physical parameters whose values must be specified by the modeler. Tables of physical constants are provided in Appendix B, along with reference citations.

Land use and soil texture dependent parameters are those that can generally be expected to be independent of the model domain location, thus literature values can reasonably constrain these parameters. Several sources of soil texture parameters were used (*Saxton and Rawls, 2006; Schaap et al., 1998; Stieglitz et al., 1997*). Values of parameters in common among these sources were averaged. Some parameters varied little from source to source, including field capacity and total porosity that differed roughly 10% among sources. Saturated hydraulic conductivity and residual

porosity varied substantially, with individual sources differing more than 50%. All other parameters values came from a single source. Land use parameters are less commonly available, therefore a single source was used for these parameters. All parameter values are listed in Table B.1. Domain-independent physical parameters relate to soil and canopy processes, but are assumed to be independent of land use or soil texture. Albedo and emissivity parameters fall into this category as well. These are listed in Tables B.5 and B.6.

Table B.7 lists those parameters whose values are domain-specific. These include the impermeable fraction of urban areas, wetland bottom hydraulic conductance, wetland recession constant, deep unsaturated zone velocity, and saturated zone domain hydraulic conductivities. Urban areas were assumed to be 20% impermeable, a value that represents the paved portion of impermeability in residential urban areas (*Akbari et al.*, 2003). Other types of imperviousness, like rooftops, quite often drain into the subsurface in residential areas. This value likely underestimates impermeability in dense urban, commercial, and industrial areas that make up a very small (and ungaged) portion of the watershed. Future improvements will parameterize urban at Level 2 or 3 classification (*Anderson et al.*, 1976) and incorporate this variability.

Subsurface hydraulic conductivity values were preliminarily calibrated to a small subset of the static water level measurements using an early steady-state version of the groundwater domain model. These are listed in Table B.9. The velocity of a wetting front in the deep unsaturated zone is closely related to these values, and was estimated using well hydrographs from the nearby Grand Traverse Bay Watershed (Table B.8).

A single wetland recession constant was derived from the Evart gage as the log-linear slope of the recession portion of the hydrograph. There are two wetland bottom hydraulic conductance values, one for wetlands connected to the water table, another for disconnected wetlands. The connected wetland value was calculated assuming a

1.5 m d⁻¹ horizontal hydraulic conductivity, with a vertical anisotropy of 10 (i.e the ratio of horizontal to vertical conductivity), and a wetland bed sediment thickness of 2 m. Conductivity values were similar to those calibrated in the groundwater model. The disconnected wetland conductance was arbitrarily chosen to be 8.6e-5 m d⁻¹.

4.4 Uncalibrated Model Results and Discussion

ILHM calculates all near-surface hydrologic fluxes and storages with hourly timesteps for each grid cell, providing a wealth of information to validate the model's predictions. There are two readily-available types of data in many regional watersheds including the MRW, static water levels collected from drinking water wells, and stream discharge values measured by gaging stations. Each static water level measurement provides a snapshot of the groundwater head at a single point in time and space. Stream discharge measurements integrate a much larger area and provide a means for bulk comparison of fluxes.

Comparing model predictions to both types of data is essential for integrated surface- and ground-water modeling. Within the groundwater domain, hydraulic head is inversely sensitive to hydraulic conductivity and groundwater recharge *Scanlon et al.* (2002). A model that reproduces measured heads will not necessarily accurately describe groundwater discharge. Similarly for the surface and near-surface domain, a model may reproduce stream discharge fluxes accurately while not accurately describing the partitioning of the water budget between groundwater recharge, overland flow runoff, and near-surface throughflow.

Here, uncalibrated model predictions of daily stream discharge and hydraulic head are compared to data described in Table 4.2. With the exception of a preliminary, steady-state manual calibration of hydraulic conductivity values, the model has not

been calibrated to specifically match either fluxes or heads. Model calibration could substantially improve fits to measured data. Nevertheless, uncalibrated results are presented here to provide a more open assessment of the capabilities of ILHM.

Of the six USGS gages in the watershed with data records longer than a few years (Table 4.5), only three, Muskegon River at Evart, Clam River at Vogel Center, and Bear Creek near Muskegon, have continuous records for the entire model period, 1980 - 2007. Daily hydrograph comparisons for those three gages are plotted in Figure 4.9. Note that they are plotted in order of descending catchment area. In all three catchments, ILHM describes the overall character of the hydrographs, particularly at the largest scale.

Close inspection of the hydrographs reveals that hydrograph recession is well-represented by ILHM. While ILHM has a site-specific recession constant for surface water bodies, the net slope of the hydrograph recession curve is the combined response of surface routing and evaporation. Particularly at the largest scales, rise times in ILHM are slower than in gage data and peak heights tend to be lower, especially during the spring. This suggests that a mechanism responsible for generating part of the peak discharge response is not fully described. Because impermeable runoff is not a significant factor in much of the watershed, the likely reason for slower rise and lower peaks within the model is that seasonal inundation is not completely captured by the simple model described above. A run-time coupling of the surface and groundwater domain modules will allow a much more detailed examination of this phenomenon.

Stream baseflows, critical for ecosystems and water resources, are well-described by ILHM. Summer low flow values depend on both modeled groundwater discharge as well as calculated evaporation from the surface drainage network. Response to precipitation during the summer is an important indicator of the sources of discharge. Greatly muted summer precipitation response relative to spring suggests that seasonal wetlands play an important role in generating peak flows. Notably, the flood

event of late summer 1986 is not well described in ILHM. This indicates that groundwater played an important role in the flood, as seasonal wetlands became inundated and contributed to a much larger runoff response. The one-way coupling currently implemented in ILHM cannot fully account for such an event. The broad base of runoff peaks are well described by this simulation, most of which is groundwater discharge. This is another indication that baseflow, and therefore groundwater recharge, are accurately simulated.

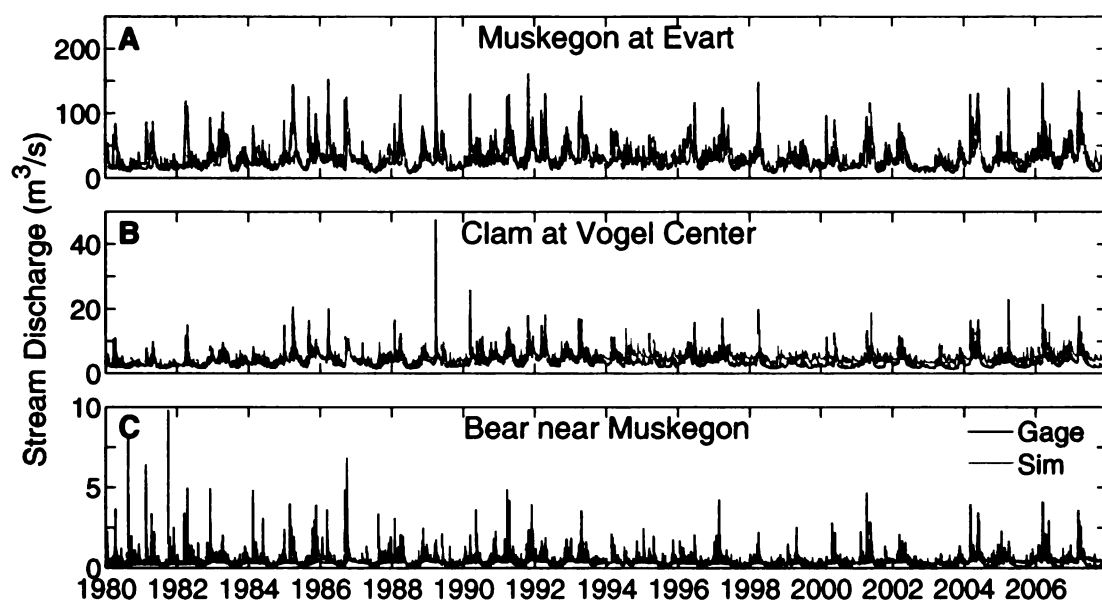


Figure 4.9: Simulated and observed (gage) hydrographs for three USGS stream gages, A) Muskegon River at Evert, B) Clam River at Vogel Center, and C) Bear Creek near Muskegon.

To more quantitatively compare ILHM stream discharge predictions to gaged values, Figure 4.10A plots simulated and observed values using an unbiased residual calculated as $(sim - obs)/obs$ from the Evert gage. This figure demonstrates that ILHM predicts the bulk of daily low-flow discharge to within $\pm 30\%$. This plot also shows the tendency for the model to respond somewhat excessively to summer precipitation events that tend to occur during low discharge periods. Similarly, the higher discharge periods that most often occur during the spring are not fully captured.

As shown in Table 4.2, three synoptic stream discharge datasets were collected at or near annual low-flow levels across the MRW. These data are plotted in Figure 4.10B, along with simulated model baseflow. Note that stream discharge varies across three orders of magnitude in this figure. In general, there is no significant trend in the data. Data collected during August of 2004 were within the central portion of the watershed. ILHM over-predicts baseflow during this period, suggesting either that the model over-predicts recharge or that the simulated water table relaxes too slowly during the growing season. Note a portion of the October 2004 and July 2003 data are under-predicted. This is likely due to small precipitation events shortly before the data were collected that continued to contribute some surface-originated discharge.

In Figure 4.10C, simulated hydraulic heads are compared to observed static water levels. A best-fit line (not shown) deviates from the 1:1 line by less than 5%, indicating no significant head-dependent trend. The region between 230-260 m, where ILHM tends to underestimate head, is due to a sequence of large impoundments along the main channel of the Muskegon River (Croton, Hardy, and Rogers dams) that are not included in the groundwater model. Other discrepancies between simulated and observed heads, such as a tendency toward over-prediction at the highest head levels, are likely due to incompletely calibrated hydraulic conductivity values.

In general, ILHM simulates hydraulic heads within ± 5 m, with little bias (Figure 4.10D). Here, the median annual head residual $sim - obs$ is shown as a heavy black line on top of a shaded region that outlines ± 1 standard deviation in annual residual from the median. ILHM slightly under-predicts heads during the 1980s, and slightly over-predicts heads during for the remainder of the model period. Overall, the mean residual bias is less than 0.20 m.

A common quantitative estimate of model performance is the Nash-Sutcliffe effi-

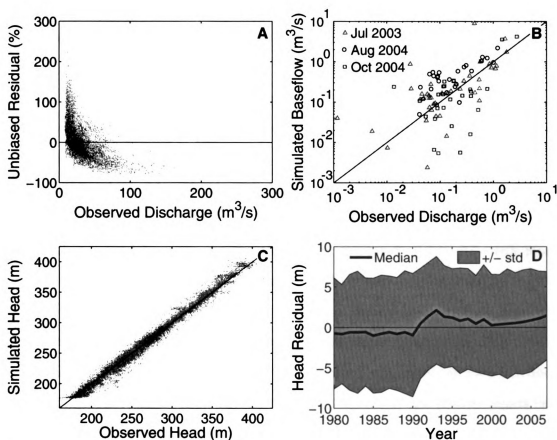


Figure 4.10: Plots of simulated and observed streamflows and groundwater head, observation data are described in Table 4.2. A) Observed daily discharge at the Evart, MI gage vs. unbiased residual $(sim - obs)/obs$, B) observed discharge from 105 points across the MRW during low-flow conditions vs. simulated baseflow, C) observed groundwater head vs. simulated head, and D) annual median and ± 1 standard deviation of the head residuals $sim - obs$. All horizontal and diagonal lines are reference lines, not linear fits.

any

the

value

value

the

the

the

the

the

the

the

the

the

the

the

the

the

the

the

the

the

the

the

the

the

the

the

ciency E Nash and Sutcliffe (1970),

$$E = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (4.13)$$

where O_i are observed values (of a time series) and P_i are predicted or simulated values. Essentially, E is a measure of the deviation of predictions from observed values normalized by the variance of the observed dataset. Here, a log-transform of the observations and predictions are used in order to more evenly weight the efficiency metric toward all flow levels (Krause *et al.*, 2005). Values of $E > 0$ indicate that the predictions describe more of the variance in the observations than the mean of the observations alone.

Annual E values of log-transformed daily stream discharge are plotted in Figure 4.11 for each of the six USGS gages. The three gages with the largest catchments, Evart, Croton, and Newaygo, have efficiencies in the range -0.2 to 0.6, with most years near 0.4. The Clam and Little Muskegon River gages, particularly from 1994 - 2007, generally have negative efficiency values. This is because of an apparent overestimation of groundwater recharge within these basins. The gage on Bear Creek, with the smallest catchment, has positive efficiencies for much of the model period, which is notable because random errors in model parameters and climate inputs can be much more significant at this scale.

Plotted in Figure 4.12A are water-year (Oct-Sep) total discharge flux error values ($sim - obs$) normalized by the area of each catchment, and expressed in cm. This flux error is directly comparable to total precipitation, which averages approximately 85 cm across the watershed. Thus, an 8.5 cm flux error in stream discharge is approximately 10% error of the total water budget. For most of the gages, excepting the early period in Bear Creek, variability in flux error is less than 10 cm annually, and most of the gages fall within 10 cm most years. In particular, the larger catchments of MR at Evart and Croton are simulated more accurately than the smaller catchments, suggesting that random errors either in input fluxes or model parameters account for

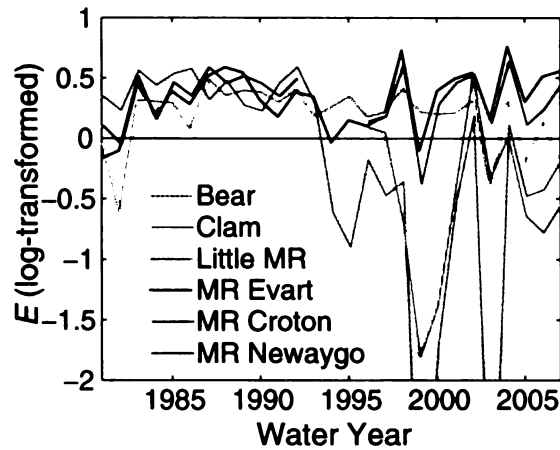


Figure 4.11: Plot of water-year (Oct - Sep) Nash-Sutcliffe efficiency of the log-transformed simulated and gage streamflow values for all six USGS gages. “MR” refers to “Muskegon River”.

much of the error in smaller gage basins.

During the early simulation period, Bear Creek (Figure 4.9) responded to precipitation events with rapid hydrograph rise. This has, over time, lessened. This may be to the requirement that new development be accompanied by retention ponds or infiltration basins to reduce peak runoff response. As more of these were installed, the system responds more slowly to precipitation, and with less intensity. This kind of time-varying land use effect can be difficult to incorporate in a regional scale model, and is a source of error across the basin.

Average monthly flux errors at each gage (Figure 4.12B) exhibit the same general behaviors discussed for Figure 4.9; spring flows are under-predicted, and summer flows slightly over-predicted. For all gages, except the Bear Creek gage which is likely influenced by anthropogenic water withdrawals near Muskegon, the general shape of monthly flux errors are similar. This implies that the processes generating monthly error variability are relatively uniform across the model. Indeed, they correspond quite closely if offset with a scalar constant at each gage. The gage-dependent bias is likely related to spatially distributed model parameter values, though no parameter sensitivity or calibration was conducted here.

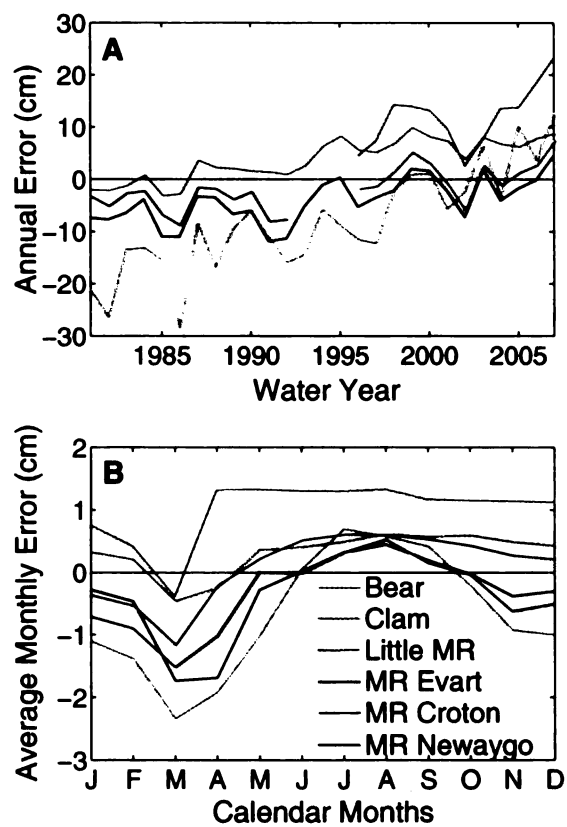


Figure 4.12: Plots of stream discharge error $sim - obs$ for all six USGS gages normalized by the area of each gage catchment. Plotted in A) are water-year error values, and B) are average monthly error values across the 27 water years in the simulation.

Over the nearly three decade model period, the influence of time-variant inputs on model error is reduced. Table 4.5 lists daily efficiency values for the entire period, along with total flux errors at each gage. Of the gages active during the entire period, the largest flux error observed is 7.9 cm, under 10% of the water budget. The larger catchment areas, including the MR @ Evart and Croton, had flux errors of 1.7 and 3.1%, respectively. This level of error compares favorably to other methods of estimating regional water budgets (*Scanlon et al.*, 2002; *Arnold and Allen*, 1996; *Winter*, 1981).

The total daily efficiencies are similar to values reported elsewhere for uncalibrated distributed rainfall-runoff models. The first Distributed Model Intercomparison Project evaluated 12 different hydrologic models for 8 catchments ranging between 65 and 2484 km² (*Smith et al.*, 2004). Model efficiencies varied among models and basins, but in general, larger basins had better efficiencies. Average uncalibrated model performance for large basins was generally less than 0.5, and in some cases negative. Calibration improved all model efficiencies, generally increasing them by 0.1 to 0.2 (*Reed et al.*, 2004).

Table 4.5: Total Nash-Sutcliffe efficiency and average annual flux error values for each of the 6 USGS gages. Efficiency values are calculated using log-transformed simulated and gaged values. Average annual flux error is *sim* – *obs* divided by the catchment area of each gage. The final column is average annual error normalized by average annual precipitation for that catchment.

<i>Gage</i>	<i>Area (km²)</i>	<i>Avg. Precip. (cm)</i>	<i>E log-transformed</i>	<i>Avg. Error (cm)</i>	<i>Avg. Error (% precip)</i>
MR @ Newaygo	6,086	81.1	0.43	-4.43	-5.46
MR @ Croton	5,991	85.4	0.43	1.24	1.45
MR @ Evart	3,711	79.0	0.45	-4.16	-5.27
Little MR	894	89.2	-0.32	12.01	13.47
Clam	629	82.4	0.11	3.69	4.48
Bear	43	87.6	0.25	-8.12	-9.26

4.5 Conclusions

The Integrated Landscape Hydrology Model (ILHM) is capable of fine-resolution, large-domain simulations. It is a fully distributed, process-based integrated hydrologic model designed to be computationally efficient and highly flexible. While there are numerous watershed and integrated hydrologic modeling codes, few are capable of fully-distributed simulations over large domains. Originally detailed in (*Hyndman et al.*, 2007), the code has been significantly improved. It now incorporates a full range of hydrologic processes covering upland and wetland domains. Currently, the coupling between surface and deeper subsurface domain modules is one-way, though full run-time coupling is under development.

ILHM is structured to handle both gaged and fully-distributed input data sources at a variety of resolutions. Heterogeneous sources can be included to provide continuous records over many decades. Extensive efforts were undertaken to provide a consistent, quality-controlled climate and landscape input dataset. Coarse cloud cover observations along with a sparse set of total solar radiation data were modeled to provide a temporally- and spatially-variable real-ground solar radiation beam and diffuse components. Methods for compositing gage and radar precipitation data are detailed, as well as a discussion of proper spatial distribution of gaged datasets. Leaf area index data that describe phenological development within each model cell came from three separate sources, including two sets of data that provided only vegetative indexes. These index values were transformed to LAI with locally-optimized empirical functions. Following this, a land use-based downscaling method was detailed to provide higher resolution LAI estimates within each coarse data cell.

Comparisons of uncalibrated ILHM predictions of stream discharge and groundwater heads to observations indicate that the code simulates most of the important hydrologic processes, and accurately describes fluxes both regionally at locally. General hydrograph shapes are well-represented, with tendencies toward under-prediction

of peaks in the spring and over-prediction of peaks during the summer months. During all periods, baseflow is reasonably estimated, with regionally-dependent over-prediction present at some gages during low-flow periods. Simulated hydraulic heads display no general biases, with modest residuals at this spatial scale. Total model flux errors are small, even for this uncalibrated simulation.

Based on these results, the current ILHM process modules are well suited for simulating unconsolidated, loosely-coupled surface- and ground-water systems. Other environments will dictate modification or development of new process modules, which the code is structured to allow relatively easily. Moreover, multi-decadal simulation periods can be run on modest desktop machines, though the surface and deep unsaturated zone modules are highly parallel, and the code is written to allow subdivision to many nodes on a cluster. The need to calibrate very few domain-specific parameters allows ILHM to be readily and rapidly applied to ungaged and remote basins, with reasonable confidence in predictions.

Acknowledgments

Funding for this work has been provided by the Great Lakes Fisheries Trust Muskegon River Initiative, the National Science Foundation through its Water Cycles program (EAR-0233648), and the Michigan State University Center for Water Sciences. We would like to thank former students including Cheryl Kendall, Jeff Eggleston, Nicklaus Welty, and Dushmantha Jayawickreme for their assistance in collecting stream discharge data and preparing input datasets, including the digitization of inland lake maps.

CHAPTER 5

Influence of Land Use, Climate, and Soils on Recharge and Evapotranspiration Across a Regional Watershed

5.1 Introduction

Climate and land use change are altering the timing, distribution, and quantity of hydrologic fluxes worldwide. During the last century, air temperatures (*NCDC Temp*, 2008), precipitation (*Center*, 2001), and stream discharge (*USGS*, 2008) have increased in the Muskegon River Watershed (MRW), a regional catchment in central lower Michigan, USA. Over the same period, land use shifted dramatically; forests expanded to cover much of their pre-settlement range, urban areas expanded considerably, and the extent of agricultural declined by half *Piajnowski et al.* (2007).

Predicting the heterogeneous local and regional responses to these changes is critical to managing water resources for both human uses and natural ecosystems. A relative abundance of shallow groundwater in Michigan is a valuable resource for community and agricultural water supplies, but also provides the bulk of annual stream discharge in this baseflow-dominated system . Numerous segments of stream channels within

the MRW provide habitats for economically-important cold-water spawning fishes due to the robust supply of groundwater discharge during summer months (*O'Neal*, 1997).

One of the primary consumptive water uses in many basins, irrigated agriculture, is not currently widespread in the MRW (*Morenz et al.*, 2005). Spring snowmelt provides a reliable source of soil moisture storage for the early growing season, complemented by relatively wet months later in the growing season. Furthermore, much of the surface soil textures across the MRW are sands or loamy sands with little plant-available water capacity. These marginal soils are farmed in areas of low relief, but are not well suited to crop systems such as seed corn that generate demand for irrigation in this region (*Morenz et al.*, 2005; *USDA-NRCS*, 2008).

During the 21st century, temperatures across the region are forecast to increase between 3 and 4 °C under the A1B emissions scenario, while annual precipitation may increase 10% (*Christensen et al.*, 2007). Higher temperatures would hasten the arrival of the growing season, and extend the period of moisture extraction from soil storage. Despite a forecast annual increase in precipitation, climate model ensembles show no change in precipitation during the growing season. Increased demand for moisture with a static supply may lead to drier soils throughout the region and threaten the viability of some non-irrigated farming practices.

Recent demand for biofuel crop production may provide economic incentive for marginal lands to shift crop systems, or perhaps be put back into production (*NASS*, 2007). Expanded production of corn and soybeans into sandier soils may increase the demand for irrigation to compensate for poor soil moisture storage capacity. Biofuel extraction and processing facilities are also major water-users, consuming 40 L of water per L of corn ethanol produced (*Pimentel and Patzek*, 2005).

Aside from shifting the consumptive demand for water resources, land use and climate change will alter the water budget of the region. Because of coarse-textured soils

an associated high infiltration capacities, most of precipitation is either evapotranspired (ET), or enters the groundwater as recharge and is subsequently discharged to surface water systems. Impervious surfaces in urban areas of the MRW increase runoff, and alter the balance between ET and groundwater recharge in upland areas. The succession of abandoned agriculture to shrub, and ultimately forests, also modifies the water budget through increased canopy interception, decreased soil evaporation, and other changes.

Effective management of existing water resources and the sustainability of land use practices requires a quantitative understanding of the interactions among climate, land use, and terrestrial hydrology. These interactions are facilitated and regulated by various aspects of the landscape, including soil textures, relief, and surface water bodies. The complexity of interactions among these factors require more than simply a conceptual or theoretical understanding of the impacts. Furthermore, the heterogeneity of landscape, land use, and climate demands a predictive capacity that addresses all scales, from hectare to basin.

The Integrated Landscape Hydrology Model (ILHM) is a fully-distributed hydrologic model capable of fine-resolution simulations over large-domain. It consists of four domain modules for surface water, canopy and root zone, deep unsaturated zone, and saturated groundwater (*Hyndman et al.*, 2007). Within those domains, sets of individual physical process models simulate hydrologic fluxes such as evaporation, transpiration, infiltration, groundwater recharge, overland flow runoff, and groundwater flow. ILHM is written to run on desktop computers, but be rapidly ported to high performance clusters if necessary. Chapter 4 described the improvement of ILHM since the publication of *Hyndman et al.* (2007) and the development of new capabilities including more sophisticated wetland process modules. Also detailed in Chapter 4 are a set of methods to develop spatially and temporally consistent input datasets to drive a fine-resolution, regional-scale model.

U

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

Uncalibrated predictions of streamflow and groundwater head for the MRW compared favorably to several different sources of observation data. Over the simulation period 1980 - 2007, the model simulated total stream discharge to within 5% of precipitation for the largest basins, and between 5 and 15% for smaller catchments. Nash-Sutcliffe efficiencies, a measure of the ability of the model simulate observation dataset relative to its variance, were generally around 0.45 for the larger basins, and between -0.3 and 0.25 for smaller watersheds. Nearly 50,000 measurements of static water table levels, collected by drinking water well drillers, were used to determine the accuracy of hydraulic head predictions within the model. The mean head residual was less than 20 cm.

This Chapter examines the spatial and temporal variability of key hydrologic fluxes including groundwater recharge and evapotranspiration. Precipitation and temperature vary significantly over the model domain, due largely to the effects of nearby Lake Michigan. Imprinted on these regional scale trends are land use and soil texture variability at the hectare scale. ILHM is capable of resolving variability across these scales, which is necessary to quantify the influences of climate, land use, and soil textures on the terrestrial hydrologic cycle.

5.2 Methods and Study Site

The Muskegon River Watershed (MRW) is within the Lake Michigan/Huron basin, part of the Laurentian Great Lakes (Figure 5.1). The MRW covers approximately 7,400 km², and encompasses several large lakes (Houghton, Higgins, Mitchell, Cadillac, Muskegon) and wetland complexes. The main stem Muskegon River is a fifth order stream, with a single fourth order tributary, the Little Muskegon River. Stream flow within the MRW is dominated by groundwater, which contributes approximately 85 percent of mean annual discharge (see Chapter 4). Its major urban centers are

Muskegon (pop. $\sim 40,000$), Big Rapids ($\sim 11,000$), and Cadillac ($\sim 10,000$). The MRW has relatively little relief, with elevations between 175 and 500 meters.

During the late 19th century, the MRW underwent a rapid period of deforestation leaving a virtually denuded landscape by the late 1890s. By the first decade of the 1900s, less than 20% of the watershed was forested, while nearly 50% was managed agriculture (Piajowski *et al.*, 2007). Subsequent abandonment of marginal agricultural lands led to significant reforestation with only moderate urbanization during the remainder of the 20th century.

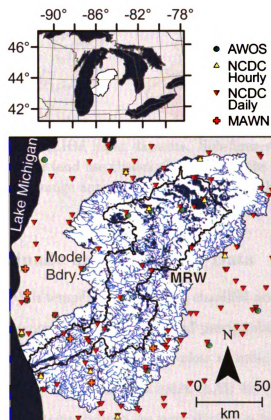


Figure 5.1: Map of the Muskegon River Watershed and surrounding model domain boundary, with an inset showing the location of the model domain within lower Michigan, USA. Filled symbols on the lower map are climate gage locations for data listed in Table 5.1.

The current land use within the MRW is predominantly forest (58%), with a significant fraction of agriculture (17%), some shrub or open lands (9.5%) and urban

(7.4%), and a significant portion of open water or wetlands (8.4%) (Torbick *et al.* (2006), see Figure 5.2). Soil textures from SSURGO data (USDA-NRCS, 2008) are predominantly sands (61%) and loamy sands (19%), with considerable areas of sandy loams (7.2%), loams (6.2%) and finer textures (6.8%). Outside the MRW, soil textures are somewhat finer, particularly in the southern model domain (see Figure 5.2).

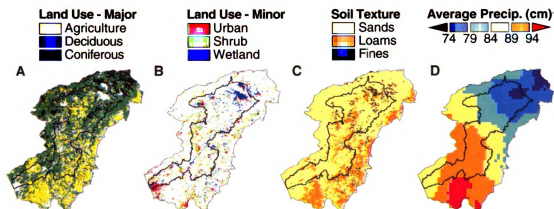


Figure 5.2: Maps of select ILHM input datasets. Sub-figures A and B show distributions of major and minor land use classes, C illustrates the dominant soil textural classes, and D plots the average annual precipitation across the model domain.

5.2.1 Input Climate and Landscape Data

The climate of the MRW is temperate and humid, classified according to the Köppen-Geiger scheme as Dfa (snow climate, fully humid precipitation, and warm summer temperatures) (Kottek *et al.*, 2006). Figure 5.3 plots monthly average temperatures, precipitation, solar radiation, and leaf area index (LAI) for the 1980 - 2007 simulation period. Average daily temperatures are typically below freezing from December-March, which typically results in a persistent seasonal snowpack that stores approximately 10 cm of water. That snowpack usually melts between late March and mid April, when daily maximum temperatures rise above freezing (Figure 5.3A).

Winter is typically the driest season, and summer the wettest, with significant potential for very wet months during the spring and autumn (see Figure 5.3B). Frontal

systems moving over Lake Michigan generate lake effect precipitation, particularly in late fall and early winter. Lake effect enhances average annual precipitation in a band near the lake by 10-20 cm annually (see Figure 5.2).

Monthly average shortwave solar radiation varies by as much as 20% from year to year during the spring and fall, with a smaller annual range during the winter months (Figure 5.3C). Watershed-average leaf area index values are greatest in forested areas (peak growing season ~ 5), and are lowest in agricultural areas (peak growing season ~ 2 , and vary typically less than 10% year-to-year (Figure 5.3D).

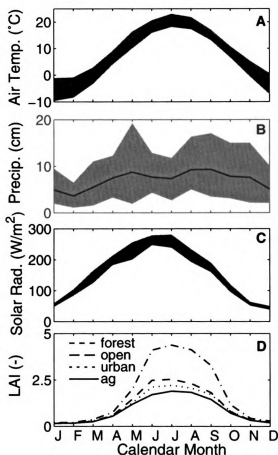


Figure 5.3: Monthly values of select climate and leaf area index (LAI) inputs. The heavy solid line is the median monthly value during the simulation period, which is surrounded by a shaded range indicating the 5% and 95% values. Plotted in A) monthly temperatures, B) monthly precipitation, C) monthly average solar radiation, and D) monthly average LAI for four land use classes.

St

2000

dat

we

Th

W

we

we

mo

T

S

1

4

1

5

11

1

1

1

1

1

1

1

1

Significant effort was invested to create a composite climate dataset from heterogeneous sources while maintaining spatial and temporal consistency (see Table 5.1 for data sources). Quality assurance, gap filling, and data consistency procedures that were used are detailed in Chapter 4 of this thesis. To spatially distribute the data, Theissen polygons were constructed for each unique combination of on-line gages. Weekly precipitation, aggregated from daily data at over 60 stations, was used to re-weight hourly precipitation values available at only 9 stations. Distributed datasets, including LAI and NEXRAD precipitation data, were bilinearly downsampled to the model resolution.

Table 5.1: Input climate and LAI data sources with the start and end dates of each source. Locations of gage data are shown in Figure 5.1.

<i>Source</i>	<i>Type</i>	<i>Start</i>	<i>End</i>
AWOS ¹	gaged climate	1980	2006
MAWN ²	gaged climate, solar radiation	1996	2007
NCDC ³		1980	2007
NEXRAD ⁴	radar precipitation	1996	2007
GIMMS ⁵	8-km NDVI	1981	2006
AVHRR ⁶	1-km NDVI	1989	2003
MODIS ⁷	LAI (v5)	2000	2007

¹(Andresen, 2007) ²(MAWN, 2008) ³(Center, 2001; NCDC Hourly, 2008)

⁴(Andresen, 2008) ⁵(Tucker et al., 2005) ⁶(Eidenshink, 1992) ⁷(Knyazikhin et al., 1999)

5.2.2 The Integrated Landscape Hydrology Model

ILHM is loosely-coupled set of fully-distributed physical process models, grouped into four domains: 1) surface water storage and routing, 2) canopy and root zone, 3) deep unsaturated zone, and 4) saturated zone groundwater. Its internal parameters are virtually all physically-measurable, and generally accessible through literature, requiring no site-specific calibration for reasonable flux estimates. This makes it well suited both for application in ungaged basins and for calibrated modeling within highly instrumented catchments.

1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
13.
14.
15.
16.
17.
18.
19.
20.
21.
22.
23.
24.
25.
26.
27.
28.
29.
30.
31.
32.
33.
34.
35.
36.
37.
38.
39.
40.
41.
42.
43.
44.
45.
46.
47.
48.
49.
50.
51.
52.
53.
54.
55.
56.
57.
58.
59.
60.
61.
62.
63.
64.
65.
66.
67.
68.
69.
70.
71.
72.
73.
74.
75.
76.
77.
78.
79.
80.
81.
82.
83.
84.
85.
86.
87.
88.
89.
90.
91.
92.
93.
94.
95.
96.
97.
98.
99.
100.

Individual domain modules within ILHM may have different spatial and temporal discretization schemes. Currently, all domains use structured grids. For this study, the surface routing, canopy and root zone, and deep unsaturated zone domains share the same hourly timesteps, with square grid cells 425 meters on a side. The saturated zone module is discretized at 1/4 the grid size of the surface domain, or approximately 106 meters, and is run with daily time steps and weekly input stress periods. The root zone is divided into 7 vertical layers, varying in thickness between 5 and 70 cm. Deep lakes have 10 vertical layers, between 30 cm and tens of meters thick. The deep unsaturated zone is a single layer that tracks multiple discrete wetting fronts. Three vertical layers are used for the MRW saturated groundwater model, ranging between 3 meters and several hundred meters thick. Neighboring cells do not necessarily have identical vertical layer structures.

Each cell in ILHM is composed of three “landunits”, uplands, wetlands, and streams. The fraction of each cell covered by the landunits are F_{up} , F_{wet} , and F_{str} . Within each landunit, hydrologic fluxes and storage volumes are independently maintained, and the total flux in a cell (for shared flux types, such as deep percolation) is calculated as $Q = Q_{up}F_{up} + Q_{wet}F_{wet} + Q_{str}F_{str}$. Only the root-zone domain module is subdivided in this fashion.

For this study, an 18,900 km² model domain of was established around the MRW. The boundaries of this domain extend outward to significant hydrologic divides, primarily the main channels of adjacent similar-order watersheds. The base of the shallow unconfined aquifer, a contact with bedrock, occurs at depths between 15 and 400 m. This contact was interpolated from thousands of oil and gas wells. Within the model, the root zone domain extends to 2.5 m depth or shallower if the water table (taken as the average annual low level from preliminary modeling) is closer to the surface. Below this, the deep unsaturated zone domain extends to the water table that defines the upper limit of the saturated groundwater domain.

5.2.3 Climate Teleconnection Indices

Numerous studies have identified correlations between low-frequency climate oscillation, or teleconnection, patterns and precipitation or temperature in the Great Lakes region (*Rodionov and Assel, 2000; Rajagopalan et al., 2000; Assel, 1992; Leathers et al., 1991*). Some of the most influential such patterns in the Northern Hemisphere are the North Atlantic Oscillation (*NAO, Visbeck et al. (2001); Walker and Bliss (1932)*), Pacific/North American (*PNA, Wallace and Gutzler (1981)*), Arctic Oscillation (*AO, Thompson and Wallace (1998)*), and El Nino Southern Oscillation (*ENSO, Trenberth (1997)*).

These patterns are summarized numerically via indices measuring the departure from normal of geopotential heights within certain geographic zones (*NAO, AO, or PNA*), or sea surface temperatures (*ENSO*). These index values are typically normalized such that monthly average values are small, typically within ± 2 for indices used here. For this study, monthly index values were obtained from the NOAA Climate Prediction Center. Note, the *ENSO* index was calculated from SST anomalies in the so-called 3.4 region, *NAO* and *PNA* using 50 hPa height anomalies, and *AO* using 100 hPa anomalies.

To understand the influence of the climate indices on fluxes in this region, correlations between monthly, seasonal, and annual index values and precipitation and temperature were evaluated. The observed correlations are discussed below, all of which are consistent with other studies. This study examines correlations between those indices and the resultant landscape hydrologic fluxes, particularly evapotranspiration and groundwater recharge. Understanding these linkages will help to further our understanding of the processes that influence the variability of terrestrial hydrologic fluxes.

5.2.4 Linear Regression Modeling

Multiple linear regression modeling is applied to extract additional information from the complex temporal and spatial variability present in the ILHM outputs. Here, annual cell-by-cell input variables are correlated to flux estimates from ILHM to produce models capable of estimating the spatial distribution of those annual fluxes. A nested linear modeling approach was used along with variables including precipitation and temperature during both the growing and non-growing seasons, field capacity of the soils, fraction of the landunit covered by each land use class, and the annual average value of the *PNA* and *NAO* indices. Linear model inputs are annual or seasonal cell-by-cell averages, except for annual mean climate index values that are spatially uniform. Due to memory limitations in the 32-bit version of the statistical package R (*R Development Core Team*, 2005) that we used, 30,000 randomly selected ILHM cells were used for the regression analyses.

To determine the quantitative improvement of each successive variable addition, the root-mean square residual and coefficient of determination, R^2 , between the annual linear model and the ILHM estimates were calculated. To compare the relative influence of each variable on the linear model, the model coefficient was multiplied by the range of values expressed by each variable. In this fashion, the maximum and minimum contributions to the linear model for each variable can be directly assessed.

5.3 Model Results and Discussion

Cell-by-cell monthly fluxes of water budget components were saved from ILHM for the 1980 - 2007 model period. The discussion below examines the spatial and temporal variability of those fluxes, to better understand the expected influences of climate, soil texture, and land use on terrestrial hydrology. In addition to direct spatial and temporal aggregation of the model results, multiple regression analysis was used to

extract information from the complex ILHM outputs.

Model parameters, as discussed in Chapter 4, were largely derived from literature values. With the exception of preliminary calibration of hydraulic conductivity values with a steady state groundwater flow simulation, the model parameters have not been adjusted to better match measured values. This uncalibrated simulation displayed no significant net basin-wide bias in stream discharge, although individual catchments had annual errors ranging from -4.3% to 13.6% of precipitation over the 28 year model period.

A wide range of model parameters could be estimated to reduce this bias and minimize the residuals between simulated and observed values, however here we choose to present uncalibrated results for several reasons. Residuals from a process-based model inform the modeler of improvements that can be made to the sources or preparation of input data, process modules, and physical parameterization of a system. This study seeks to better understand the interactions among hydrologic processes and climate and landscape inputs. Calibration in place of properly simulating physical processes can mis-represent the role of individual processes within the overall system. Finally, by presenting uncalibrated results and explicitly acknowledging the model bias, the reader can make a more honest assessment of ILHM's capabilities to predict flows in ungaged basins as well as for application to other problems.

The results below focus primarily on the near-surface processes that drive spatial and temporal variability of hydrologic fluxes in the MRW. Because of this deep percolation (DP), or the flux of water below the root zone, is presented rather than groundwater recharge to the water table. In the coarse-textured, high-conductivity soils of the model domain, the near-surface water budget is dominated by DP and evapotranspiration (ET). Note also that all annual fluxes are calculated for 26 complete water years, from October - September (e.g., Oct 1980 - Sep 1981 is the 1981 water year).

5.3.1 Spatial and Temporal Variability of DP and ET

Figure 5.4 shows the average cell-by-cell annual ET and DP along with the coefficient of variation of the annual fluxes. The coefficient of variation c_v is a measure of the inherent variability of a dataset, and is calculated as $c_v = \sigma_x / \bar{x}$, where σ_x is the standard deviation of the dataset, and $\bar{x}$ its mean. Note that, due to lateral subsurface flows, the sum of ET and DP may exceed precipitation.

ET (Figure 5.4A) displays no significant regional trends, with the exception of high values in the upper part of the MRW due to a concentration of wetlands. Large lakes and concentrated urban areas have lower ET, as does the broad floodplain of the Muskegon River main channel. Areas of higher ET at the 10-50 km scale (see Figure 5.1 for a scale) tend to have finer-textured soils (see Figure 5.2). Patterns in mean ET due to land use are more subtle, and relative relationships between classes depend on underlying soil textures.

By contrast, deep percolation (Figure 5.4B) has a strong regional trend. From the southwest - northeast, DP decreasing by 50% within cells of similar land use and soil type. The underlying factors driving this regional trend are discussed in greater depth below, but the broad spatial pattern of DP is indicative of lake-effect precipitation during the fall and winter. ET responds much more directly to the more spatially-uniform growing season precipitation. Certain wetland classes also produce significant DP, including the complex in the northeast corner of the watershed. Coarse-textured soils and sediments, and concentrated urban areas also lead to higher DP.

In general, annual DP is more variable than ET (Figure 5.4C&D); model average c_v of DP is 32.4%, and ET is 14.0%. The DP c_v map exhibits significantly more variability at the 5-10 km scale than ET as well. ET c_v exhibits some geometric patterns that are artifacts of the nearest-neighbor spatial interpolation of solar radiation and temperature.

Regions with high c_v of DP tend to have low c_v of ET. In uplands, deep percolation

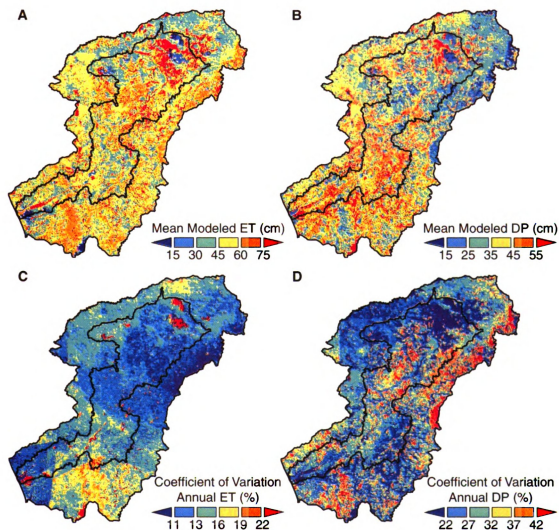


Figure 5.4: Maps of total cell mean annual modeled A) evapotranspiration (ET), B) mean annual modeled deep percolation (DP), C) coefficient of variation (ratio of standard deviation to the mean) of annual ET, D) coefficient of variation, c_v of annual DP. Note the color scales differ for each sub-figure.

occurs only when root zone soil moisture exceeds field capacity; whether and to what degree this occurs is controlled largely by spring and autumn rains, which vary significantly inter-annually. Much of ET is determined by the storage of moisture in the soil root zone, nominally the difference between the field capacity and wilting point of a soil. Most years, soils are near or at field capacity at the start of the growing season, leaving summer precipitation as the remaining factor governing inter-annual variability for ET. Regions with low c_v of ET tend to have finer textured soils, and thus greater soil moisture storage, reducing dependence on summer precipitation. These same regions also require more soil moisture to exceed field capacity, decreasing the likelihood of significant DP occurring during any given year.

To further examine the temporal variability of ET and DP fluxes across the MRW, Figure 5.5 plots annual watershed-averaged fluxes. In Figure 5.5A, total watershed ET, storage and runoff are plotted. ET dominates the watershed water balance, which is typically 60-70% of precipitation (Figure 5.5C). The annual watershed runoff time series is quite steady, compared to the variability in other water budget components. Steady annual watershed runoff is facilitated by storage of water within the watershed, primarily in the saturated groundwater system. Storage is the most variable component of the MRW water budget, and is capable of storing and releasing as much as 25% of annual precipitation in any given year.

Discussing longer-term trends within a time span of just a few decades can be difficult, particularly when the model input data sources change during the simulation period (see Chapter 4 of this thesis). Nonetheless, simulated watershed ET seems to be declining steadily, while runoff is increasing. Some of this apparent trend may be driven by the years 2004 - 2007, which are the four highest runoff years in the simulation period. As a fraction of the total water budget, runoff seems to be increasing as well.

The water budget of upland areas is plotted in Figure 5.5B & D. The only two

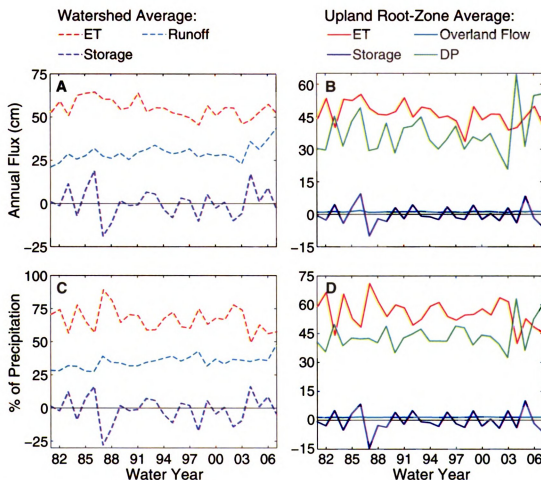


Figure 5.5: Plots of annual watershed-averaged fluxes for the MRW. On the left, watershed-average ET, storage, and runoff for the MRW are plotted, while on the right upland ET, overland flow, root-zone storage, and DP within upland landunits are shown. Parts A & B plot annual fluxes in cm. Those quantities, divided by annual precipitation, are plotted in parts C & D. Note, for storage, negative values indicate release of water from storages in the model.

dominant components at the watershed scale are ET and DP, runoff is a small fraction except in urban land uses. Storage within the root zone is an important component of the annual variability in fluxes, but is minimal long-term. Within uplands, ET remains the largest component of the water budget, though deep percolation exceeds ET in some years. Over the entire model period, ET makes up 53% of the water budget, and DP 43%. The 2004, 2006, and 2007 water years experienced the greatest DP during the 28 year model period, as a result of heavy spring and fall rains. In absolute terms, DP is annually more variable than ET, though the opposite is true as a fraction of precipitation.

Strong 2-3 year cycles are evident in both ET and DP that may be indicative of the influence of broader climatic patterns. Table 5.2 summarizes a nested linear regression model of annual watershed-averaged upland ET and DP with annual means of the four climate teleconnection indices *PNA*, *NAO*, *ENSO*, and *AO*. Four models were tested, each adding one additional index. The ordering was chosen based on single regressions for each dataset, with highly correlated indices included earlier in the nested linear models. The final five years (water years 2003-2007) were excluded from this analysis because of behavior that is significantly different from the remainder of the model period.

A model including just *NAO* + *PNA* explains between 15% and 24% of the variability in average annual upland DP and ET, respectively. The addition of *ENSO* has little effect on either model, while adding *AO* seems to significantly increase overall model fit. The significance of this, however, is suspect as the total number of data points (22) may be over-parameterized by a four-component linear model. Note, since the range of index values is approximately equal for all indexes, the values of coefficients from the models indicate the relative influence of each parameter.

The positive phase of *NAO* (and the closely-related *AO*) brings warmer than average annual temperatures to the region, along with slightly increased average precipi-

tation. In addition to the continental effects induced by modifications to mean storm tracks, discussed in *Hurrell et al.* (2003), increased average temperatures may promote decreased ice cover and increased lake effect precipitation. According to these ILHM simulations, both annual upland DP and ET tend to be higher when *NAO* is in its positive phase.

A positive phase of *PNA* corresponds to reduced upland ET, but slightly increased upland DP. As observed in *Rodionov and Assel* (2000); *Leathers et al.* (1991), winter and spring temperatures tend to be cooler during a positive *PNA* cycle, while summer and fall precipitation is reduced. These lower winter temperatures prolong the seasonal snowpack, which leads to a more significant pulse of meltwater infiltration in the spring. Reduced summer and fall precipitation limits ET significantly during a period when surface moisture storage is typically at or near annual minimums.

Table 5.2: Tables of multiple regression models for annual watershed-average A) Upland ET, and B) Upland DP. Columns include the root-mean square residual between the multiple regression model and ILHM, along with the correlation coefficient R^2 , and the model parameters. Note, the final five years of the simulation were excluded from this analysis because DP seems to behave exceptionally during this period compared the rest of the simulation.

A (Upland ET)		R^2	Intercept	Coefficients			
Model	RMSR (cm)			NAO	PNA	ENSO	AO
NAO	4.44	0.20	46.3	6.88			
+ PNA	4.33	0.24	46.8	5.81	-2.93		
+ ENSO	4.31	0.25	46.7	5.68	-3.10	0.64	
+ AO	3.47	0.51	45.5	17.5	-4.87	1.07	-11.8

B (Upland DP)		R^2	Intercept	Coefficients			
Model	RMSR			NAO	PNA	ENSO	AO
NAO	5.20	0.08	33.9	4.79			
+ PNA	5.00	0.15	33.3	6.34	4.24		
+ ENSO	5.00	0.15	33.2	6.32	4.22	0.08	
+ AO	4.90	0.18	32.8	10.9	3.53	0.24	-4.52

There is disagreement over which of the *AO* and *NAO* patterns better represent fundamental atmospheric modes (*Ambaum et al.*, 2001; *Thompson and Wallace*, 1998). In any case, the two indices are highly cross-correlated in this study (Table 5.3) and

others (*Ambaum et al.*, 2001). *NAO* is virtually uncorrelated with either *PNA* or *ENSO*, whereas *AO* is somewhat correlated with both. This strong cross-correlation between *AO* and *NAO* partially explains the behavior of the fourth model for Upland ET in Table 5.2A. As expected, including both *NAO* and *AO* alters the model coefficients of *NAO* because of interaction between the two variables. Including *AO* also increases the model R^2 for upland ET dramatically, while having little effect for upland DP. Based on Tables 5.2 and 5.3, only *PNA* and *NAO* seem to offer unique correlative information about upland ET and DP.

Table 5.3: Cross-correlation values between monthly climate indices.

	<i>PNA</i>	<i>NAO</i>	<i>ENSO</i>	<i>AO</i>
<i>PNA</i>	1	-0.003	0.257	-0.264
<i>NAO</i>	-0.003	1	0.001	0.608
<i>ENSO</i>	0.257	0.001	1	-0.125
<i>AO</i>	-0.264	0.608	-0.125	1

Average monthly storage values provides an informative view of the interactions among model domain processes (Figure 5.6). From January through March, water is stored both in the snowpack and root-zone soil. At the onset of the growing season, a large pulse of deep percolation (Figure 5.6B) balances extraction of water by ET. As this pulse of deep percolation and, eventually, groundwater recharge drains from the system, water is removed from groundwater storage until August. At this point, streams are near annual low flows, and ET has depleted most of the soil moisture storage capacity. Fall precipitation at the end of the growing season then begins adding water to root-zone storage and deep percolation.

On a monthly basis, upland ET and DP vary inversely (Figure 5.6B). Most DP occurs following spring snowmelt in late March and early April, while a second, smaller peak occurs in October and November after leaf senescence. Deep percolation (and therefore groundwater recharge) is small during the winter months because the snowpack stores precipitation, and during the growing season when transpiration and soil

evaporation consume most of the available moisture.

Watershed runoff (Figure 5.6A) is the sum of groundwater discharge and overland flow runoff, minus evaporation from the drainage network. Direct overland flow runoff accounts for a small fraction of total runoff from the MRW. While still weighted towards higher runoff during the spring, storage in unsaturated and saturated zones dampen the sharp DP peaks significantly.

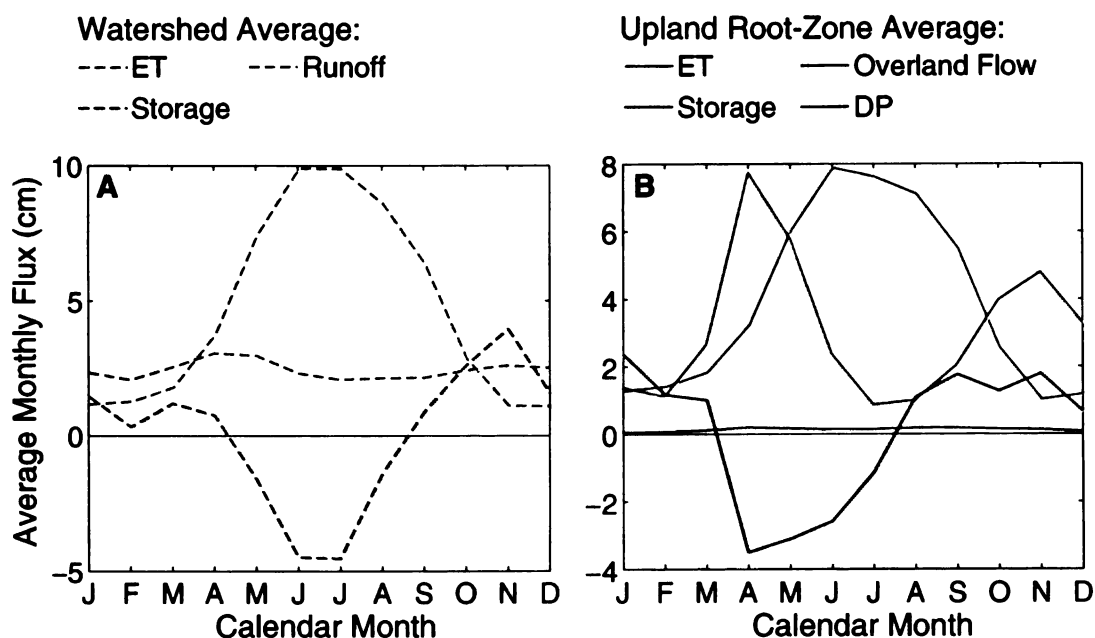


Figure 5.6: Average monthly fluxes of A) watershed ET, runoff (overland + groundwater discharge), and storage, and B) upland ET, overland flow, root-zone storage, and DP within upland landunits.

Evapotranspiration is the sum six transpiration and evaporation components, as shown in Figure 5.7, and Tables 5.4 & 5.5. Upland and wetland transpiration are the dominant components, followed by canopy evaporation, soil evaporation, snowpack sublimation, and finally open water evaporation. Canopy evaporation during the winter months is due in large part to evergreen coniferous interception of snowfall. Snowpack sublimation peaks in February because the cold, dry winter air has a high vapor pressure deficit. Open water evaporation during the late summer and autumn occurs primarily in large lakes that accumulated heat during the summer. This sim-

ulation may under-represent open water evaporation in these lakes, as modeled lake temperatures tend to be lower than observed.

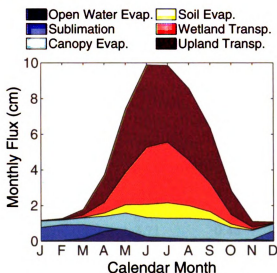


Figure 5.7: Filled area plot of watershed-average monthly evaporation and transpiration components for the MRW. Shown are upland transpiration, wetland transpiration, soil evaporation, canopy evaporation, snow and ice sublimation, and open water evaporation.

Listed in Tables 5.4 & 5.5 are watershed-average annual mean, standard deviation, and coefficient of variation, c_v , of hydrologic flux components. Each table consists of three groups of rows. The first is annual mean precipitation, the second are the components of the water budget for the watershed (Table 5.4) and for upland landunits (Table 5.5). The third group includes notable fluxes within ET and runoff components in the second group, ranked in order of mean annual flux. Standard deviation, here, is calculated as the standard deviation of annual watershed-average or landunit-average fluxes.

As in Figure 5.5, storage is the most variable component of the water budget apart from precipitation (Table 5.4), over the model period, total change in storage is minimal (0.3 cm). Most of this storage occurs within the saturated groundwater domain, and is controlled largely by the variability in annual deep percolation (Table

Table 5.4: Watershed-average annual mean, standard deviation, and coefficient of variation, c_v , of hydrologic fluxes.

<i>Flux Type</i>	<i>Mean (cm)</i>	<i>Std (cm)</i>	<i>c_v %</i>
Precip	84.6	12.4	14.6
Total ET	55.1	5.26	9.5
Total Runoff	29.3	4.64	15.9
Total Storage	0.26	8.60	3308
Baseflow	25.6	4.15	16.2
Upland T	23.2	2.36	10.2
Wetland T	14.1	1.66	11.8
Canopy E	8.67	1.48	17.1
Soil E	3.71	0.45	12.1
Snow and Ice Sub.	3.06	1.92	62.7
Wetland E	2.37	0.45	19.0
Wetland Run	2.73	0.57	20.9
Baseflow	0.96	0.18	18.8

Table 5.5: Watershed-averaged mean, standard deviation, and coefficient of variation, c_v , of upland water budget components.

<i>Flux Type</i>	<i>Mean (cm)</i>	<i>Std (cm)</i>	<i>c_v %</i>
Precip	84.6	12.4	14.6
Upland ET	46.3	5.10	11.0
Upland DP	37.6	9.95	26.5
Upland Run	1.15	0.21	18.6
Upland T	29.0	2.95	10.2
Canopy E	8.8	1.15	17.1
Soil E	4.71	0.57	12.0
Snow Sub	3.84	2.40	62.5

5.5). With the exception of snow and ice sublimation, the c_v values for all ET and runoff sub-components are between 10 and 20%.

5.3.2 Influence of Climate, Soils and Land use on Upland Hydrologic Fluxes

Lake Michigan plays a major role in the hydrology of the MRW and the surrounding region. To examine the effect of proximity to Lake Michigan on upland water budgets, the distance of each cell in the model to Lake Michigan along the prevailing wind direction was calculated. Wind rose diagrams from the MAWN gages (not shown) demonstrate that the prevailing winds are roughly from the SSW at an azimuth of 60°.

Figure 5.8 illustrates the lake effect on average annual precipitation, ET, and DP. Shown are monthly upland fluxes within 25 km distance bands downwind from the lake. Here, all cells containing greater than 75% deciduous forests, and sand as the dominant soil texture both horizontally and vertically are used.

There are two distinct lake effect precipitation modes (Figure 5.8A), warm-lake and cool-lake *Andresen and Winkler (2009)*. The warm-lake effect, when the lake is warmer than the air above, occurs most strongly in late September through early December. During these times, Lake Michigan significantly enhances precipitation near the lake shore, to approximately 150 km in the downwind direction. When the lake is cooler than the air above, which occurs from late May through early September, near-lake precipitation to 50 km is reduced as the cool water inhibits the formation of convective thunderstorm cells.

The cool-lake effect is visible in Upland ET (Figure 5.8B) during the summer months. ET is roughly uniform from September - April. Then, the combined effects of decreased precipitation, and slightly cooler temperatures near the lake reduce ET by roughly 10% within 50 km of the lake shore.

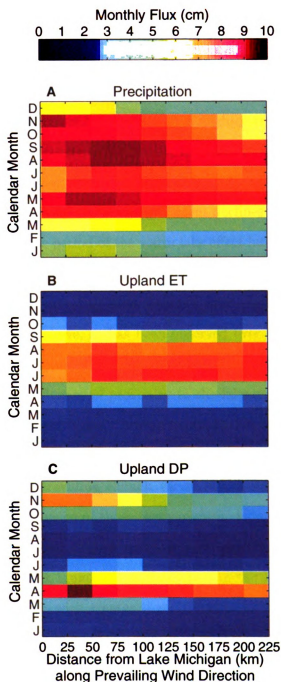


Figure 5.8: Filled value plots of average monthly A) precipitation, B) upland evapotranspiration (ET), and C) upland deep percolation (DP) vs. distance from Lake Michigan along the prevailing wind direction. All values are averages of only cells within the model domain containing greater than 75% deciduous forests, with sand as the dominant texture. Values are averaged over the entire model period.

Upland DP exhibits both cool-lake and warm-lake effects (Figure 5.8C). From October - December, and again in March - May, deep percolation is strongly biased toward the lake shore. Near-lake DP may be 50% greater than inland values during those months because of the warm-lake effect. January and February do not exhibit strong trends. From June - September, near-lake DP is slightly reduced by the cool-lake effect, though not as significantly as ET during the same period.

Although not specifically examined here, the role that Lake Michigan plays in the water budget of the MRW may be changing. Recent declines in lake ice cover (*Assel et al.*, 2003) may extend the period of the warm lake effect later into the winter. This may, in turn, increase winter precipitation and lead to greater DP (*Burnett et al.*, 2003). The cool-lake effect may decline as summer water surface temperatures warm more rapidly than the air above, as has been observed over the last few decades at Lake Michigan tide gages (*Austin and Colman*, 2007).

Because of the strong influence of proximity to the lake on upland fluxes, Figures 5.9 & 5.10 are calculated within the distance bands 75-150 km from the lake shore. That band was chosen to provide relatively uniform ET and DP conditions according to Figure 5.8, while still providing enough variability in soil texture and land use classes to examine the influence of those two variables. However, most of the sandy loam and loam soils are in agricultural areas as these soils are more productive.

Figure 5.9 plots upland ET and DP as a function of the majority horizontal and vertical soil texture in cells with greater than 75% agricultural land use within the 75-150 km downwind distance band. To reduce the impact of vertical discontinuities in soil texture, cells with mismatched majority horizontal and vertical textures were not included. Finer-textured soils influence hydrologic fluxes in two ways: 1) reduced hydraulic conductivity slows the loss of water out of the root zone as deep percolation, and 2) higher field capacity stores more water after the spring. This increased storage capacity is clearly evident in Figure 5.9A. Finer-textured soils evaporate and transpire

much more water than coarser sands. Note that the fines class in Figure 5.9A includes soils with a wide range of textures including silty loams, muck, and sandy-clay loams, so the behavior is less uniform than the other three texture classes.

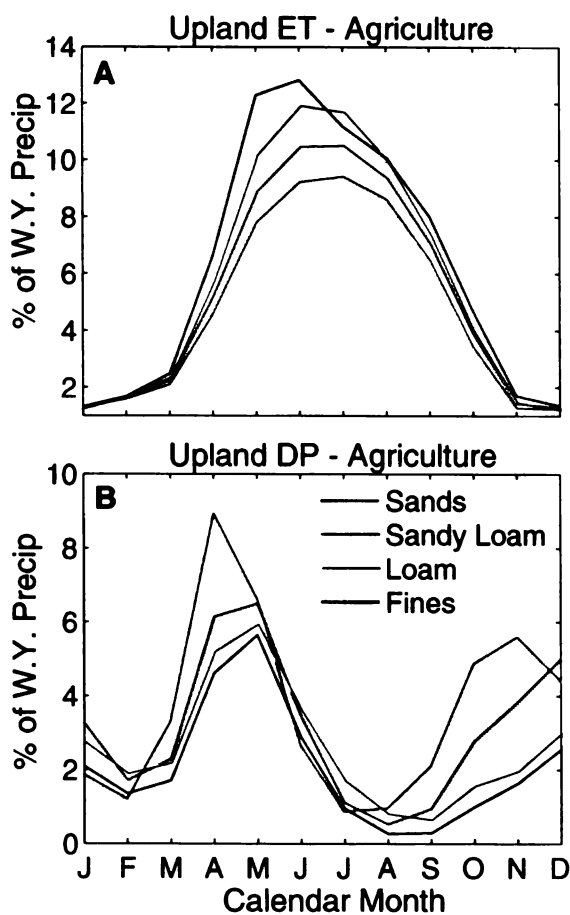


Figure 5.9: Plots of monthly average upland ET and DP across generalized soil texture classes as a percent of water-year (Oct - Sep) precipitation. Values are from cells within the model domain containing greater than 75% agricultural land use between 75-150 km from the lake shore along the prevailing wind direction.

Soil texture influences both the quantity and timing of deep percolation, as shown in Figure 5.9B. Less water is lost during spring snowmelt in finer textured soils. For textures except the fines class, DP during the early growing season is somewhat greater as silt and clay content increases because the reduced hydraulic conductivity in these soils slows transport of moisture through the root zone profile. As the growing season progresses, greater utilization of water by plants again reduces DP in finer

textures relative to sands. Regardless of soil texture, by the end of the summer soil moisture is near the wilting point most years (not shown here). At the end of the growing season, a greater amount of water is required to restore fine-textured soils to their field capacity. This has the effect of dampening the autumn DP peak seen in the sandy soils.

Table 5.6 presents the annual average upland ET and DP fluxes plotted in Figure 5.9. Compared to sandy-textured soils, finer soil textures can increase upland ET by as much as 30%, and decrease upland DP by almost 40%. Note here than the fines texture class has the fewest members, though the differences shown are still statistically significant. This very strong dependence on soil texture highlights the importance of using higher-resolution (preferably vertically-variable) soil texture information available from the digitized SSURGO data. Many studies rely on the older, coarser, but more manageable STATSGO data (*USDA-SCS*, 1993; *Wang and Melesse*, 2006) that lacks much of the detailed spatial variability present in SSURGO.

Table 5.6: Average annual upland ET and DP by soil texture in agricultural areas between 75-150 km from the lake shore along the prevailing wind direction.

<i>Texture</i>	<i>Upland ET (cm)</i>	<i>Upland DP (cm)</i>
Sands	47.6	37.3
Sandy Loam	52.8	32.4
Loam	58.5	27.6
Fines	62.0	21.9
<i>range</i>	14.4	15.4

To examine the influence of land use on the terrestrial water budget, upland ET and DP are plotted in Figure 5.10 as a function of the dominant land use class within sandy soils (sub-figures A and B) and loam or sandy loam soils (sub-figures C and D). For each cell to be included in this analysis, at least 75% of it had to be covered by a single land use class. Sandy soil textures dominate the model domain, and tend to be more vertically homogeneous than loam soils, so the interrelationships among land use classes are perhaps more robust for the sands. As in Figure 5.9, only model

cells within the 75-150 km distance band are included.

Land use affects the ILHM-simulated fluxes through a variety of mechanisms, most prominently through canopy interception of moisture and radiation, and through transpiration differences. Note, there are several factors not yet included in ILHM that may influence the land use relationships shown here, including sub-canopy microclimate effects, soil temperature variation, ground litter, and differences among land uses in plant soil water extraction efficiency. Furthermore, the influence that a particular land use class has on terrestrial hydrology may vary depending on the texture of soil beneath it (compare Figures 5.10 A & B to C & D).

Within both soil textures, coniferous forests tend to evapotranspire more than other land use classes, with the exception of sandy agriculture. This is due primarily to interception of precipitation during periods when plants in other land covers are dormant and their leaves senescent. Urban areas, in both sands and loams had the least ET and DP, due to several factors. ET in urban areas may be under-predicted due to the difficulty of retrieving urban LAI from remote sensing platforms. Even so, LAI in urban areas tends to be lower than shrub or forested land uses, as space is appropriated for structures and transportation. Furthermore, runoff in urban areas is significant due to impermeable surfaces, which for this simulation were assumed to cover 20% the urban fraction of each cell.

Next to urban, shrub land uses have the lowest ET. This is perhaps because the shrub land use class includes rangeland and other managed vegetation that maintain persistently low LAI throughout the year (see Figure 5.3), while the resistance to transpiration in woody shrubs in particular is much higher than in agriculture crops. Low runoff in shrub land uses leads to higher DP than other classes. Agriculture and deciduous forests, the two dominant land use classes in the watershed, tend to fall in between other classes, with the exception of DP in sandy-soil agriculture.

Though a less important factor than soil texture, understanding the impacts of

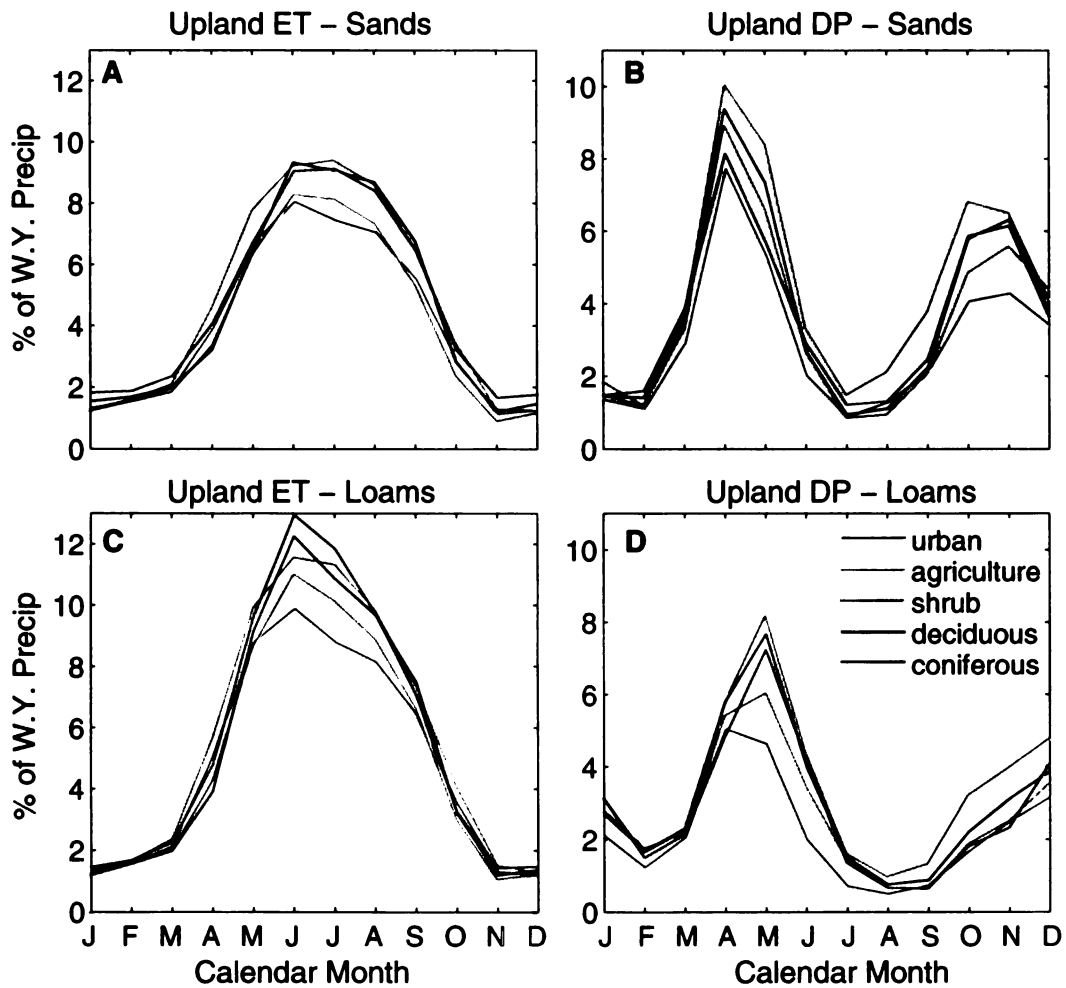


Figure 5.10: Plots of monthly average upland ET and DP across land use classes as a percent of water-year precipitation. Values are from cells within the model domain located between 75-150 km from the lake shore along the prevailing wind direction with either sands (A & B) or loams (C & D) as the dominant soil texture class both vertically and horizontally.

land use on hydrologic fluxes is critical because relative changes in fluxes are more important both ecologically and for water resources than the absolute values of fluxes. Listed in Table 5.7 are annual average upland ET and DP fluxes across land use classes within sandy soils in the 75-150 km down wind distance band. At the extremes, land use in sand soils produces a range of 7.8 cm for upland ET and 14.0 cm for upland DP. From Table 5.6, soil textures induced a range within agricultural areas of 14.4 cm for both upland ET and DP. For the most part, the influence of soils is important only spatially, while land use has both spatial and temporal implications.

Table 5.7: Average annual upland ET and DP by land use in sandy soils 75-150 km from the lake shore along the prevailing wind direction.

<i>Land Use</i>	<i>Upland ET (cm)</i>	<i>Upland DP (cm)</i>
urban	40.3	31.2
agriculture	47.6	37.3
shrub	39.8	45.2
deciduous	44.7	39.2
coniferous	45.8	36.6
<i>range</i>	7.8	14.0

5.3.3 Linear Regression Analysis of ILHM Outputs

Traditionally, one of the few tools available for mapping the spatial and temporal variability of groundwater recharge and ET over regional domains has been linear regression modeling. The only available estimates of spatially-variable groundwater recharge in the MRW were generated by correlating baseflow-separated recharge estimates from stream gages with land cover, geographic, topographic, and climatic variables within catchments (*MDEQ*, 2009; *Holtzschlag*, 1996).

Two problems with these estimates are that 1) they rely on baseflow separation to determine groundwater recharge, a method that can be problematic in catchments with significant wetland area, and 2) variables that co-vary at scales smaller than the resolution of catchment-based methods will inevitably confuse the linear models. An example of this second case is agricultural lands and soil textures. Within the MRW

and surrounding model domain, agricultural land use is highly positively correlated with field capacity of the soil (correlation coefficient of 0.46).

Linear modeling of ILHM-calculated ET and DP suffers neither of those two problems. ILHM calculates ET and groundwater recharge (or DP) directly, and at the scale of the ILHM model cells in this study, roughly 425 meters (cell area of 18 hectares), there are sufficient instances of all variables in all combinations within cells that the co-variation problem is greatly reduced. Therefore, linear regression models derived from the ILHM outputs can potentially provide broadly- and readily-applicable annual estimates of ET and DP. Furthermore, these models provide additional information about the relative influence of each variable without needing to painstakingly isolate each variable as in the previous section. We refer to this analysis as process-inferred statistical modeling.

Here five nested linear models are calculated separately for upland ET and DP. The first includes only seasonal precipitation, the second adds seasonal temperatures. Cell-average soil field capacity values are added next, followed by within-landunit land use fractions, and finally climate index values. A sixth model, labeled “P” to indicate “parsimonious” has been calculated omitting seasonal temperatures and climate cycles. Note that the climate variables vary both in space and time (except the indices, which are annually single-valued), while the landscape parameters only vary spatially.

As shown in Tables 5.8 and 5.9, each additional set of variables offered an improvement in model fit, ultimately resulting in the linear models explaining 44 and 62% of the annual variability in upland ET and DP, respectively. The root-mean square residuals for DP are somewhat higher, due to the inherently more variable nature of this quantity. The influence of each variable relative to the mean field (the model intercept) is given in part C of Tables 5.8 and 5.9, and is calculated as the product of the range of the parameter and its model coefficient. The most influential

variable for DP by a significant margin is non-growing season precipitation, followed by growing season precipitation, field capacity of soils, coniferous forests, and urban land use. The least influential are the climate indices *NAO* and *PNA*, and air temperatures. Upland ET is most influenced by soil field capacity, followed by growing season precipitation, forest, and agricultural land use. Here, non-growing season air temperature, urban land use, and the two climate indices were the least influential.

The primary limitations of a linear model is that the system may respond non-linearly to input variables, and threshold responses are difficult to describe. Such responses are evident in the ILHM outputs, particularly for soil texture and precipitation. Additionally, the variables may interact with each other, which a linear model cannot accurately represent. Ideally, the residuals between the linear model and ILHM would be randomly distributed. Non-linearities and interactions among parameters will produce non-random distributions.

Figure 5.11 maps water-year average upland ET and DP predicted by the fifth linear model, along with the cell-by-cell root-mean square annual residuals. Parts A and B capture most of the spatial patterns evident in the upland ET and DP maps (Figure 5.4 is somewhat comparable, but includes wetland ET and DP components as well). The residual maps (Figure 5.11C&D) display a combination of both random and non-random residuals. The non-random portion of the residuals are associated with both finer soil textures and agricultural land use, suggesting that these variables either interact, or behave non-linearly, or both (likely). There is also a region of high residual in the southern portion of the model domain similar to the high c_v region present in Figure 5.4C that may be an artifact of a single outlier gage or climate station. Overall, however, the linear model does a reasonable job of reproducing the spatial and temporal variability in the ILHM outputs. The mean RMSR of annual upland ET and DP are 7.7 cm and 8.8 cm respectively.

Table 5.8: Nested linear multiple-regression models for upland ET. A) displays the model intercepts, coefficients, correlation coefficients, and root-mean square residuals (*RMSR*) between the ILHM results and the linear models, listed in B) are the ranges of values for each linear model parameter, C) shows the influence of each parameter relative to the mean field (the intercept value, abbreviated *Intcpt.*), calculated as the product of the range of each parameter and the model coefficients from part A. *P_{GS}*, *P_{NGS}*, *T_{GS}* and *T_{NGS}* are growing season and non-growing season precipitation and temperature, respectively. *FC* is soil field capacity, and all land use classes are abbreviated.

A Multiple-Regression Models														
Model	RMSR (cm)	R ²	Intcpt.	P _{GS}	P _{NGS}	T _{GS}	T _{NGS}	FC	Urb.	Ag.	Shr.	Dec.	Con.	NAO PNA
1	10.91	0.11	27.61	0.37	0.09									
2	10.82	0.13	-5.08	0.39	0.10	1.84	-1.23							
3	9.46	0.33	-10.32	0.37	0.08	1.48	-1.21	0.95						
4	8.81	0.42	-29.76	0.37	0.09	1.60	-1.09	1.07	4.62	16.32	9.44	17.65	20.34	
5	8.63	0.44	-34.91	0.31	0.10	2.03	-1.02	1.07	4.28	16.16	9.49	17.70	20.51	1.96 -3.47
P	8.90	0.41	-1.37	0.35	0.08			1.07	4.61	16.40	9.51	17.62	20.26	

B Parameter range														
Parameter	P _{GS}	P _{NGS}	T _{GS}	T _{NGS}	FC	Urb.	Ag.	Shr.	Dec.	Con.	NAO	PNA		
Range	68.70	81.61	4.56	6.66	27.00	1.00	1.00	1.00	1.00	1.00	1.41	1.84		
Units	cm	cm	°C	°C	cm/m	-	-	-	-	-	-	-		

C Parameter influence (cm)														
Model	Intcpt.	P _{GS}	P _{NGS}	T _{GS}	T _{NGS}	FC	Urb.	Ag.	Shr.	Dec.	Con.	NAO	PNA	
1	27.61	25.73	7.19											
2	-5.08	27.10	8.43	8.40	-8.22									
3	-10.32	25.32	6.90	6.75	-8.08	25.73								
4	-29.76	25.53	7.60	7.30	-7.26	28.88	4.62	16.32	9.44	17.65	20.34			
5	-34.91	21.55	7.96	9.25	-6.77	29.00	4.28	16.16	9.49	17.70	20.51	2.75	-6.38	
P	-1.37	24.30	6.41			29.02	4.61	16.40	9.51	17.62	20.26			
Rank		2	8	7	9	1	11	5	6	4	3	12	10	

Table 5.9: Nested linear multiple-regression models for upland DP. A) displays the model intercepts, coefficients, correlation coefficients, and root-mean square residuals between the ILHM results and the linear models, listed in B) are the ranges of values for each linear model parameter, C) shows the absolute value of the relative influence of each parameter, calculated as the product of the range of each parameter and the model coefficients from part A. Column headers as in Table 5.8.

A Multiple-Regression Models															
Model	RMSR (cm)	R ²	Intcpt.	P _{GS}	P _{NGS}	T _{GS}	T _{NGS}	FC	Urb.	Ag.	Shr.	Dec.	Con.	NAO	PNA
1	11.37	0.45	-16.85	0.43	0.84										
2	11.31	0.46	2.74	0.41	0.81	-1.08	1.22								
3	9.97	0.58	8.06	0.44	0.83	-0.71	1.20	-0.97							
4	9.54	0.62	27.04	0.44	0.83	-0.77	1.19	-1.12	-20.06	-15.90	-8.54	-17.30	-20.29		
5	9.49	0.62	29.85	0.47	0.83	-1.01	1.17	-1.12	-19.90	-15.83	-8.59	-17.33	-20.37	-1.68	1.75
P	9.62	0.61	13.06	0.45	0.86			-1.13	-19.48	-15.54	-8.51	-17.31	-20.55		

B Parameter range												
Parameter	P _{GS}	P _{NGS}	T _{GS}	T _{NGS}	FC	Urb.	Ag.	Shr.	Dec.	Con.	NAO	PNA
Range	68.70	81.61	4.56	6.66	27.00	1.00	1.00	1.00	1.00	1.00	1.41	1.84
Units	cm	cm	°C	°C	cm/m	-	-	-	-	-	-	-

C Parameter influence (cm)													
Model	Intcpt.	P _{GS}	P _{NGS}	T _{GS}	T _{NGS}	FC	Urb.	Ag.	Shr.	Dec.	Con.	PNA	
1	-16.85	29.39	68.48										
2	2.74	28.39	66.34	-4.93	8.11								
3	8.06	30.19	67.89	-3.26	7.97	-26.10							
4	27.04	29.90	67.67	-3.50	7.94	-30.25	-20.06	-15.90	-8.54	-17.30	-20.29		
5	29.85	32.40	67.55	-4.59	7.79	-30.31	-19.90	-15.83	-8.59	-17.33	-20.37	-2.36	
P	13.06	30.69	69.95			-30.52	-19.48	-15.54	-8.51	-17.31	-20.55	3.22	
Rank		2	1	10	9	3	5	7	8	6	4	12	11

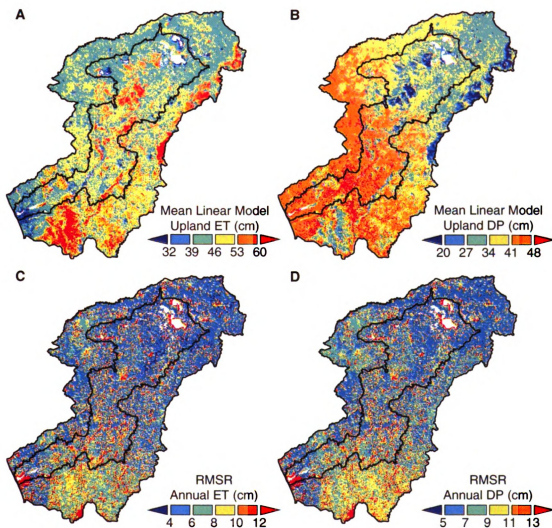


Figure 5.11: Maps of linear regression modeled annual average A) Upland ET and B) Upland deep percolation (DP) for 1980 - 2007. Parts C) and D) show mean root mean square residuals (RMSR) between the ILHM annual fluxes and the linear regression modeled values of ET and DP respectively.

5.4 Conclusions

This study primarily examines the spatio-temporal variability of groundwater recharge (or deep percolation) and watershed evapotranspiration which dominate the water budget of uplands in regions with high-conductivity soils. The region surrounding the Muskegon River Watershed was simulated using the Integrated Landscape Hydrology Model (ILHM), a novel hydrologic modeling suite. Twenty-eight years of hourly hydrologic fluxes were simulated at 425 m resolution over a 19,000 km² area.

In the MRW, groundwater recharge is more variable than ET both spatially and temporally. Over the period 1980 - 2007, cell-by-cell recharge (or DP) varied approximately 32% annually, with ET varying somewhat less at 14%. In the sandy soils of the MRW, virtually all precipitation in excess of soil field capacity becomes groundwater recharge. In contrast, ET variability is buffered by root-zone storage of moisture to capacity following spring snowmelt, and the subsequent summer rains. Regionally, recharge exhibited a strong decreasing trend away from the lake shore.

Two climate teleconnection indices, the Pacific/North American (*PNA*) and the North Atlantic Oscillation (*NAO*) were observed to correlate with both ET and recharge. In the MRW, a positive *PNA* phase indicates cooler winter and spring temperatures, and reduced summer and fall precipitation. This leads to increased annual groundwater recharge in a positive *PNA* phase as the snowpack persists longer in the season, and decreased annual ET due to late-summer moisture deficits. Positive *NAO* index values are associated with warmer average annual temperatures, and slightly increased annual precipitation. As a result, both annual ET and recharge correlate positively with *NAO*.

The proximity of Lake Michigan is critical to the hydrologic behavior of the MRW. Because of lake effect precipitation during the fall and winter, groundwater recharge near the lake shore is nearly 50% greater than inland values. If the lake effect is considered, soil texture plays the next most significant role in controlling the spatial

variability of recharge and ET. ET and DP are oppositely affected by finer-textured soils; ET increases by up to 30% and recharge decreases by as much as 39% relative to sands. Land use plays an important role as well, particularly because unlike soil texture, land use changes over decadal timescales. Variability between land use classes in sandy soils is almost 20% of upland ET and 45% of recharge.

Spatially-variable linear models of annual recharge and ET were calculated from the ILHM outputs. These process-inferred statistical models incorporate seasonal precipitation and temperature, soil field capacity, land cover, and climate index values. They explained approximately 40% and 60% of the spatial and temporal variability of ET and recharge, respectively. Critically, the resolution and domain size of this simulation allowed for strongly co-variate parameters like soil texture and agricultural land cover to be resolved independently. The three most important factors in the linear models were soil field capacity, seasonal precipitation, and coniferous land cover.

Recharge and ET exhibited considerable variability at all scales, bounded only by the extent of the model domain and the resolution of the grid cells. The presence of significant spatial and temporal variability even at the smallest model scales highlights the importance of explicitly modeling the fully-distributed, non-linear processes that drive landscape hydrologic fluxes. Future efforts will explicitly examine the impact of model resolution (both through parameterization and numerical solutions) on predicted fluxes to quantify importance of finer resolution simulations for regional and larger scale modeling.

Acknowledgments

Funding for this work has been provided by the Great Lakes Fisheries Trust Muskegon River Initiative, the National Science Foundation through its Water Cycles program (EAR-0233648), and the Michigan State University Center for Water Sciences.

APPENDIX A

ILHM Model Development

A.1 Evaporation and Transpiration

The basis for our potential evaporation and transpiration calculations is the modified Penman-Monteith equation (*Monteith*, 1965) presented by *Chen et al.* (2005a):

$$PE, PT = \frac{\Delta F + \rho c_p \frac{e_s - e}{r_{ai}}}{\lambda_v \left(\Delta + \gamma \left(1 + \frac{r_{ci}}{r_{ai}} \right) \right)} \quad (\text{A.1})$$

Variables appearing in Equation A.1 and others that are not explicitly defined in the text are explained in Table A.1. In Equation A.1, F is the net radiation flux (W m^{-2}), which is the product of total solar radiation measured at the *MAWN* gage multiplied by the albedo of each cell. Here we assume that albedo varies seasonally from a leaf-off “brown albedo”, a_b , condition to a peak growing season “green albedo”, a_g value. When $LAI = 0$, albedo equals a_b and increases linearly to a_g until canopy closure is complete, which we assume occurs at $LAI = 3$. Thus albedo

$$a = \begin{cases} a_g & LAI \geq 3 \\ \frac{LAI}{3} (a_g - a_b) + a_b & LAI < 3 \end{cases} \quad (\text{A.2})$$

Values of a_b and a_g are provided in Table A.2. ρ (kg m^{-3}) is the density of moist air given by the ideal gas law:

$$p = \left(\frac{P_{atm}}{R_d T} + \frac{e_s}{R_v T} \right). \quad (\text{A.3})$$

The barometric pressure, P_{atm} (Pa) is calculated according to the barometric formula (*Berberan-Santos et al.*, 1997):

$$P_{atm} = P_o \exp \left(\frac{-Mg_0z}{R_gT} \right) \quad (A.4)$$

where z is the elevation relative to mean sea level (m) given by the DEM. e_s (Pa) is calculated from the Goff-Gratch equation (*Goff and Gratch*, 1946):

$$\begin{aligned} \log_{10} e_s = & -7.90298 (T_{st}/T - 1) + 5.02808 \log_{10} (T_{st}/T) . \\ & - 1.3816 \times 10^{-7} \left(10^{11.344(1-T/T_{st})} - 1 \right) \\ & + 8.1328 \times 10^{-3} \left(10^{-3.49149(T_{st}/T-1)} - + \log_{10} e_{st} \right) \end{aligned} \quad (A.5)$$

Δ (Pa K⁻¹) is the slope of the saturated vapor pressure- temperature curve, calculated as the numerical derivative of the Goff-Gratch equation; e , the product of e_s and measured fractional relative humidity, is the ambient water vapor pressure; λ_v is the latent heat of vaporization for water (J kg⁻¹) (*Harrison*, 1963):

$$\lambda_v = 10^3 (2500.5 - 2.359T) \quad (A.6)$$

and γ is the psychrometric coefficient (Pa K⁻¹) (*Brunt*, 1952)

$$\gamma = 1.61c_p P_{atm} / \lambda_v \quad (A.7)$$

The canopy resistance to vapor transport, r_{ci} , (m s⁻¹) is calculated as

$$r_{ci} = \frac{1}{g_s LAI} \quad (A.8)$$

where g_s is the stomatal conductance (s m⁻¹). Values for maximum stomatal conductance were taken from *Schulze et al.* (1994). Unlike *Chen et al.* (2005a), the aerodynamic resistance, r_{ai} (m s⁻¹) is calculated based on canopy properties and height-adjusted gaged wind speed *Allen et al.* (1998)

$$r_{ai} = \log_e \left(\frac{h_{mw} - h_0}{l_m} \right) \cdot \log_e \left(\frac{h_{mv} - h_0}{l_v} \right) \cdot (k^2 v_{wa})^{-1} \quad (A.9)$$

Table A.1: Parameters used in calculating PET

Symbol	Definition	Units	Value	Source
R_d	Gas constant of dry air	$\text{J kg}^{-1}\text{K}^{-1}$	287.05	6
R_v	Gas constant of water vapor	$\text{J kg}^{-1}\text{K}^{-1}$	461.495	6
P_0	Reference air pressure at mean sea level	Pa	101,325	6
M	Molar mass of dry air	kg	0.029	6
g_0	Gravitational acceleration (msl, 44.5 °N)	m s^{-2}	9.80665	7
R_g	Molar gas constant of dry air	$\text{J mol}^{-1}\text{K}$	8.314	5
c_p	Constant-pressure specific heat of dry air	$\text{J kg}^{-1}\text{K}^{-1}$	1006	6
T_{st}	Steam-point temperature	K	373.15	3
e_{st}	e_s at the steam-point temperature	Pa	101,325	3
h_{mv}	Measurement height of relative humidity	m	2	-
h_{m0}	Measurement height of windspeed	m	2	-
f_0	Arbitrary windspeed-at-height offset factor	-	2	-
k	Von-Karman constant	-	0.41	1
α_w	Empirical windspeed-at-height power law coefficient	-	0.143	2
τ	Soil tortuosity	-	2	4
D_v	Molecular diffusion coefficient for water vapor	m^2s^{-1}	2.5×10^{-5}	4

¹ Allen et al. (1998) ² Elliot et al. (1986) ³ Goff and Gratch (1946) ⁴ Choudhury and Monteith (1988) ⁵ CODATA (2002) ⁶ Cengel and Boles (2001) ⁷ Halliday et al. (2004)

Because windspeed was only measured at one height, the effective measurement height is adjusted for canopy height. Here we assume that the height to which windspeed is adjusted, h_{mw} is given by

$$h_{mw} = \max(h_{m0}, f_0 h_c) \quad (\text{A.10})$$

where h_c is the canopy height assumed constant for a given land cover (Table A.2), and f_0 is a factor to move the adjusted wind height some distance above the canopy. The zero displacement height h_0 is assumed to be $2/3h_c$ (*Allen et al.*, 1998). l_m is the roughness length for momentum transport (m) taken as (*Allen et al.*, 1998) (same as the zero displacement), and l_v is the roughness length for vapor and heat transport (m) assumed to be (*Allen et al.*, 1998). v_{wa} is the measured wind speed (m s^{-1}) adjusted for measurement height according to the wind profile power law assumption (*Elliot et al.*, 1986).

$$v_{wa} = v_w \left(\frac{h_{mw}}{h_{m0}} \right)^{\alpha_w} \quad (\text{A.11})$$

where v_w is the raw measured wind speed (m s^{-1}).

In order to calculate evaporation from leaf surfaces E_c , surface depressions (E_d), and the soil (E_s), Equation A.1 is used but the two conductance terms are modified. r_{ci} is set to 0 for evaporation from leaf surfaces and surface depressions, and r_{ai} is calculated using a crop height of 2 cm for surface depression evaporation. To calculate surface soil layer evaporation, r_{ci} is replaced by r_s given by *Choudhury and Monteith* (1988):

$$r_s = \frac{\tau l_e}{\Phi D_v} \quad (\text{A.12})$$

where l_e is the depth from the surface to the top of the evaporative layer of water (m), here assumed to be half the depth of the top soil layer, and Φ is the total porosity.

The total transpiration in each cell T (m) is the sum of the root water uptake from each biologically active soil layer, T_1 and T_2 . Following *Manfreda et al.* (2005), we assume in this version of the ILHM code that the actual transpiration is calculated

Table A.2: Land-Use parameters

Symbol	Definition	Units	Land-use Type							Source
			Urban	Ag	Shrub	Forest	Wetlands	Water	Bare	
g_s	Stomatal Conductance	s m^{-1}	10.0	11.6	7.0	5.0	3.5	0.0	0.0	1
h_c	Canopy Height	m	0.8	1.0	1.0	22.0	1.6	0.01	0.01	2
β	Root Beta	-	0.943	0.961	0.964	0.97	0.96	0.95	0.95	3
a_g	Green Albedo	-	0.2	0.22	0.2	0.16	0.12	0.001	0.001	2
a_b	Brown Albedo	-	0.36	0.4	0.28	0.24	0.4	0.001	0.001	2
μ	Canopy-fraction parameter	-	0.1	0.45	0.35	0.65	0.5	0.0001	0.0001	4

¹ *Schulze et al. (1994)* ² *Walke and Tremback (2005)* ³ *Jackson et al. (1996)* ⁴ *Manfreda et al. (2005)*

from the potential value by linearly interpolating between 0 at the permanent wilting point and the potential rate at 75% of saturation according to:

$$T_i = (S_i > \Phi_{-33}) \cdot \min \left[1, \frac{4}{3} S_i (\Phi_i l_i)^{-1} \right] \cdot PT \cdot \frac{l_i}{\sum_i l_i} \quad (\text{A.13})$$

where S_i is the soil moisture (m) of the i^{th} soil layer, Φ_{-33} is the permanent wilting point of the soil (Table A.3) and l_i is the thickness of the i^{th} soil layer (m). The term $S_i > \Phi_{-33}$ is a logical statement that returns a value of “1” if true and “0” if false. Porosity values, Φ_i are taken as a function of soil type as given by Table A.3. Biologically active soil thickness is calculated as the depth above which 90% of the root mass lies using the asymptotic equation (*Gale and Grigal*, 1987):

$$y = 1 - \beta^d \quad (\text{A.14})$$

where y is the cumulative root fraction at depth $d = 0.9$ (cm) for this study; β is a land cover-dependent parameter (Table A.1). We use Equation A.14 to solve for d with a fixed cumulative root fraction $y = 0.9$,

$$l = \sum_i l_i = \log_{\beta} (0.9) . \quad (\text{A.15})$$

Total evaporation E (m) is the sum of canopy evaporation, E_c , soil evaporation E_s , and surface depression evaporation E_d . Soil evaporation is calculated according to *Chen et al.* (2005a) as

$$E_s = \min (PE_s, d_s) \cdot (1 - f_c) \quad (\text{A.16})$$

where d_s is the soil-controlled exfiltration depth (m) calculated by

$$d_s = s_e \cdot \Delta t^{-1/2} \quad (\text{A.17})$$

and Δt is the model timestep length (s), while s_e is the soil desorptivity ($\text{m s}^{-1/2}$), calculated as in *Entekhabi and Eagleson* (1989):

$$s_e = \left[\frac{8\Phi_1 k_{sat} \phi_b}{3(1+3m)(1+4m)} \right]^{1/2} S_0^{(m/2+2)} \quad (\text{A.18})$$

where k_{sat} is the saturated hydraulic conductivity of the first soil layer (m s^{-1}), m is the pore size distribution index assumed to be a function of soil texture (Table A.3), and $S_0 = \theta_1 / (\Phi_1 l_1)$ is the fractional saturation of the first soil layer. As in *Manfreda et al.* (2005), the closed canopy fraction (f_c) is defined by the empirical relationship (*Eagleson*, 1982)

$$f_c = 1 - e^{-\mu \cdot LAI} \quad (\text{A.19})$$

where μ is a constant for a given land cover type given by Table A.1.

Calculating E_c requires a full canopy water balance model. The canopy water balance is calculated using:

$$\Delta S_c = P - Int - E_c \quad (\text{A.20})$$

where ΔS_c is canopy water storage (m). Incoming rainfall is first subjected to interception up to the water holding capacity of the canopy (m) given by (*Dickinson et al.*, 1991):

$$S_{c \max} = 1 \times 10^{-4} LAI \quad (\text{A.21})$$

where LAI is the leaf area index (m^2/m^2). The available interception capacity of the canopy is then given by

$$\delta S = S_{c \max} - S_c^{t-1}. \quad (\text{A.22})$$

Additionally, we modify the model of *Chen et al.* (2005a) to allow some water to penetrate the canopy at all times based on the assumption that the canopy is not completely closed. Interception at time t is then

$$Int = \min(\delta S, P \cdot f_c). \quad (\text{A.23})$$

Canopy evaporation, E_c is then calculated as in *Manfreda et al.* (2005) with

$$E_c = \min \left[(S_c / S_{c \max})^{2/3} \cdot P E_c, S_c \right]. \quad (\text{A.24})$$

Surface depression evaporation, E_d occurs only when water is stored in surface depressions. At each time step, any water stored in surface depressions S_d from the

previous timestep is added to throughfall from the canopy, or snowmelt from the UEB model, such that precipitation excess runoff, R_e is given by

$$R_e = \min \left(0, P + S_d^{t-1} - Inf - Int \right) \quad (\text{A.25})$$

where infiltration, Inf , is calculated as discussed below. S_d is then calculated as

$$S_d = \min (S_{d\max}, R_e + Ex_1) \quad (\text{A.26})$$

where Ex_1 is the exfiltration out of the first soil layer and the depression storage capacity. $S_{d\max}$ is assumed to be constant for a given combination of slope, land cover, and soil type (see table in *Manfreda et al.* (2005)). Depression evaporation, E_d is then given by

$$E_d = \min [S_d, (1 - F_c) \cdot PE_d]. \quad (\text{A.27})$$

A.2 Infiltration, Percolation, Throughflow, and Exfiltration

The next three terms of the water balance (Equation 3.1), percolation, Pc ; throughflow, Tr ; and exfiltration Ex are calculated within the soil water balance model. First, the outputs of the canopy model, snowmelt model, and depression storage model are used to calculate infiltration into the surface soil layer

$$Inf = \min \left(P - Int + S_d^{t-1}, i_{s\max} \right). \quad (\text{A.28})$$

The infiltration capacity, $i_{s\max}$ is a function of the moisture content of the surface soil layer and is calculated according to

$$i_{s\max} = \max \left(\Phi_1 l_1 - S_1^{t-1}, i_{sat} \right) \quad (\text{A.29})$$

where l_1 is the thickness of the first soil layer as defined previously, S_1 is the moisture stored within the first soil layer, and i_{sat} is the saturated infiltration capacity, which can vary with time due to the influence of impermeable frozen soils.

Table A.3: Soil properties

Symbol	Definition	Units	Soil Texture Class						
			Sand	Loamy Sand	Sandy Loam	Silty Loam	Loam	Muck	Source
k_{sat}	Saturated Conductivity	$\times 10^6 \text{ m s}^{-1}$	58.3	17.0	7.19	1.9	3.67	1.19	1
Φ	Total Porosity	-	0.437	0.437	0.453	0.501	0.463	0.398	1
θ_r	Residual Water Content	-	0.020	0.035	0.041	0.015	0.027	0.068	1
i_{sat}	Infiltration Capacity	$\times 10^6 \text{ m s}^{-1}$	58.3	17.0	7.19	1.9	3.67	1.19	2
Φ_{-33}	Field Capacity	-	0.14	0.15	0.2	0.28	0.25	0.25	2
$\Phi_{-10,000}$	Wilting Point	-	0.07	0.07	0.09	0.10	0.11	0.16	2
α	van Genuchten parameter	m^{-1}	14.5	12.4	7.5	2.0	3.6	5.9	3
N	van Genuchten parameter	-	2.68	2.28	1.89	1.41	1.56	1.48	3
ϕ_b	Bubbling Presure	m	0.073	0.087	0.147	0.208	0.112	0.281	1
m	Pore Size Distribution	-	0.695	0.553	0.378	0.234	0.252	0.319	1

Infiltration is applied to the first soil layer, which can then percolate into the second layer. First, the soil moisture storage at the end of each timestep in the first layer is calculated as

$$S_1^{t+} = S_1^{t-1} + Inf - E_s - T_1 + Tr_1 - P_1 + Ex_2 - Ex_1 \quad (A.30)$$

where P_1 is the percolation of water from the first soil layer to the second. Note that calculating T_1 requires knowing S_1 . To avoid having to solve the coupled equations, T_1 is calculated using an intermediate value of $S_1^{t+} = S_1^{t-1} + i_s - E_s$. Then $S_1 = S_1^{t-} - T_1$, and $S_1^{t+} = S_1 + Tr_1 - P_1 + Ex_2 - Ex_1$. Given S_1 , percolation into the second layer, P_1 is given by

$$P_1 = \max(S_1 - S_{1\max}, P_{DREAM}) \quad (A.31)$$

where $S_{1\max} = \Phi_1 l_1$ and P_{DREAM} is the percolation calculated according to *Manfreda et al.* (2005) given by:

$$P_{DREAM} = \begin{cases} 0 & S_i(t) \leq \Phi_{-33} l_i \\ S_{i\max} \left(S_i - \left[\frac{\Delta t \cdot k_{sat}(\gamma-1)}{S_{i\max}} + \left(\frac{S_i}{S_{i\max}} \right)^{1-\gamma} \right]^{1/(1-\gamma)} \right) & S_i(t) > \Phi_{-33} l_i \end{cases} \quad (A.32)$$

where $\gamma = (2 + 3m)/m$. P_{DREAM} effectively allows percolation only when soil moisture exceeds the field capacity given by $\Phi_i l_i$.

The second-layer soil moisture at the end of the timestep, (S_2), is calculated similarly to Equation A.30:

$$S_2^{t+} = S_2^{t-1} + P_1 - P_2 - T_2 + Tr_2 - Ex_2 \quad (A.33)$$

where T_2 is calculated from the values of S_2^{t-1} at the previous timestep.

$$P_c = P_2 = P_{DREAM} \quad (A.34)$$

is calculated from an intermediate value of $S_1^{t-} = S_1^{t-1} + P_1 - T_2$.

Throughflow out of a cell is calculated as

$$Tr_{out\,i} = l_i \cdot \Delta x \cdot \left(\frac{2}{1/k_{\theta i} + 1/k_{\theta i,down}} \right) \cdot \left[\frac{\Delta z}{\Delta x} + \left(\frac{\Delta \psi}{\Delta \theta} \right)_i \cdot \left(\frac{\theta_{i,down} - \theta_i}{\Delta x} \right) \right] \quad (\text{A.35})$$

where the effective unsaturated hydraulic conductivity for subsurface flow in layer i is taken as the harmonic average of the unsaturated hydraulic conductivity in the cell, $k_{\theta i}$ (m s^{-1}) and the down slope value $k_{\theta i,down}$; Δx is the model cell resolution (m); $\Delta z/\Delta x$ is the vertical gradient in the down-slope direction; $(\Delta \psi/\Delta \theta)_i$ is the slope in meters of the moisture retention curve in layer i , and

$$\theta = \frac{S_i(t)/S_{i\max} - \theta_r}{\Phi - \theta_r} \quad (\text{A.36})$$

where θ_r is the residual volumetric moisture content assumed to be a soil-texture dependent property (see Table A.3). The assumption that flow only occurs parallel to the dip of the slope requires that $Tr_{out} \geq 0$. Tr_i is then calculated as

$$Tr_i = Tr_{out,up} - Tr_{out,i}. \quad (\text{A.37})$$

Finally, Ex_2 is calculated as the soil water in excess of saturation given by

$$Ex_2 = \max(0, S_2 + Tr_2 - S_{2\max}). \quad (\text{A.38})$$

This is then applied to the first layer prior to calculating

$$Ex_1 = \max(0, S_1 + Tr_1 + Ex_2 - S_{1\max}). \quad (\text{A.39})$$

A.3 Runoff Routing

Water exfiltrated from layer one is then applied to the surface depression model, thus R (m) is simply

$$R = R_e - S_d \quad (\text{A.40})$$

which is then routed to the streams using an approach modified from that presented by *Manfreda et al.* (2005). Once generated, runoff cannot infiltrate and is instead

routed overland and through streams according to the D8 flowdirection algorithm in ARC (*ESRI*, 2003) with runoff times given by the velocities in each cell along the flowpath. Runoff is assumed to travel overland at a velocity given by the Kerby time of concentration equation (*Kerby*, 1959)

$$t_{cell} = 86.735 \left(\frac{l_{cell} \cdot n}{\sqrt{s}} \right)^{0.467} \quad (\text{A.41})$$

where t_{cell} is the time required to completely traverse a model cell (s), l_{cell} is the length of the model cell (m), n is the dimensionless Manning's Roughness coefficient (values from *McCuen* (2004)) and s is the fractional slope of the cell in the downslope direction. The velocity (m s^{-1}) is then

$$v_{land} = \frac{l_{cell}}{t_{cell}}. \quad (\text{A.42})$$

Once the runoff enters the stream channel its velocity is calculated using Manning's Equation (*McCuen*, 2004)

$$v_{stream} = \frac{1}{n} r^{2/3} s^{1/2} \quad (\text{A.43})$$

where r is the hydraulic radius (m) given by the ratio of the stream cross-sectional area to the wetted perimeter.

APPENDIX B

ILHM Regional Upscaling

This appendix provides additional equations from those in Appendix A and Chapter 4. In some instances, variable names and symbols have been changed from Appendix A to be more consistent.

B.1 Surface Domain

B.1.1 Shortwave Radiation

Incoming total shortwave solar radiation is provided to the model as two constituents: beam and diffuse radiation, Q_{SB} and Q_{SD} [W/m^2]. For this study, the beam and diffuse fractions were obtained from real ground solar radiation modeling (see Solar Radiation Preparation, below), though one could use the method of *Chen et al.* (2005a). The canopy then intercepts a portion of each both constituents, transmitting to the ground beam radiation $Q_{GB} = Q_{SB} \times \tau_{CB}$ and diffuse radiation $Q_{GD} = Q_{SD} \times \tau_{CD}$ where τ_{CB} is the canopy transmissivity for beam radiation, and τ_{CD} for diffuse. These transmissivities are defined as

$$\tau_{CB} = \frac{F_G}{\cos(\theta_B)}, \text{ and} \quad (B.1)$$

$$\tau_{CD} = \frac{F_G}{\cos(\theta_D)}, \quad (B.2)$$

where θ_B is the zenith angle (degrees from vertical) of beam radiation, and θ_D is the effective diffuse radiation angle, defined as $\theta_D = 0.537 + 0.025L$ (Chen *et al.*, 2005a), and F_G is the gap fraction of the canopy, defined as

$$F_G = \exp [-0.5\Omega (LAI + SAI)] \quad (B.3)$$

where Ω is the canopy clumping index, a measure of the distribution of leaves within the canopy. Ω values near 0.5 indicate that leaves are distributed in rows, and values closer to 1 that the leaves are randomly distributed within the canopy. LAI is the leaf area index [m^2/m^2] of the canopy representing the leaf area per unit surface area, and SAI is the stem area index.

The ground then reflects back to the canopy a portion of the transmitted shortwave radiation, given by

$$Q_{GC} = (F_W + F_S) \times (\alpha_{WB}Q_{GB} + \alpha_{WD}Q_{GD}) + F_U (\alpha_{UB}Q_{GB} + \alpha_{UD}Q_{GD}) \quad (B.4)$$

where F_{WS} is the fraction of each cell covered by streams and wetlands, and F_U the fraction by uplands; these terms are related by $1 = F_W + F_S + F_U$. The α terms refer to shortwave albedos (weighted average of visible and near IR) for wetlands and uplands to beam and diffuse radiation. It is assumed that the canopy then reflects back to the ground a negligible portion of this radiation, thus the infinite series of reflected radiation terms is truncated after the first reflection (ground - canopy). The net radiation absorbed by the canopy in each cell is then given by

$$Q_{CN} = (1 - \alpha_{CD}) \times (Q_{SD} - Q_{GD} + Q_{GC}) + (1 - \alpha_{CB}) \times (Q_{SB} - Q_{GB}). \quad (B.5)$$

Here, the ground-reflected shortwave radiation Q_{GC} is assumed to be diffuse radiation. The ground absorbs

$$Q_{GN} = (Q_{GB} + Q_{GD}) - Q_{GC}. \quad (B.6)$$

From this equation, the net shortwave radiation in the upland portion of each cell is $Q_{UN} = Q_{GN}F_U$, while for the wetland and stream portions $Q_{WN} = Q_{GN}(F_W + F_S)$.

For this study, the beam and diffuse radiation terms for the canopy are assumed to be equal, and are given as a linear mixing of two end members: the brown and green albedo of a given vegetative canopy type as in *Walko and Tremback (2005)*.

$$\alpha_{CD} = \alpha_{CB} = \begin{cases} \alpha_{C,gr} & LAI \geq 3 \\ \frac{LAI}{3} (\alpha_{C,gr} - \alpha_{C,br}) + \alpha_{C,br} & LAI < 3 \end{cases}, \quad (B.7)$$

with $\alpha_{C,gr}$ and $\alpha_{C,br}$ are the albedos of the canopy when fully green and fully brown, respectively. Frozen water stored on the canopy can have a significant effect on the albedo, which is described by the following two equations:

$$\begin{aligned} \alpha_{CB} &= \alpha_{CB}(1 - F_{CS}) + \alpha_{SN}F_{CS}, \text{ and} \\ \alpha_{CD} &= \alpha_{CD}(1 - F_{CS}/2) + \alpha_{SN}(F_{CS}/2). \end{aligned} \quad (B.8)$$

In Equation B.8, α_{SN} is the albedo of new snow, and F_{CS} is the fraction of canopy covered by snow, calculated as $F_{CS} = S_C/S_{CM}$ where S_C is the total water stored on the canopy, which is assumed to be frozen when the air temperature is below freezing, and S_{CM} is the maximum canopy water storage. The factor of 2 in the denominator of diffuse albedo calculation arises from the fact that snow only covers the top half of each leaf, and only a small portion of the total stem area. Since beam radiation primarily sees the tops of leaves and stems, no modification is necessary.

Net upland albedo is calculated according to the net soil albedo and the beam and diffuse components of snow albedo. Bare ground upland albedo is calculated via the linear combination of dry and wet soil albedos, as $\alpha_{SO} = \alpha_{SO,dry}(1 - F_{SAT,1}) + \alpha_{SO,wet}F_{SAT,1}$. α_{SO} is the net soil albedo, $\alpha_{SO,dry}$ is the dry soil albedo, $\alpha_{SO,wet}$ is the wet soil albedo, and $F_{SAT,1}$ is the soil saturation fraction in the first soil layer. The soil saturated fraction in layer i is defined as $F_{SAT,i} = S_{S,i}/S_{S,max,i}$, where $S_{S,i}$ is the soil water storage in a given layer [m], and $S_{S,max,i}$ is the maximum water

storage. $S_{S,\max,i} = \Phi_{S,i} D_{S,i}$ where $\Phi_{S,i}$ is the total soil porosity in a given layer and $D_{S,i}$ is the thickness of that soil layer [m]. Diffuse and beam snow albedo on the upland portion of a cell, α_{SND} and α_{SNB} , are a function of the age of the snow at the air interface. They are calculated as in the UEB snow model *Tarboton and Luce* (1996), which itself is derived partially from the land surface model BATS *Yang et al.* (1997). This formulation includes a modification of the albedo to account for shallow snow and partial transmission of the visible and near IR components of shortwave radiation. The net upland albedo is calculated as

$$\begin{aligned} \alpha_{UB} &= \begin{cases} \alpha_{SOB} & D_{SN} = 0 \\ \alpha_{SNB} & D_{SN} > 0 \end{cases} \\ \alpha_{UD} &= \begin{cases} \alpha_{SOD} & D_{SN} = 0 \\ \alpha_{SND} & D_{SN} > 0 \end{cases} \end{aligned} \quad (B.9)$$

Here, D_{SN} is the depth of snow in the upland portion of each cell.

In this version of the model, snow cover in wetland and stream portions of cells is ignored, thus wetland albedo depends only on the albedo of water and any ice cover. The diffuse albedo of open water is α_{WOD} , and the beam albedo is $\alpha_{WOB} = \alpha_{WOD}/(\cos(\theta_B)^{1.7} + 0.15)$ (*Bonan*, 1996). Ice albedo is assumed constant, $\alpha_{WID} = \alpha_{WIB}$. The net wetland albedo is then

$$\begin{aligned} \alpha_{WB} &= \begin{cases} \alpha_{WOB} & D_{ICE} = 0 \\ \alpha_{WIB} & D_{ICE} > 0 \end{cases} \\ \alpha_{WD} &= \begin{cases} \alpha_{WOD} & D_{ICE} = 0 \\ \alpha_{WID} & D_{ICE} > 0 \end{cases} \end{aligned} \quad (B.10)$$

Table B.1: Land use dependent parameters in ILHM

<i>Land Use</i>	$\beta[-]$	$h_{C,\max}[m]$	$h_{C,\min}[m]$	$\alpha_{C,\text{gr}}[-]$	$\alpha_{C,\text{br}}[-]$	$SAI[-]$	$\Omega[-]$	$g_{s,\max}[s/m \times 10^3]$
urban	0.954	6.0	6.0	0.20	0.36	0.20	0.75	10.0
agriculture	0.961	1.0	0.0	0.22	0.40	0.30	0.73	11.6
shrubland	0.950	1.0	0.5	0.20	0.28	0.10	0.71	7.0
deciduous	0.966	15.0	15.0	0.20	0.24	0.50	0.69	4.6
water	0.950	0.01	0.01	0.00	0.00	0.00	0.00	0.10
wetland	0.970	1.60	1.60	0.17	0.24	0.30	0.65	3.5
barren	0.950	0.10	0.10	0.21	0.15	0.00	0.00	0.10
coniferous	0.976	15.0	15.0	0.41	0.24	0.70	0.62	5.5
<i>source</i>	1	2	2	2	2	3	4	5

¹(*Jackson et al.*, 1996) ²(*Walke and Tremback*, 2005) ³(*Bonan*, 1996) ⁴(*Chen et al.*, 2005b) ⁵*Schulze et al.* (1994)

Table B.2: Depression capacity, $S_{DEP,max}$ [$m \times 10^3$]. Values are from *Manfreda et al.* (2005), assuming no soil-texture dependence, and taking values for clay soil textures.

<i>Land Use</i>	<i>Slope [%]</i>			
	0 - 0.5	0.5 - 5.0	5.0 - 10.0	> 10.0
urban	1.80	1.37	0.94	0.51
agriculture	1.50	1.08	0.66	0.23
shrubland	2.00	1.46	0.93	0.39
deciduous	2.50	1.88	1.25	0.63
water	0	0	0	0
wetland	0	0	0	0
barren	1.00	0.71	0.42	0.12
coniferous	2.50	1.88	1.25	0.63

B.1.2 Snowpack Accumulation and Melt

In Appendix A, the snow depth was calculated at only a single point, and averaged across the watershed. Here, the code was modified to be fully-explicit, and snow depths are calculated for each cell in the model.

B.1.3 Infiltration and Root-zone Moisture Vertical Redistribution

Water falling through the canopy may infiltrate, according to

$$W_{INF} = \min (W_{THR}F_{IM}, W_{INF,max}) , \quad (B.11)$$

where $W_{INF,max}$ is the maximum infiltration capacity, defined as

$$W_{INF,max} = \begin{cases} K_{SAT} & T_1 \geq T_{SF} \\ 0.1 \cdot K_{SAT} & T_1 < T_{SF} \end{cases} , \quad (B.12)$$

and F_{IM} is the fraction of the cell surface covered by impermeable materials. $F_{IM} = F_{IMS}F_{URB}$ where F_{IMS} is the specific urban impermeable fraction, and F_{URB} is the fraction of each cell that is classified as urban land use. In Equation B.12, T_1 is the temperature of the 1stsoil layer [C], in this study given by input data, T_{SF} is the soil freezing threshold temperature [C], and K_{SAT} is the saturated hydraulic conductivity [m/s].

Table B.3: Soil-texture dependent parameters. Texture name abbreviates S, L, Si, C, and M are sand, loam, silt, clay, and muck, respectively. If multiple sources are listed for a particular parameter, values obtained from multiple sources were averaged

Texture	$\alpha_{S,dry}$ [—]	θ_{WP} [%]	θ_{FC} [%]	θ_R [%]	$\theta_{S \text{ or } \Phi}$ [%]	K_{sat} [$m/s \times 10^6$]	K_0 [$m/s \times 10^6$]	α [m^{-1}]	N [—]	L [—]	ϕ_b [m]	m [—]
S	0.23	5.0	9.6	3.7	42.4	54.2	2.83	3.52	3.18	-0.93	0.07	0.69
LS	0.27	5.0	12.3	4.2	42.9	18.7	2.82	3.48	1.75	-0.87	0.09	0.55
SL	0.24	8.0	19.4	4.0	42.4	8.54	1.79	2.67	1.45	-0.86	0.15	0.38
L	0.25	14.0	27.5	4.4	44.1	3.12	0.43	1.11	1.47	-0.37	0.11	0.25
SiL	0.25	11.0	32.0	4.0	47.3	2.83	0.20	0.51	1.66	0.37	0.21	0.23
Si	0.15	6.0	30.0	5.0	48.5	5.59	0.39	0.66	1.68	0.62	0.20	0.20
SCL/M	0.09	17.0	26.3	6.6	40.4	1.95	0.80	2.11	1.33	-1.28	0.28	0.32
CL	0.23	22.0	33.9	7.7	46.2	0.93	0.58	1.58	1.42	-0.76	0.26	0.24
SiCL	0.25	22.0	37.3	6.5	48.8	1.10	0.26	0.84	1.52	-0.16	0.33	0.18
SC	0.15	25.0	35.0	11.3	41.8	0.68	0.50	3.34	1.21	-3.67	0.29	0.22
SiC	0.30	27.0	39.9	8.4	49.3	0.80	0.37	1.62	1.32	-1.29	0.34	0.15
C	0.30	30.0	40.8	9.4	47.8	0.73	0.34	1.50	1.25	-1.56	0.37	0.17
<i>source</i>	1	2,3	2	3,4	2,3,4	2,3,4	4	4	4	4	3	3

¹(USDA-NRCS, 2008) ²(Sarton and Rawls, 2006) ³(Stieglitz et al., 1997) ⁴(Schaap et al., 1998)

Table B.4: Table listing physical constants

<i>Symbol</i>	<i>Variable Description</i>	<i>Units</i>	<i>Value</i>
ρ_{air}	density of dry air @ 10 °C	kg m ⁻³	1.247
ρ_{ice}	density of ice	kg m ⁻³	920
ρ_{min}	density of mineral soil matrix	kg m ⁻³	2650
ρ_{h2o}	density of water	kg m ⁻³	1000
R_{air}	gas constant dry air	J kg ⁻¹ K ⁻¹	287.05
R_{h2o}	gas constant water vapor	J kg ⁻¹ K ⁻¹	461.5
g	gravitational acceleration @ msl and 44.5°N	m s ⁻²	9.80665
$C_{p,air}$	heat capacity of moist air	J kg ⁻¹ K ⁻¹	1013
$C_{p,ice}$	heat capacity of ice	J kg ⁻¹ K ⁻¹	2114
$C_{p,mnrl}$	heat capacity of mineral soil matrix	J kg ⁻¹ K ⁻¹	754.7
$C_{p,org}$	heat capacity of organic soil matrix	J kg ⁻¹ K ⁻¹	1923
$C_{p,h2o}$	heat capacity of water	J kg ⁻¹ K ⁻¹	4181.3
ν_{air}	kinematic viscosity of air	m ² s ⁻¹ × 10 ⁵	1.5
H_f	latent heat of fusion	J kg ⁻¹ × 10 ⁻⁵	3.34
H_s	latent heat of sublimation	J kg ⁻¹ × 10 ⁻⁵	28.34
w_{air}	molecular weight dry air	kg mol ⁻¹ × 10 ²	2.8964
w_{h2o}	molecular weight water	kg mol ⁻¹ × 10 ²	1.8015
σ	Stefan Boltzmann constant	J s ⁻¹ m ⁻² K ⁻¹ × 10 ⁸	5.6704
$K_{q,air}$	thermal conductance air	W m ⁻¹ K ⁻¹	0.025
$K_{q,ice}$	thermal conductance ice	W m ⁻¹ K ⁻¹	2.2
$K_{q,mnrl}$	thermal conductance mineral soil matrix	W m ⁻¹ K ⁻¹	2.9
$K_{q,org}$	thermal conductance organic soil matrix	W m ⁻¹ K ⁻¹	0.25
$K_{q,h2o}$	thermal conductance water	W m ⁻¹ K ⁻¹	0.57
k_vK	von Karman's constant	-	0.41

Table B.5: Table listing albedo and emissivity parameters, and sources

<i>Symbol</i>	<i>Variable Description</i>	<i>Units</i>	<i>Value</i>	<i>Source</i>
$\alpha_{SN,vis}$	albedo snow vis	-	0.95	1
$\alpha_{SN,nir}$	albedo snow nir	-	0.72	1
α_{WOD}	albedo water diff	-	0.06	2
α_{WOB}	albedo water beam	-	0.06	2
$\alpha_{WI,dry}$	albedo ice dry	-	0.50	3
$\alpha_{WI,wet}$	albedo ice wet	-	0.25	4
ϵ_{SN}	emissivity snow	-	0.98	5,6
ϵ_{SO}	emissivity soil	-	0.96	5
ϵ_{h2o}	emissivity water	-	0.96	4
ϵ_{ice}	emissivity ice	-	0.97	4
$D_{\alpha SN}$	albedo extinction depth, snow	m	0.10	6

¹(Aoki et al., 2003) ²(Hostetler and Bartlein, 1990) ³(Yongjiu et al., 2001) ⁴(Patterson and Harnblin, 1988) ⁵(Bonan, 1996) ⁶(Tarboton and Luce, 1996)

Table B.6: Table of domain-independed physical parameters and sources

<i>Symbol</i>	<i>Variable Description</i>	<i>Units</i>	<i>Value</i>	<i>Source</i>
$h_{R,wet}$	roughness height, wetland	m	0.10	-
$h_{R,so}$	roughness height, soil	m	0.01	-
T_{SOF}	soil freezing temperature	°C	-0.25	-
T_{CD}	canopy dormant temperature	°C	1.0	1
SCC	specific canopy capacity	mm m ⁻² (leaf)	0.15	2
η_W	wetland light attenuation coefficient	m ⁻¹	0.70	3
P_0	wetland neutral Prandtl number	-	1.0	4
α_W	wind-offset power	-	0.14	5
f_0	wind-offset factor	-	2.0	-

¹(Chen et al., 2005a) ²Average of Dickinson et al. (1991) and Yongjiu et al. (2001)

³Modified from Bonan (1996) ⁴(Hostetler and Bartlein, 1990) ⁵(Allen et al., 1998)

Table B.7: Table of domain-dependent physical parameters, values are explained in Chapter 4

<i>Symbol</i>	<i>Variable Description</i>	<i>Units</i>	<i>Value</i>
F_{ISO}	soil frozen impermeable factor	-	0.90
F_{IURB}	specific urban impermeability	-	0.20
C_{WETD}	conductivity, disconnected wetlands	m m ⁻¹ (head) hr ⁻¹ × 10 ⁶	3.60
C_{WETC}	conductivity, connected wetlands	m m ⁻¹ (head) hr ⁻¹ × 10 ³	3.125
A_{WT}	water table amplitude	m	1.25
F_{RWET}	recession constant, wetlands	m m ⁻¹ (head) hr ⁻¹ × 10 ³	7.0

Movement of water vertically in the shallow root zone can be calculated according to one of two methods: a 1-D Richard's equation (*Hillel*, 1980), or via a free-drainage model similar to that in Appendix A. For this study, the free-drainage model was chosen because of its stability and computational efficiency. For this model, the movement of water from one soil layer to the next is calculated as:

$$W_{S,i} = \min \left(\max \left(S_{S,i} - S_{SFC,i}, 0 \right), K_{\theta,i} \right). \quad (\text{B.13})$$

Here, $S_{S,i}$ is the water stored in layer i [m], $S_{SFC,i}$ is the field capacity of the soil layer [m], and $K_{\theta,i}$ is the unsaturated hydraulic conductivity in that layer. $K_{\theta,i}$ is calculated via a modified form of the Maulem-vanGenuchten equations (*Maulem*, 1976; *van Genuchten*, 1980; *Schaap and Leij*, 2000)

$$K_{\theta} = K_0 \theta' \left[1 - \left(1 - \theta'^{1/M} \right)^M \right]^2. \quad (\text{B.14})$$

In Equation B.14, K_0 is a soil-textural dependent value that is less than K_{SAT} , which provides a more accurate description of unsaturated conductivity according to *Schaap and Leij* (2000). $\theta' = (\theta - \theta_R)/(\theta_S - \theta_R)$ where θ is the volumetric soil moisture [m^3/m^3], θ_R is the residual volumetric soil moisture, and $\theta_S = \Phi_S$ is the total porosity in a given soil layer. Also, in Equation B.14, $M = 1/N$, where N is the van Genuchten empirical parameter. Flux out of the bottom soil layer into the deep unsaturated zone is terms deep percolation, W_{DPU} .

B.1.4 Canopy Transpiration

Upland root-zone transpiration has been slightly modified from Appendix A to account for the relative fraction of root mass within each soil layer. The new formulation is

$$W_{TU,i} = \begin{cases} \min(1, 4/3 \cdot F_{SAT}) \cdot W_{TP} \cdot F_{SR,i} & S_{S,i} > S_{SW,i} \\ 0 & S_{S,i} \leq S_{SW,i} \end{cases}, \quad (\text{B.15})$$

where W_{TP} is potential canopy transpiration [m/s], calculated as in Appendix A, $F_{SR,i}$ is the fraction of root mass in the root zone within layer i , and S_{SW} is the wilting capacity of the layer [m]. The fraction of the root mass within each layer is calculated by assuming that root mass declines exponentially with depth, according to the following equation, modified from *Gale and Grigal* (1987)

$$F_{SR,i} = \beta^{(D_{S,i-1}+D_{S,i})/2/100} - \beta^{(D_{S,i}+D_{S,i+1})/2/100}, \quad (B.16)$$

dealing appropriately with the first and last soil layers. It can be necessary to normalize $F_{SR,i}$ by the sum $\sum_{i=1}^{N_S} F_{SR,i}$ for cases when soils are thin relative to the root depth. The parameter β is a land-cover dependent parameter, and $D_{S,i}$ refers to the cell-center depth of layer i [m].

In wetlands, transpiration occurs at a rate equal to the potential, W_{PT} , unless limited by available moisture. For wetlands connected to the water table, transpiration is removed directly from the phreatic zone at the potential rate as $W_{TW,P} = W_{TP}$ [m/s]. Transpiration is removed surface storage in disconnected wetlands up to the limit of available moisture,

$$W_{TW,S} = \begin{cases} W_{TP} & S_{WS} > S_{WS,\min} \\ 0 & S_{WS} \leq S_{WS,\min} \end{cases}. \quad (B.17)$$

B.1.5 Inundated Area Extent

Inundated areas are defined as wetland landunits containing ponded water at the surface. This can occur in two situations: 1) wetland sediments have low conductance, and are disconnected from the water table, and 2) the water table in connected wetlands is above the bed elevation. The first situation is handled simply by tracking the storage in disconnected wetlands. The second situation requires some capability to describe the transient depth of the water table. Provided a fully-coupled surface

and groundwater system, this becomes straightforward. As this capability is not yet available in ILHM, another approach is required.

Generally speaking, the water table depth varies annually roughly sinusoidally, peaking in the spring and early summer in response to snowmelt and spring infiltration, and reaching an annual minimum sometime near October. Locations with thicker unsaturated zones exhibit smaller amplitudes of variability and phase shift relative to shallow unsaturated zone areas. For shallow unsaturated zone areas, the water table is assumed to vary sinusoidally as:

$$D_{WT} = A_{WT} \sin \left(\frac{2\pi (DD - DD_0)}{366} \right) \quad (\text{B.18})$$

where $D_{WT}[m]$ is the depth to the water table, $A_{WT} [m]$ is the amplitude of water table fluctuation in areas of shallow water table, $DD [day]$ is the decimal day of the year, and $DD_0 [day]$ is an offset to control the phase shift of the water table. For this study, $DD_0 = 0$, $A_{WT} = -1.25$ which results in a peak water table elevation (a minimum depth) in April, and a maximum depth in October.

To then calculate which wetlands are inundated as a function of water table depth, each wetland is assigned an average depth to water (Table 4.3). The depth to water in a particular wetland is then

$$D_{WW,i} = D_{WA,i} - D_{WT} - D_{WS,i}, \quad (\text{B.19})$$

with depth to water in wetland cell i equal to $D_{WW,i} [m]$, $D_{WA} [m]$ is the average depth to water in the wetland cell i , and D_{WS} is the thickness of water stored in that wetland cell, calculated as:

$$D_{WS,i} = \begin{cases} W_{WS,i} & W_{WS,i} \geq 0 \\ W_{WS,i} \Phi_{1,i} & W_{WS,i} < 0 \end{cases}. \quad (\text{B.20})$$

Note that W_{WS} , the water stored in wetlands, may be negative, indicating water stored in the soil within a non-inundated wetland cell. In Equation B.20, $\Phi_{1,i}$ is the

soil porosity in the wetland landunit in the first layer of cell i . If the depth to water $D_{WW} < 0$, the wetland cell is inundated.

B.1.6 Wetland Evaporation, Lake Temperature, Ice Pack

Wetland evaporation is calculated directly from the lake temperature, air temperature and aerodynamic resistance of each wetland cell. First, the sensible heat flux is calculated as

$$Q_H = \frac{\rho_{m,air} C_{p,air}}{\tau_{ai} (T_{air} - T_W)} \quad (B.21)$$

where Q_s [W/m^2] is the sensible heat flux, $\rho_{m,air}$ [kg/m^3] is the moist air density, $C_{p,air}$ is the heat capacity of air, τ_{ai} [s/m] is the aerodynamic resistance (calculated via Equation A.9), T_{air} [$^{\circ}C$] is the input air temperature, and T_W [$^{\circ}C$] is the wetland surface layer temperature. The moist air density is calculated via Equation A.3. After calculating sensible heat flux, latent heat flux Q_L [W/m^2] is calculated by

$$Q_L = \lambda_{ev} PE_W \quad (B.22)$$

with λ_{ev} the latent heat of evaporation, and PE_W [m/s] the potential evaporation rate of wetlands. The latent heat of evaporation is

$$\lambda_{ev} = \frac{\Delta_e}{e_s} \cdot R_{h2o} \cdot (T_W + 273.15) \quad (B.23)$$

where Δ_e is the gradient of the vapor-pressure temperature curve, e_s is the saturated vapor pressure at the water temperature T_W , and R_{h2o} is the gas constant of water vapor.

Lake temperature is calculated by solving the energy balance Equation 4.6 using the method of *Hostetler and Bartlein* (1990). Ice pack thickness is calculated in this process module as discussed in Chapter 4.

Table B.8: Lookup table of unsaturated zone wetting front velocities.

K_{sat} [m/d]	V_{WF} [d/m]
1.7	4.18
2.5	3.13
5.0	2.47
10	1.81
20	1.62

B.2 Deep Unsaturated Zone

Unsaturated zone routing has been modified from Appendix A, adding variable wetting-front velocity. Wetting fronts are assumed to move with constant velocity through the deep unsaturated zone. All moisture leaving the root-zone is accumulated for the length of the groundwater domain stress period, currently 7 days. This pulse of water then moves through the unsaturated zone at a fixed rate, depending on the calibrated saturated conductivity of the deep sediments.

No data from monitoring wells in the MRW were available for this study, so wells instrumented in the nearby Grand Traverse Bay Watershed were used to calibrate a 1-D Richard's equation model. For this model, a saturated conductivity of 5 m d^{-1} was assumed. van Genuchten equation parameters α and N were calibrated such that the observed mean wetting front velocity V_{WF} [d/m] (determined by the mean arrival time of a pulse of water traveling through a simulated 10 m homogeneous column) matched that observed in wells across the GTBW, roughly 2.5 d m^{-1} . These values are $\alpha = 1.76$ and $N = 3.2$. These values were then held constant as the saturated hydraulic conductivity, K_{sat} , was varied to build a lookup table spanning the range of values used in the MRW saturated groundwater domain module. Table B.8 lists the values obtained through this procedure.

Table B.9: Table of preliminarily-calibrated saturated hydraulic conductivities for deep sediments

Sediment Type	Conductivity [m/d]
Water	100
Peat and Muck	5
Dune Sand	50
Lacustrine Fines	5
Lacustrine Sand and Gravel	20
Outwash and Alluvium	15
Ice-Contact Outwash	20
Morainal Tills	8
Layers 2 and 3	2.5

B.3 Saturated Zone

Discharge from each “drain” type cell in the MODFLOW model is recorded at every timestep. After completing the MODFLOW run, these discharge values are then summed by sub-basin (there are approximately 40 sub-basins in the MRW). Evaporation demand not satisfied by surface storage is then passed to this step. This demand is summed at the same sub-basin level as discharge, and subtracted from discharge as detailed in Equation 4.10.

BIBLIOGRAPHY

- Ahuja, L., J. Hanson, K. Rojas, and M. Shaffer (Eds.) (1999), *The Root Zone Water Quality Model*, Water Resources Publications, CO USA.
- Akbari, H., L. Rose, and H. Taha (2003), Analyzing the land cover of an urban environment using high-resolution orthophotos, *Landscape and Urban Planning*, 63, 1–14.
- Allen, R., L. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration - guidelines for computing crop water requirements - FAO irrigation and drainage paper 56, *Tech. rep.*, Food and Agriculture Organization (FAO) of the United Nations, Rome.
- Ambaum, M., B. Hoskins, and D. Stephenson (2001), Arctic Oscillation or North Atlantic Oscillation?, *Journal of Climate*, 14(16), 3495–3507.
- Anderson, J., E. Hardy, J. Roach, and R. Witmer (1976), A land use and land cover classification system for user with remote sensor data, *Tech. rep.*, US Geological Survey Professional Paper 964.
- Andresen, J., and J. Winkler (2009), *Michigan Geography and Geology*, chap. Chapter 19: Weather and Climate, Pearson Custom Publishing, Boston, MA.
- Andresen, J. S. (2004), Level-III NEXRAD data, Archived from the National Weather Service.
- Andresen, J. S. (2007), Automated Weather Observing System (AWOS) climate data, Archived from the National Weather Service.
- Andresen, J. S. (2008), Level-iii nexrad data, Archived from the National Weather Service.
- Aoki, T., A. Hachikubo, and M. Hori (2003), Effects of snow physical parameters on shortwave broadband albedos, *Journal of Geophysical Research*, 108(D19), 4616.
- Arnold, J., and P. Allen (1996), Estimating hydrologic budgets for three illinois watersheds, *Journal of Hydrology*, 176(1-4), 57–77.
- Arnold, J., P. Allen, and G. Bernhardt (1993), A comprehensive surface-groundwater flow model, *Journal of Hydrology*, 142, 47–69.

- Assel, R. (1992), Great Lakes winter-weather 700-hPa PNA teleconnections, *Monthly Weather Review*, *120*(9), 2156–2163.
- Assel, R., K. Cronk, and D. Norton (2003), Recent trends in Laurentian Great Lakes ice cover, *Climatic Change*, *57*, 185–204.
- Austin, J., and S. Colman (2007), Lake Superior summer water temperatures are increasing more rapidly than regional air temperatures: A positive ice-albedo feedback, *Geophysical Research Letters*, *34*, L06,604.
- Avnir, D., O. Biham, D. Lidar, and O. Malacai (1998), Is the geometry of nature fractal?, *Science*, *279*, 39–40.
- Bach, M., and T. Meigen (1999), Do's and don'ts in Fourier analysis of steady-state potentials, *Documenta Ophthalmologica*, *99*, 69–82.
- Baret, F., and G. Guyot (1991), Potentials and limits of vegetation indices for LAI and APAR assessment, *Remote Sensing of Environment*, *35*, 161–173.
- Berberan-Santos, M., E. Bodunov, and L. Pogliani (1997), On the barometric formula, *American Journal of Physics*, *65*(5), 404–412.
- Beven, K., and M. Kirkby (1979), A physically-based variable contributing area model of basin hydrology, *Hydrology Science Bulletin*, *24*(1), 43–69.
- Blackman, R., and J. Tukey (1958), *The measurement of power spectra, from the point of view of communications engineering*, New York: Dover Publications.
- Bogena, H., R. Kunkel, T. Schobel, H. Schrey, and F. Wendland (2005), Distributed modeling of groundwater recharge at the macroscale, *Ecological Modelling*, *187*, 15–26.
- Bonan, G. (1996), A Land Surface Model (LSM version 1.0) for ecological, hydrological, and atmospheric studies: Technical description and user's guide, *Tech. Rep. NCAR/TN-417+STR*, NCAR Technical Note.
- Bouraoui, F., G. Vachaud, L. Li, H. Le Treut, and T. Chen (1999), Evaluation of the impact of climate changes on water storage and groundwater recharge at the watershed scale, *Climate Dynamics*, *15*(2), 153–161.
- Boutt, D., D. W. Hyndman, B. Pijanowski, and D. Long (2001), Identifying potential land use-derived solute sources to stream baseflow using ground water models and GIS, *Ground Water*, *39*, 24–34.
- Bras, R., and I. Rodriguez-Iturbe (1985), *Random Functions and Hydrology*, New York: Dover Publications.

Brunt, D. (1952), *Physical and Dynamical Meteorology*, Cambridge University Press: Cambridge.

Burnett, A., M. Kirby, H. Mullins, and W. Patterson (2003), Increasing Great Lake-effect snowfall during the twentieth century: A regional response to global warming?, *Journal of Climate*, 16, 3535–3542.

Cengel, Y., and M. Boles (2001), *Thermodynamics: An Engineering Approach*, McGraw-Hill.

Center, N. C. D. (2001), NESDIS, NOAA, U.S. Department of Commerce, U.S. daily surface data (DS-3200 and DS-3210), *Tech. rep.*, National Climatic Data Center, Asheville, NC.

Chen, J., X. Chen, W. Ju, and X. Geng (2005a), Distributed hydrological model for mapping evapotranspiration using remote sensing inputs, *Journal of Hydrology*, 305, 15–39.

Chen, J., C. Menges, W. Ju, and X. Geng (2005b), Global mapping of foliage clumping index using multi-angular satellite data, *Remote Sensing of Environment*, 97, 447–457.

Chiew, F., T. McMahon, and I. O'Neill (1992), Estimating groundwater recharge using an integrated surface and groundwater modelling approach, *Journal of Hydrology*, 131, 151–186.

Choudhury, B., and J. Monteith (1988), A four-layer model for the heat budget of homogeneous land surfaces, *Quarterly Journal of the Royal Meteorological Society*, 114, 373–398.

Christensen, J., et al. (2007), *Climate Change 2007: The Physical Sciences Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chap. Regional Climate Projections, Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA.

CODATA (2002), CODATA internationally recommended values for the fundamental physical constants.

Cooley, J., and J. Tukey (1965), An algorithm for the machine calculation of complex fourier series, *Mathematics of Computation*, 19, 297–301.

Coulibaly, P., and D. Burn (2004), Wavelet analysis of variability in annual Canadian streamflows, *Water Resources Research*, 40(3), W03,105.

- Crawford, N., and R. Linsley (1966), *Digital simulation in hydrology: Stanford Watershed Model IV*, Stanford University Department of Civil Engineering, Tech Report 39, Palo Alto, CA.
- Dahlstedt, K., and H. Jensen (2004), Fluctuation spectrum and size scaling of river flow and level, *Physica A*, *348*, 596–610.
- Dai, A., K. T.R., B. Sun, and K. Trenberth (2005), Recent trends in cloudiness over the United States: A tale of monitoring inadequacies, *Bulletin of the American Meteorological Society*, *87*, 597–606.
- Dickinson, R., A. Henderson-Sellers, C. Rosenzweig, and P. Sellers (1991), Evapotranspiration models with canopy resistance for use in climate models, a review, *Agriculture and Forest Meteorology*, *54*, 373–388.
- Dickinson, R., A. Henderson-Sellers, and P. Kennedy (1993), Biosphere-Atmosphere Transfer Scheme (BATS) version 1e as coupled to the NCAR Community Climate Model, *Tech. rep.*, NCAR Technical Note NCAR/TN-387+STR, Boulder, CO.
- Downer, C., and F. Ogden (2003), GSSHA user's manual: Gridded surface subsurface hydrologic analysis, version 1.43 for WMS 6.1, *Tech. rep.*, U.S. Army Corps of Engineering, Engineering Research and Development Center.
- Eagleson, P. (1982), Climate, soil and vegetation, 5: A derived distribution of storm surface runoff, *Water Resources Research*, *18*(2), 325–340.
- Eidenshink, J. (1992), The 1990 conterminous U.S. AVHRR data set, *Photogrammetric Engineering and Remote Sensing*, *58*, 809–813.
- Elliot, D., C. Holladay, W. Barchet, F. H.P., and W. Sandusky (1986), Wind energy resources atlas of the United States, *Tech. rep.*, Pacific Northwest Laboratory: Richland, WA.
- Entekhabi, D., and P. Eagleson (1989), Land surface hydrology parameterization for the atmospheric general circulation models including subgrid scale spatial variability, *Journal of Climate*, *2*, 816–831.
- ESRI (2003), ArcInfo workstation, *Tech. rep.*, Environmental Research Systems Institute: Redlands, CA.
- Farrand, W., and D. Bell (1982), Quaternary geology of southern Michigan, The University of Michigan.
- Fleming, S., A. Lavenue, A. Aly, and A. Adams (2002), Practical applications of spectral analysis to hydrologic time series, *Hydrological Processes*, *16*, 565–574.

- Flerchinger, G. (2000), *The Simultaneous Heat and Water (SHAW) Model: Technical Documentation*, Northwest Watershed Research Center, USDA Agricultural Research Service.
- Foley, J., I. Prentice, N. Ramankutty, S. Lewis, D. Pollard, S. Sitch, and A. Haxeltine (1996), An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics, *Global Biogeochemical Cycles*, 10(4), 603–628.
- Frigo, M., and S. Johnson (1998), FFTW: An adaptive software architecture for the FFT, in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1381–1384.
- Fulton, R., J. Breidenbach, D.-J. Seo, and D. Miller (1998), WSR-88D rainfall algorithm, *Weather Forecasting*, 13, 377–395.
- Gale, M., and D. Grigal (1987), Vertical root distributions of northern tree species in relation to successional status, *Canadian Journal of Forest Research*, 17, 829–834.
- Gaucherel, C. (2002), Use of wavelet transform for temporal characterization of remote watersheds, *Journal of Hydrology*, 269, 101–121.
- Gedney, N., P. Cox, R. Betts, O. Boucher, C. Huntingford, and P. Stott (2006), Detection of a direct carbon dioxide effect in continental river runoff records, *Nature*, 439(7078), 835–838.
- Gesch, D., M. Oimoen, S. Greenlee, C. Nelson, M. Steuck, and D. Tyler (2002), The National Elevation Dataset, *Journal of the American Society for Photogrammetry and Remote Sensing*, 68(1), 5–11.
- Goff, J., and S. Gratch (1946), Low-pressure properties of water from -160 to 212f, *Transactions of the American Society of Heating and Ventilating Engineers*, 52, 95–122.
- Halliday, D., R. Resnick, and J. Walker (2004), *Fundamentals of Physics*, Wiley.
- Hameed, S. (1984), Fourier analysis of Nile flood levels, *Geophysical Research Letters*, 1(9), 843–845.
- Harbaugh, A., E. Banta, M. Hill, and M. McDonald (2000), MODFLOW-2000, the u.s. geological survey modular ground-water model- user guide to modularization concepts and the ground water flow process, *Tech. rep.*, Open-File Report 00-92, USGS: Reston, VA.

- Harris, F. (1978), On the use of windows for harmonic analysis with the Discrete Fourier Transform, *Proceedings of the IEEE*, 66, 67–67.
- Harrison, L. (1963), *Humidity and moisture*, Vol 3, chap. Fundamental concepts and definitions relating to humidity, Reinhold Publishing Co., NY.
- HDF (2008), *HDF5 User's Guide*, The HDF Group.
- Healy, R., and P. Cook (2002), Using groundwater levels to estimate recharge, *Hydrogeology Journal*, 10(1), 91.
- Hillel, D. (1980), *Introduction to Soil Physics*, Academic Press Inc.: San Diego, CA.
- Holtschlag, D. (1996), A generalized estimate of ground-water-recharge rates in the lower peninsula of Michigan, *Tech. rep.*, USGS Water-Supply Paper 2437.
- Hostetler, S., and P. Bartlein (1990), Simulation of lake evaporation with application to modeling lake level variations of Harney-Malheur lake, Oregon, *Water Resources Research*, 26(10), 2603–2612.
- Hurrell, J., Y. Kushnir, M. Visbeck, and G. Ottersen (2003), *The North Atlantic Oscillation: Climate Significance and Environmental Impact*, chap. An overview of the North Atlantic Oscillation, pp. 1–35, AGU Geophysical Monograph Series 134.
- HydroGeoLogic (2002), MODHMS–MODFLOW-based hydrologic modeling system: Documentation and user's guide, *Tech. rep.*, HydroGeoLogic Inc., Herndon, VA.
- Hyndman, D. W., A. D. Kendall, and N. R. H. Welty (2007), *Subsurface Hydrology: Data Integration for Properties and Processes*, chap. Evaluating Temporal and Spatial Variations in Recharge and Streamflow Using the Integrated Landscape Hydrology Model (ILHM), pp. 121–141, AGU Geophysical Monograph Series 171.
- IPCC (2001), *Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, chap. IPCC 2001, Climate Change 2001: The Scientific Basis, Cambridge University Press: Cambridge.
- Jackson, R., J. Canadell, and J. Ehleringer (1996), A global analysis of root distributions for terrestrial biomes, *Oecologia*, 108, 389–411.
- Jayawickreme, D., and D. W. Hyndman (2007), Evaluating the influence of land cover on seasonal water budgets using Next Generation Radar (NEXRAD) rainfall and streamflow data, *Water Resources Research*, 43(2), W02,408.

- Johanson, R., J. Kittle, and A. Donigian (1984), *Hydrological Simulation Program-FORTRAN (HSPF): User's Manual for Release 8.0*, EPA-600/3-84-066 U.S. Environmental Research Laboratory, Athens, GA.
- Jukic, D., and V. Denic-Jukic (2004), A frequency domain approach to ground-water recharge estimation in karst, *Journal of Hydrology*, 289, 95–110.
- Kendall, A. D., and D. W. Hyndman (2007), *Subsurface Hydrology: Data Integration for Properties and Processes*, chap. Examining Watershed Processes Using Spectral Analysis Methods Including the Scaled-Windowed Fourier Transform, pp. 183–200, AGU Geophysical Monograph Series 171.
- Kerby, W. (1959), Time of concentration for overland flow, *Civil Engineering*, 29(3), 60.
- Kirchner, J. (2005), Aliasing in $1/f^{\alpha}$ noise spectra: Origins, consequences, and remedies, *Physical Review*, 71(6), 066,110(16).
- Kirchner, J., X. Feng, and C. Neal (2000), Fractal stream chemistry and its implications for contaminant transport in catchments, *Nature*, 403, 524–526.
- Knyazikhin, Y., et al. (1999), MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by vegetation (FPAR) product (MOD15) algorithm, *Tech. rep.*, Theoretical Basis Document, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.
- Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel (2006), World map of the Koppen-Geiger climate classification updated, *Meteorologische Zeitschrift*, 15(3), 259–263.
- Krause, P., D. Boyle, and F. Base (2005), Comparison of different efficiency criteria for hydrological model assessment, *Advances in Geosciences*, 5, 89–87.
- Labat, D., Y. Godderis, J. L. Probst, and J. L. Guyot (2004), Evidence for global runoff increase related to climate warming, *Advances in Water Resources*, 27, 631–642.
- Lawrence, M. (2005), The relationship between relative humidity and the dew-point temperature in moist air: A simple conversion and applications, *Bulletin of the American Meteorological Society*, 86(2), 225–233.
- Leathers, D., B. Yarnal, and M. Palecki (1991), The Pacific/North American teleconnection pattern and the United States climate. part i: Regional temperature and precipitation associations, *Journal of Climate*, 4(5), 517–528.

Leavesley, G., R. Lichty, B. Troutman, and L. Saindon (), *Precipitation-Runoff Modeling System—User's manual*, US Geological Survey Water-Resources Investigations Report 83-4238.

Lee, J., and K. Lee (2000), Use of hydrologic time series data for identification of recharge mechanism in a fractured bedrock aquifer system, *Journal of Hydrology*, 229, 190–201.

Liang, X., D. Lettenmaier, E. Wood, and S. Burges (1994), A simple hydrologically based model of land surface water and energy fluxes for general circulation models, *Journal of Geophysical Research*, 99(D7), 14,415–14,428.

Manfreda, S., M. Fiorentino, and V. Iacobellis (2005), DREAM: a distributed model for runoff, evapotranspiration, and antecedent soil moisture simulation, *Advances in Geosciences*, 2, 31–39.

Mann, M. (2004), On smoothing potentially non-stationary climate time series, *Geophysical Research Letters*, 31, L07,214.

Markstrom, S. L., R. G. Niswonger, R. S. Regan, and D. E. Prudic (2008), *GSFLOW—Coupled Ground-Water and Surface-Water Flow Model Based on the Integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005)*, chap. D1, USGS Techniques and Methods 6.

Mathworks (2008), *MATLAB Computational Environment*, The Mathworks, Inc.

Maulem, Y. (1976), A new model predicting the hydraulic conductivity of unsaturated porous media, *Water Resources Research*, 12, 513–522.

MAWN (2008), Hourly climate data, <http://www.agweather.geo.msu.edu/mawn>.

Maxwell, R., and N. Miller (2005), Development of a coupled land surface and groundwater model, *Journal of Hydrometeorology*, 6, 233–247.

McCuen, R. (2004), *Hydrologic Analysis and Design*, Pearson Prentice Hall: Upper Saddle River New Jersey.

MDEQ (2008), WelLogic, statewide ground water database.

MDEQ (2009), Michigan estimated groundwater recharge.

MDEQ Bathymetry (2008), Inland lake bathymetry.

MDNR (2001), Integrated forest monitoring assessment and prescription (IFMAP), *Tech. rep.*, Michigan Department of Natural Resources: Lansing, MI.

MDNR (2009), Inland lake maps.

Monteith, J. (1965), Evaporation and environment, *Symposia for the Society of Experimental Biology*, *XIX*, 205–234.

Morenz, M., R. van Til, and C. Luukkonen (2005), Water use for irrigation in Michigan, 2001, USGS Fact Sheet FS 2005-3079.

Myeni, R., et al. (2002), Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, *Remote Sensing of Environment*, *83*, 214–231.

Naff, R., and A. Gutjahr (1983), Estimation of groundwater recharge parameters by time series analyses, *Water Resources Research*, *19*(6), 1531–1546.

Nash, J., and J. Sutcliffe (1970), River flow forecasting through conceptual models. part i: A discussion of principles, *Journal of Hydrology*, *10*, 282–290.

NASS (2007), Prospective plantings, *Tech. rep.*, National Agriculture Statistical Service, USDA.

NCDC Hourly (2008), Hourly precipitation gauge data, <http://www.ncdc.noaa.gov/>.

NCDC Temp (2008), Daily temperature data, <http://www.ncdc.noaa.gov/>.

Neitsch, S., J. Arnold, J. Kiniry, R. Srinivasan, and J. Williams (2002), *Soil and Water Assessment Tool User's Manual Version 2000*, GSWRL Report 02-02 BRC Report 02-06 Texas Water Resources Institute TR-192, College Station, TX.

Niswonger, R., D. Prudic, and R. Regan (2006), Documentation of the Unsaturated-Zone Flow (UZFI) package for modeling unsaturated flow between the land surface and the water table with MODFLOW-2005, in *U.S. Geological Techniques and Methods Book 6, Chapter A19*, p. 62.

O'Callaghan, J., and D. Mark (1984), The extraction of drainage networks from digital elevation data, *Computer Vision, Graphics, and Image Processing*, *28*, 323–344.

O'Neal, R. (1997), Muskegon River Watershed assessment, *Tech. rep.*, Michigan Department of Natural Resources, Fisheries Special Report 19.

- Panday, S., and P. Huyakorn (2004), A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow, *Advances in Water Resources*, 27(4), 361–382.
- Patterson, J., and P. Hamblin (1988), Thermal simulation of a lake with winter ice cover, *Limnology and Oceanography*, 33(3), 323–338.
- Pelletier, J., and D. Turcotte (1997), Long-range persistence in climatological and hydrological time series: Analysis, modeling and application to drought hazard assessment, *Journal of Hydrology*, 203, 198–208.
- Percival, D., and A. Walden (1993), *Spectral Analysis for Physical Applications*, Cambridge: Cambridge University Press.
- Peterson, E., and J. Hennessey, Jr. (1978), On the use of power laws for estimates of wind power potential, *Journal of Applied Meteorology*, 17, 390–394.
- Piajnowski, B., D. Ray, A. Kendall, J. Duckles, and D. Hyndman (2007), Using backcast land-use changes and groundwater travel-time models to generate land-use legacy maps for watershed management, *Ecology and Society*, 12(2), 25 [online].
- Pimentel, D., and T. Patzek (2005), Ethanol production using corn, switchgrass, and wood; biodiesel production using soybean and sunflower, *Natural Resources Research*, 14(1), 65–76.
- Press, W., S. Teukolsky, W. Vetterling, and B. Flannery (1992), *Numerical Recipes in FORTRAN 77*, Cambridge: Cambridge University Press.
- Probst, J., and Y. Tardy (1987), Long range stream-flow and world continental runoff fluctuations since the beginning of this century, *Journal of Hydrology*, 94, 289–311.
- Prudic, D., L. Konikow, and E. Banta (2004), A new Stream-Flow Routing (SFR1) package to simulate stream-aquifer interaction with MODFLOW-2000, *Tech. rep.*, Open-File Report 2004-1042, USGS: Reston, VA.
- Qu, Y., and C. Duffy (2007), A semidiscrete finite volume formulation for multiprocess watershed simulation, *Water Resources Research*, 43, W08,419.
- Querner, E. (1997), Description and application of the combined surface and groundwater flow model MOGROW, *Journal of Hydrology*, 192(1-4), 158–188.
- R Development Core Team (2005), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Rajagopalan, B., E. Cook, U. Lall, and B. Ray (2000), Spatiotemporal variability of ENSO and SST teleconnections to summer drought over the United States during the twentieth century, *Journal of Climate*, 13, 4244–4255.

Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, D.-J. Seo, and D. Participants (2004), Overall distributed model intercomparison project results, *Journal of Hydrology*, 298(1-4), 27–60.

Refsgaard, J., and B. Storm (1995), *Computer models of watershed hydrology*, chap. MIKE SHE, pp. 809–846, Water Resources Publication Co. USA.

Ritchie, J. (1972), Model for predicting evaporation from a row crop with incomplete cover, *Water Resources Research*, 8(5), 1204–1213.

Rodionov, S., and R. Assel (2000), Atmospheric teleconnection patterns and the severity of winter in the Laurentian Great Lakes basin, *Atmosphere-Ocean*, 38(4), 601–635.

Said, A., D. Stevens, and G. Sehlke (2005), Estimating water budget in a regional aquifer using HSPF-MODFLOW integrated model, *Journal of the American Water Resources Association*, 41(1), 55–66.

Saxton, K., and W. Rawls (2006), Soil water characteristic estimates by texture and organic matter for hydrologic solution, *Soil Science Society of America Journal*, 70, 1569–1578.

Scanlon, B., R. Healy, and P. Cook (2002), Choosing appropriate techniques for quantifying groundwater recharge, *Hydrogeology Journal*, 10(1), 18–39.

Schaap, M., and F. Leij (2000), Improved prediction of unsaturated hydraulic conductivity with the Maulem-van Genuchten model, *Soil Science Society of America Journal*, 64, 843–851.

Schaap, M., F. Liej, and M. van Genuchten (1998), Neural network analysis for hierarchical prediction of soil water retention and saturated hydraulic conductivity, *Soil Science Society of America Journal*, 62, 847–855.

Schaetzl, R., and D. Tomczak (2001), Wintertime temperatures in the fine-textured soils of the Saginaw Valley, Michigan, *The Great Lakes Geographer*, 8(2), 87–99.

Schulze, E., F. Kelliher, C. Körner, J. Lloyd, and R. Leuning (1994), Relationships among maximum stomatal conductance, ecosystem surface conductance, carbon assimilation rate, and plant nitrogen nutrition: a global ecology scaling exercise, *Annual Reviews of Ecological Systems*, 25, 629–660.

Seo, D. (1998), Real-time estimation of rainfall fields using radar rainfall and rain gage data, *Journal of Hydrology*, 208, 37–52.

Seo, H., J. Šimůnek, and E. P. Poeter (2007), Documentation of the HYDRUS package for MODFLOW-2000, the U.S. Geological Survey modular ground-water model, *Tech. Rep. GWMI 2007-01*, Int. Ground Water Modeling Center.

Simunek, J., M. van Genuchten, and M. Sejna (2005), *The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media*, Department of Environmental Sciences, University of California Riverside.

Smith, M., D.-J. Seo, V. Koren, S. Reed, Z. Zhang, Q. Duan, F. Moreda, and S. Cong (2004), The distributed model intercomparison project (DMIP): motivation and experiment design, *Journal of Hydrology*, 298(1-4), 4–26.

Smith, S. (1997), *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Publications.

Sophocleous, M., and S. Perkins (2000), Methodology and application of combined watershed and ground-water models in Kansas, *Journal of Hydrology*, 236, 185–201.

Stieglitz, M., D. Rind, J. Famiglietti, and C. Rosenzweig (1997), An efficient approach to modeling the topographic control of surface hydrology for regional and global climate modeling, *Journal of Climate*, 10, 118–137.

Tang, Z., B. Engel, B. Pijanowski, and K. Lim (2005), Forecasting land use change and its environmental impact at a watershed scale, *Journal of Environmental Management*, 76, 35–45.

Tarboton, D. (1997), A new method for the determination of flow directions and upslope areas in grid digital elevation models, *Water Resources Research*, 33(2), 309–319.

Tarboton, D., and C. Luce (1996), Utah Energy Balance snow accumulation and melt model (UEB), computer model technical description and user's guide, *Tech. rep.*, Utah Water Research Laboratory and USDA Forest Service Intermountain Research Station.

Tessier, Y., S. Lovejoy, P. Hubert, and D. Schertzer (1996), Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions, *Journal of Geophysical Research*, 101(D21), 26,427–26,440.

Therrien, R., R. McLaren, E. Sudicky, and S. Panday (2004), *HydroGeoSphere, A Three-Dimensional Numerical Model Describing Fully-Integrated Subsurface and Surface Flow and Solute Transport*, Groundwater Simulation Group, University of Waterloo, Ont. Canada.

Thompson, D., and J. Wallace (1998), The Arctic Oscillation signature in the wintertime geopotential height and temperature fields, *Geophysical Research Letters*, 25(9), 1297–1300.

Thoms, R., R. Johnson, and R. Healy (2006), User's guide to the Variably Saturated Flow (VSF) process for MODFLOW, *Tech. rep.*, U.S. Geological Survey Techniques and Methods 6-A18.

Torbick, N., J. Qi, G. Roloff, and R. Stevenson (2006), Investigating impacts of land-use land cover change on wetlands in the Muskegon River Watershed, Michigan, USA, *Wetlands*, 26(4), 1103–1113.

Torrence, C., and G. Compo (1997), A practical guide to wavelet analysis, *Bulletin of the American Meteorological Society*, 79, 61–78.

Trenberth, K. (1997), The definition of El Nino, *Bulletin of the American Meteorological Society*, 78(12), 2771–2777.

Tucker, C., J. Pinzon, M. Brown, D. Slayback, E. Pak, R. Mahoney, E. Vermote, and N. El Saleous (2005), An extended AVHRR 8-km NDVI data set compatible with MODIS and SPOT vegetation NDVI data, *International Journal of Remote Sensing*, 26(20), 4485–5598.

USACE-HEC (1998), *HEC-HMS Hydrologic Modeling System user's manual*, U.S. Army Corps of Engineers Hydrologic Engineering Center, Davis, CA.

USDA-NRCS (2008), Soil survey geographic (SSURGO) database.

USDA-SCS (1993), *State Soil Geographic (STATSGO) Database*, Miscellaneous Publication No. 1492, U.S. Government Printing Office, Washington, D.C.

USGS (2005), National Water Information System (NWISWeb), data available on the World Wide Web.

USGS (2008), National Water Information System (NWISWeb), data available on the World Wide Web.

van Genuchten, M. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Science Society of America Journal*, 44, 892–898.

VanderKwaak, J. (1999), Numerical simulation of flow and chemical transport in integrated surface-subsurface hydrologic systems, Ph.D. thesis, Dept. of Earth Sciences, University of Waterloo: Ontario, Canada.

VanderKwaak, J., and K. Loague (2001), Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model, *Water Resources Research*, 37(4), 999–1013.

Visbeck, M., J. Hurrell, L. Polvani, and H. Cullen (2001), The North Atlantic Oscillation, past present and future, *Proceedings of the National Academy of Sciences*, 98(23), 12,876–12,877.

Vorosmarty, C., P. Green, J. Salisbury, and R. Lammers (2000), Global water resources: vulnerability from climate change and population growth, *Science*, 298, 284–288.

Voss, C., and A. Provost (2002), A model for saturated-unsaturated variable-density ground-water flow with solute or energy transport, *Water-Resources Investigations Report 02-4231*, U.S. Geological Survey.

Walker, G., and E. Bliss (1932), World weather V, *Memorandum of the Royal Meteorological Society*, 4, 53–84.

Walko, R., and C. Tremback (2005), Modifications for the transition from LEAF-2 to LEAF-3, *Tech. rep.*, ATMET Technical Note.

Walko, R., et al. (2000), Coupled atmosphere-biophysics-hydrology models for environmental modeling, *Journal of Applied Meteorology*, 39, 931–944.

Wallace, J., and D. Gutzler (1981), Teleconnections in the geopotential height field during the northern hemisphere winter, *Monthly Weather Review*, 109(4), 784–812.

Wang, X., and A. M. Melesse (2006), Effects of STATSGO and SSURGO as inputs on SWAT model's snowmelt simulation, *Journal of the American Water Resources Association*, 42(5), 1217–1236.

Westjohn, D., and T. Weaver (1998), Hydrogeologic framework of the Michigan Basin regional aquifer system, *Tech. rep.*, US Geological Survey Professional Paper 1418.

Wigmosta, M., L. Vail, and D. Lettenmaier (1994), A distributed hydrology-vegetation model for complex terrain, *Water Resources Research*, 30(6), 1665–1679.

- Wiley, M., and P. Richards (2006), Unpublished streamflow data, cedar Creek.
- Winter, T. (1981), Uncertainties in estimating the water balance of lakes 1, *Journal of the American Water Resources Research*, 17(1), 82–115.
- Yang, Z., and D. Han (2006), Derivation of unit hydrograph using a transfer function approach, *Water Resources Research*, 42, W01,501.
- Yang, Z.-L., R. Dickinson, A. Robock, and K. Vinnikov (1997), Validation of the snow sub-model of the Biosphere-Atmosphere Transfer Scheme with Russian snow cover and meteorological observational data, *Journal of Climate*, 10, 353–373.
- Yeh, G., and G. Huang (2003), A numerical model to simulate water flow in watershed systems of 1-d stream-river network, 2-d overland regime, and 3-d subsurface media (WASH123D: Version 1.5), *Technical report*, Dept. of Civil and Environmental Engineering University of Central Florida, Orlando, Florida.
- Yongjiu, D., Z. Xubin, and R. Dickinson (2001), *Common Land Model (CLM): Technical Documentation and User's Guide*, National Center for Atmospheric Research.
- Yu, Z., D. Pollard, and L. Cheng (2006), On continental-scale hydrologic simulations with a coupled hydrologic model., *Journal of Hydrology*, 331(1-2), 110–124.
- Zhang, Y., and K. Schilling (2004), Temporal scaling of hydraulic head and river base flow and its implications for groundwater recharge, *Water Resources Research*, 40(3), W03,504.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03062 5986